Fast Passivity Enforcement for Pole-Residue Models by Perturbation of Residue Matrix Eigenvalues

Bjørn Gustavsen, Senior Member, IEEE

Abstract—Rational models must be passive in order to avoid unstable time domain simulations. This paper introduces a fast approach for passivity enforcement of pole-residue models. This is achieved by perturbing the eigenvalues of the residue matrices, as opposed to the existing approach of perturbing matrix elements. This leads to large savings in computation time with only a small increase of the modeling error. This fast residue perturbation (FRP) approach is merged with the Modal Perturbation technique, leading to fast modal perturbation (FMP). Usage of FMP over FRP achieves to retain the relative accuracy of the admittance matrix eigenvalues. A complete approach is obtained by combining the passivity enforcement step with passivity assessment via the Hamiltonian matrix eigenvalues and a robust iteration scheme, giving a guaranteed passive model. Application of FMP to a six-port power transformer shows that the approach is able to remove large out-of-band passivity violations without corrupting the in-band behavior. This is shown to mitigate an unstable simulation. The approach is also demonstrated for a high-speed interconnect and a transmission line.

Index Terms—Macromodel, passivity enforcement, rational model, stability, transformer, vector fitting.

I. INTRODUCTION

The modeling of frequency-dependent components for use in electromagnetic transient studies is usually based on calculating a rational function that reproduces a given frequency domain behavior. The usage of rational functions leads to recursive convolution in the time domain and thus fast computations [1]. This procedure is routinely applied in the frequency dependent modeling of transmission lines and cables by the Method of Characteristics [2], [3]. Rational modeling is the main ingredient for the calculation of frequency-dependent network equivalents (FDNEs) with respect to a set of ports (terminals) [4], [5]. This black-box approach can also be used for the wide band modeling of transformers, starting from computed [6] or measured responses [7], [8]. The rational modeling can be easily achieved via the pole relocating algorithm known as vector fitting (VF) [9] with recent enhancements [10]–[13], or via polynomial fitting with frequency partitioning [5].

The Achilles heel of the black-box approach is the need for ensuring passivity when the model can interact with the adjacent system over its ports. Failure to comply with the passivity requirement can easily lead to an unstable simulation. Passivity enforcement can be incorporated in the fitting process using convex optimization [14], but the computation time can be excessive. A much faster approach is residue perturbation (RP) [15], [16] where passivity is enforced by perturbing the residues of a pole-residue model (or state-space model), as a post processing step. The perturbation is done so as to minimize the change to the model in the least-squares (LS) sense, at a set of frequency samples. These samples are usually those used in the fitting process (in-band). Usage of modes (modal perturbation (MP) [17]) allows to retain the relative accuracy of the admittance matrix eigenvalues, thus mitigating the problem of error magnification with arbitrary terminal conditions. A faster approach is obtained by perturbing the poles instead of the residues [18], but this leads to a more constrained problem and thus a larger model perturbation. A fast approach is also achieved by the alternative RP approach in [19], which minimizes the change to the system impulse energy. The latter approach does however not distinguish between in-band and out-of-band frequencies. This makes it difficult to remove out-of-band passivity violations without corrupting the in-band model behavior.

Although the RP and MP approaches lead to small model perturbations, they remain demanding in computation time. The computational efficiency can be greatly improved by solving the associated constrained LS problem using sparse computations. This requires a sparse implementation for quadratic programming (QP). Sparse QP solvers are only available in specialized software (e.g., CPLEX as used in [17], [20]), but they are often costly. The computation time can also be reduced by using a subset of the residues as free variables, but this significantly increases the model perturbation.

In this paper we introduce a straightforward procedure which greatly reduces computation time and memory requirements for RP and MP, with only a small increase of the model perturbation. This is achieved by taking as free variables the eigenvalues of the individual residue matrices. It is shown how to incorporate this idea in both the RP and MP approaches, leading to fast RP (FRP) and fast MP (FMP). The FRP/FMP approaches are combined with precise passivity assessment via the Hamiltonian matrix eigenvalues and a robust iterative scheme. The FMP-based approach is demonstrated for a power transformer model which has large out-of-band passivity violations, and for an interconnect model which has large in-band violations. Finally, the results by RP/FRP/MP/FMP are compared in terms of model perturbation size and computation time when applied to an FDNE model of a transmission line.
II. POLE-RESIDUE MODELING

A. Rational Fitting From Frequency Domain Data

The modeling starts from a given port-admittance matrix \( \mathbf{Y}(s) \) (1), which defines the relation between port voltages \( \mathbf{v} \) and currents \( \mathbf{i} \). This matrix can be obtained via calculations or measurements

\[
\mathbf{i}(s) = \mathbf{Y}(s) \mathbf{v}(s),
\]

(1)

It is assumed that \( \mathbf{Y}(s) \) has been fitted by a model of pole-residue form (2). This model uses a common pole set for all matrix elements of \( \mathbf{Y}(s) \)

\[
\mathbf{Y}(s) \cong \mathbf{Y}_{\text{rat}}(s) = \sum_{n=1}^{N} \frac{\mathbf{R}_m}{s - \alpha_m} + \mathbf{D} + s\mathbf{E}.
\]

(2)

Physicality of the model leads to the following requirements:
1) \( \mathbf{Y} \) is a symmetric matrix. Hence, \( \{\mathbf{R}_m\}, \mathbf{D}, \) and \( \mathbf{E} \) are symmetric.
2) \( \mathbf{D} \) and \( \mathbf{E} \) are real matrices.
3) The poles and residues are real or come in complex conjugate pairs.
4) The poles are in the left half plane.
5) The model is passive, i.e., it cannot generate power. This implies (3) [22], [15], which at infinite frequency gives (4)

\[
eig(\Re(\mathbf{Y}_{\text{rat}}(s))) = \eig(\mathbf{G}_{\text{rat}}(s)) > 0 \Rightarrow \eig(\mathbf{D}) > 0,
\]

(3)

(4)

6) The capacitance matrix \( \mathbf{E} \) has positive eigenvalues,

\[
eig(\mathbf{E}) > 0.
\]

(5)

Unfortunately, there is no efficient method available that can calculate the approximation (2) while at the same time satisfy requirements 1)–6). For instance, vector fitting can only enforce conditions 1)–4). A practical solution is therefore to enforce conditions 5)–6) by a perturbation of the model.

III. PERTURBATION

A. Fast Residue Perturbation (FRP)

Using the ideas in [15], passivity is enforced by perturbing the elements of the residue matrices \( \{\mathbf{R}_m\} \) and \( \mathbf{D} \). In addition, it is enforced that \( \mathbf{E} \) is positive definite (has positive eigenvalues). This leads to the constrained optimization problem

\[
\Delta \mathbf{Y} = \sum_{m=1}^{N} \frac{\Delta \mathbf{R}_m}{s - \alpha_m} + \Delta \mathbf{D} + s\Delta \mathbf{E} \succeq 0 \quad (6a)
\]

\[
eig \left( \Re \left( \mathbf{Y} + \sum_{m=1}^{N} \frac{\Delta \mathbf{R}_m}{s - \alpha_m} + \Delta \mathbf{D} + s\Delta \mathbf{E} \right) \right) > 0 \quad (6b)
\]

\[
eig(\mathbf{D} + \Delta \mathbf{D}) > 0 \quad (6c)
\]

\[
eig(\mathbf{E} + \Delta \mathbf{E}) > 0, \quad (6d)
\]

The first part (6a) minimizes the change to the admittance matrix elements while the second part (6b) enforces that the perturbed model meets the passivity criterion (3). The third (6c) and fourth (6d) parts enforce that \( \mathbf{D} \) and \( \mathbf{E} \) become positive definite.

Similarly as in [15], first-order perturbation is used in (6b)–(6d) for relating the perturbation of a matrix \( \mathbf{F} \) to its eigenvalues. Since the matrices \( (\mathbf{G}_{\text{rat}}, \mathbf{D}, \mathbf{E}) \) are real and symmetric, the inverse of the associated eigenvalue matrix is equal to its transpose. With \( \mathbf{v}_i \) denoting a right eigenvector of (the unperturbed) \( \mathbf{F} \), we get for the eigenvalue perturbation

\[
\Delta \lambda_i = \frac{\mathbf{v}_i^T \Delta \mathbf{F} \mathbf{v}_i}{\mathbf{v}_i^T \mathbf{v}_i}.
\]

(7)

The number of free (perturbed) variables is reduced by individually diagonalizing the residue matrices \( \{\mathbf{R}_m\} \), and the \( \mathbf{D} \) and \( \mathbf{E} \) matrix, and perturbing only their eigenvalues (8) [21]. This achieves to reduce the problem size while still having the flexibility to perturb all elements of all residue matrices. The resulting approach will be denoted fast residue perturbation (FRP) (in the case of complex residue matrices, their real and imaginary parts are diagonalized separately)

\[
\Delta \mathbf{R}_m = \mathbf{S}_m \Delta \mathbf{G}_{\text{rat}} \mathbf{S}_m^T \quad (8a)
\]

\[
\Delta \mathbf{D} = \mathbf{S}_D \Delta \mathbf{G}_{\text{rat}} \mathbf{S}_D^T \quad (8b)
\]

\[
\Delta \mathbf{E} = \mathbf{S}_E \Delta \mathbf{G}_{\text{rat}} \mathbf{S}_E^T \quad (8c)
\]

The implementation of (6)–(8) leads to the form (9) where \( \Delta \mathbf{x} \) is a vector that holds the perturbed parameters. This problem is solved using quadratic programming (QP)

\[
\min_{\Delta \mathbf{x}} \frac{1}{2} (\Delta \mathbf{x}^T \mathbf{A}_{\text{sys}}^T \mathbf{A}_{\text{sys}} \Delta \mathbf{x}) \quad (9a)
\]

\[
\mathbf{B}_{\text{sys}} \Delta \mathbf{x} < \mathbf{c}, \quad (9b)
\]

After solving (9), the corrections for \( \{\mathbf{R}_m\}, \mathbf{D} \) and \( \mathbf{E} \) are recovered by (8).

B. Fast Modal Perturbation (FMP)

In [17], it was proposed to perturb the rational model such that the eigenvalues of \( \mathbf{Y} \) are perturbed in relation to their size. Diagonalizing \( \mathbf{Y} \) gives

\[
\mathbf{Y} = \mathbf{T} \Delta \mathbf{T}^{-1}.
\]

(10)

Postmultiplying (10) with \( \mathbf{T} \) and taking first order derivatives gives for each eigenpair \( (\lambda_i, \mathbf{t}_i) \)

\[
\Delta \mathbf{Y} \mathbf{t}_i + \mathbf{Y} \Delta \mathbf{t}_i = \Delta \lambda_i \mathbf{t}_i + \lambda_i \Delta \mathbf{t}_i.
\]

(11)

Ignoring terms involving \( \Delta \mathbf{t}_i \) and replacing \( \Delta \mathbf{Y} \) with (6a) gives

\[
\left( \sum_{m=1}^{N} \frac{\Delta \mathbf{R}_m}{s - \alpha_m} + \Delta \mathbf{D} + s\Delta \mathbf{E} \right) \mathbf{t}_i(s) \simeq \Delta \lambda_i(s) \mathbf{t}_i(s) \simeq 0.
\]

(12)
The perturbation size is made inversely proportional to the eigenvalue size by using a weighting that is equal to the inverse of the given frequency sample (eigenvalue minimum), all eigenvalues as smooth functions of frequency within the vicinity of crossover frequencies is based on eigenvalues with positive real parts, since the complex eigenvalues appear in conjugate pairs.

\[
M = \begin{bmatrix}
    A - B(D + D^T)^{-1}C & B(D + D^T)^{-1}B^T \\
    -C^T(D + D^T)^{-1}C & -A^T + C^T(D + D^T)^{-1}B^T
\end{bmatrix}
\]

where \( A, B, C \) are the matrices of the state space model associated, with (2)

\[
Y(s) = C(sI - A)^{-1}B + D + sE.
\]

The expansion of (2) into (15) is straightforward as shown in [25]. This conversion gives a state-space model with complex \( A \) and \( C \). The model is next converted into a real-only model as shown in [26], since the computation of eigenvalues of a real matrix is much faster than for a complex matrix. The assessment of crossover frequencies is based on eigenvalues with positive imaginary parts, since the complex eigenvalues appear in conjugate pairs.

In reality, crossover frequencies do not exactly correspond to purely imaginary eigenvalues as a small real part will be present. In this work, an eigenvalue \( \lambda \) is deemed to be imaginary if it satisfies the criterion

\[
\frac{\text{Re}(\lambda)}{\text{Im}(\lambda)} < tol
\]

where \( tol \) is a small quantity. A fairly large value for \( tol \) is used and the obtained list of frequencies is treated as prospective crossover frequencies that are checked by assessment of the eigenvalues of \( G_{rat} \) (3). One could treat the imaginary part of all complex eigenvalues as prospective crossover frequencies, but that would lead to a less efficient approach due to the need for assessing \( G(s) \) at many frequency samples.

Negative eigenvalues of \( D \) are enforced to become slightly positive since the calculation of the matrix \((D + D^T)^{-1}\) in (14) requires a nonsingular \( D \).

C. Robust Iterations

Enforcing passivity at only a few frequencies (eigenvalue minima) will often result in that new passivity violations arise at other frequencies. This makes it necessary to repeatedly perturb the model in order to remove all violations. In order to avoid divergence, a robust iterative procedure is used, see Fig. 2. This procedure makes use of an inner loop which adds more constraints if new passivity violations are detected, without updating the model. The outer loop generates a list of frequencies \( s_2 \) where passivity is to be enforced. Only violating eigenvalues are included in the constraint (6b). The inner loop generates a list of frequencies \( s_3 \) where new, negative eigenvalue minima appear. All eigenvalues at \( s_3 \) are added to the constraint, thereby preventing the new violations from occurring. In practice, the nonlinearity of the problem will result in that new violations are frequently detected, since the shift of eigenvalue minima in Fig. 1 is not exactly vertical. In the implementation, we therefore terminate the inner loop after a fixed number of iterations. In all examples, we used a maximum of three iterations. \( \Delta D \) and \( \Delta E \) are removed from (6) as soon as they become positive definite.

In order to increase the computational efficiency, \( A_{\text{offs}} \) in (9a) is built only a single time and is not updated during the iterations. This was shown to have a negligible impact on the final result [20].

![Sample selection for passivity enforcement.](image-url)
D. Auxiliary Frequency Samples

The conditioning of (9) is improved by adding auxiliary frequency samples to the least-squares part (9a) at out-of-band frequencies that correspond to out-of-band poles [20]. Such samples are given a small weighting in the least-squares problem (e.g., 0.001). This leads to a better defined problem without significantly impairing the quality of the perturbed model at in-band frequencies.

V. EXAMPLE: DISTRIBUTION TRANSFORMER

A. Pole-Residue Modeling

In this example, we demonstrate the ability of FMP to handle large out-of-band violations without corrupting the model in-band behavior. The modeling starts from a measured admittance matrix of a two-winding transformer (six ports) [7], see Fig. 3.

An 80th-order pole-residue model (2) is identified using vector fitting [9] with relaxation [12], with a nonzero \( D \) and inverse magnitude LS weighting. The symmetry and common pole property is achieved by fitting simultaneously all elements of the upper triangle of \( Y \). The resulting approximation is shown in Fig. 4, demonstrating a highly accurate result.

B. Passivity Enforcement by FMP

The obtained model is subjected to perturbation by FMP in combination with robust iterations and passivity assessment via the Hamiltonian matrix eigenvalues (Section IV).

Fig. 5 shows the eigenvalues of \( G_{raf}(s) \), before and after passivity enforcement. It is seen that \( G_{raf}(s) \) of the original model has negative, large eigenvalues at out-of-band frequencies, implying that the model is non-passive by criterion (3). The passivity enforcement is seen to make all eigenvalues positive, thereby making the model passive.

Fig. 6 shows a close-up of the three small eigenvalues in Fig. 5, within the fitting band. It is seen that the passivation does not adversely corrupt these eigenvalues, thanks to the ability of the (F)MP approach of retaining the relative accuracy of eigenvalues.

The actual passivity assessment is done via the Hamiltonian matrix \( M \), as described in Section IV. With a six port, 80th order model, \( M \) has size 960 × 960. Fig. 7 shows the eigenvalues with positive imaginary parts, sorted by the real part divided by the

---

**Figure 2:** Robust iterations.

**Figure 3:** Distribution transformer (30 kVA).

**Figure 4:** Rational approximation (80th order).

**Figure 5:** Eigenvalues of \( G_{raf}(s) \).
imaginary part. (Only the last few eigenvalues in the list are included in the plot). It is seen that the passivity enforcement removes the presence of (nearly) purely imaginary eigenvalues. This means that the eigenvalues of $G(s)$ do not have any zero crossings, which is a consequence of them being positive at all frequencies.

Fig. 8 shows the diagnostic output from the implemented passivity enforcement routine as executed in Matlab. The bracketed values denote the iteration count of the outer ($\bar{x}$) and inner ($y$) loops of the robust iteration scheme in Fig. 2. (A maximum of three iterations was allowed for the inner loop). It is seen that a total of eight FMP calls were used, giving a total computation time of 475 sec. (1.3-GHz Pentium processor). $D$ and $E$ of the original model had negative eigenvalues but the passivity enforcement results in that all eigenvalues become positive by a small amount $tol_D = 1E-6$ and $tol_E = 1E-12$, see Table I.

### C. Comparison With Time Domain Measurement

In a laboratory test, a near step voltage was applied to terminal 4 with terminals 5 and 6 grounded. The voltage responses on terminals 1, 2, and 3 were recorded, see Fig. 9 [7]. In Fig. 10 is shown the recorded waveforms on terminals 1, 2, and 3, as well as the applied voltage on terminal 4. In the same plot is shown the simulated voltage waveforms [27] when taking the recorded voltage on terminal 4 as an ideal voltage source. The original model is seen to give an unstable simulation result, while the
Passivated model gives a stable result that agrees well with the recorded waveforms.

VI. EXAMPLE: HIGH-SPEED INTERCONNECT

In this example, we demonstrate the ability of FMP to handle large in-band passivity violations. The $2 \times 2$ admittance matrix $\mathbf{Y}$ of a 100 mm single conductor interconnect is calculated via the Enhanced Transmission Line Model [28]. The $\mathbf{Y}$-matrix is fitted by a 50th-order pole-residue model with a nonzero $\mathbf{D}$, calculated by the (relaxed) VF algorithm.

Fig. 11 shows that the eigenvalues of $\mathbf{G}_{\text{rel}}(s)$ are substantially negative, thereby requiring a quite large perturbation. (The rational fitting was highly accurate—the passivity violations were present in the data). The rational model is next subjected to passivity enforcement by FMP. As can be seen in Fig. 11, the procedure removes all passivity violations with only a moderate change to the eigenvalues where they are positive. The impact on the Hamiltonian matrix eigenvalues is shown in Fig. 12.

Several iterations were needed in order to arrive at this result. The computation time was 5.4 s for passivity checking and 7.2 s for the passivity enforcement.

VII. EXAMPLE: TRANSMISSION LINE

In this example we compare the model perturbation size and the computation time when using either RP, FRP, MP, or FMP. As in the previous examples, the perturbation step is combined with passivity checking via the Hamiltonian matrix and the robust iteration scheme (Section IV).

The terminal admittance matrix $\mathbf{Y}$ of the transmission line in Fig. 13 is computed in the frequency domain, from 10 Hz to 10 kHz. A 30th-order pole-residue model (1) is calculated for the six-port $\mathbf{Y}$ by fitting all elements simultaneously using VF.

The resulting model is nonpassive by criterion (3) as several eigenvalues of $\mathbf{G}_{\text{rel}}(s)$ are negative at out-of-band frequencies, see Fig. 14. Thus, the objective is to perturb the model such that all eigenvalues are positive, while at the same time the change to $\mathbf{Y}_{\text{rel}}(s)$ is minimal in the fitting range (10 Hz–10 kHz).

Fig. 15 shows the change to the eigenvalues of $\mathbf{G}_{\text{rel}}(s)$ when perturbing by either RP or FRP. It can be seen that both approaches result in positive eigenvalues and thus a passive model. The perturbation within the fitting band is with both approaches quite small, despite the large correction for the out-of-band passivity violations.

Fig. 16 shows the deviation from the eigenvalues of the original model $\mathbf{Y}(s)$, in the fitting band. It can be seen that FRP gives only a slightly larger perturbation of the eigenvalues than RP. The increase is remarkably small, considering that the number of free unknowns per residue matrix has been reduced from 21 to 6. In the same plot is also shown the result by RP when perturbing only diagonal elements of the residue matrices. This is seen to cause a much larger perturbation.
The examples in this paper are all characterized by large passivity violations. They were chosen so as to put the algorithms to a real test, since most approaches can easily fix small violations without corrupting the model behavior. It is remarked that...
the severity of violations can often be substantially reduced by throwing out high frequency out-of-band poles, followed by a refitting of the residues [29]. Unfortunately, this approach will often impair the model accuracy within the fitting band. It has also been proposed to reduce the need for passivity corrections by enforcing asymptotic passivity during the fitting process [15], [29].

IX. CONCLUSIONS

This paper has introduced an improvement to the existing RP and MP approaches for passivity enforcement of pole-residue models. By taking the residue matrix eigenvalues as free variables, a significant reduction is obtained for the computation time and memory requirements. This is achieved with only a small increase of the model perturbation. The FRP/FMP approaches are combined with a robust iteration scheme and passivity checking via the Hamiltonian matrix eigenvalues, giving a reliable approach that produces a guaranteed passive model. Calculated results for a distribution transformer demonstrate that the approach can handle quite large out-of-band violations, without corrupting the model in-band behavior.

ACKNOWLEDGMENT

The author thanks Dr. L. De Tommasi (University of Antwerp, Belgium) for providing the interconnect data case. Useful discussions with C. Heitz (Zürcher Hochschule, Switzerland) and M. Tiberg (ABB, Switzerland) are much appreciated.

REFERENCES


Bjarne Gustavsen (M’94–SM’03) was born in Harstad, Norway in 1965. He received the M.Sc. degree in 1989 and the Dr.Ing. degree in 1993, both from the Norwegian Institute of Technology (NTN), Trondheim.

Since 1994, he has been with SINTEF Energy Research, Trondheim. His interests include simulation of electromagnetic transients and modeling of frequency-dependent effects. He spent 1996 as a Visiting Researcher at the University of Toronto, Toronto, ON, Canada, and the summer of 1998 at the Manitoba HVDC Research Centre, Winnipeg, MB, Canada. He was a Marie Curie Fellow at the University of Stuttgart, Germany, during August 2001–August 2002.