Postprint of: Fotyga G., Czarniewska M., Lamęcki A., Mrozowski M.: Reliable Greedy Multipoint Model-Order Reduction Techniques for Finite-Element Analysis, IEEE Antennas and Wireless Propagation Letters, Vol. 17, iss. 5 (2018), pp. 821-824, DOI: 10.1109/LAWP.2018.2817391

Reliable Greedy Multipoint Model-Order Reduction Techniques for Finite Element Analysis

Grzegorz Fotyga, Martyna Czarniewska, Adam Lamecki, *Senior Member, IEEE* Michał Mrozowski, *Fellow, IEEE*

***Abstract*—A new greedy multipoint model-order reduction algorithm for fast frequency-domain finite element method simulations of electromagnetic problems is proposed. The location of the expansion points and the size of the projection basis are determined based on a rigorous error estimator. Compared to previous multipoint methods, the quality of the error estimator is significantly improved by ensuring the orthogonality of the projection basis vectors at each stage of the model-order reduction algorithm. Numerical studies show that the new algorithm yields compact and highly accurate reduced-order models.**

***Index Terms*—model-order reduction, a posteriori error estimator, finite element method.**

I. Introduction

In recent years, extensive research has aimed to increase the efficiency of Finite Element Method (FEM) frequency-domain simulations. One approach to speeding up frequency sweeps involves projecting the original system of equations onto a low-order space, thus producing so-called reduced-order models (ROMs). Such models approximate the frequency response of the original system, but at lower numerical cost.

To create a ROM, a set of field solutions (snapshots) at many frequency points is collected, a singular-value decomposition (SVD) is applied to remove the redundancy from the set of vectors, and finally the Galerkin projection onto a low-order subspace is used [1], [2]. This approach may still be time consuming, as it requires many factorizations of a large system matrix. This is avoided in techniques that make use of the moment-matching property of Krylov subspace methods [3]–[5], but at the cost of larger projection bases, which becomes a problem when the frequency band grows. To address this issue, multipoint model-order reduction approaches have been proposed in [6]–[8], where the moments of the transfer functions are matched at many expansion (frequency) points. In automated greedy multipoint model-order reduction (GM-MOR) [8], the subsequent expansion points, as well as the number of block moments in each of the points, are selected automatically based on an *a posteriori* error estimator [9]. However, in this case, the subsequent block moments added to the basis are not orthogonalized with respect to the previous ones. This leads to a severe loss of orthogonality of the vectors, significantly deteriorating the quality of the error estimator. In effect, high-accuracy reduced-order models cannot be generated and the resulting projection basis is unnecessarily large. Moreover, input parameters, such as the number of frequency sub-ranges, as well as the maximum number of block moments in each of the sub-ranges have to be selected arbitrarily, which makes the GM-MOR approach non-automated.

This problem is addressed in the present paper. A fully automated reliable greedy multipoint model-order reduction (RGM-MOR) approach is presented. The subsequent block moments added to the projection basis are orthogonalized by means of the Modified Gram-Schmidt method and SVD. Moreover, the size of a projection basis is kept as small as possible, thanks to a compression technique. Although the orthogonalization and compression process slightly increases the computational time, a significant improvement in performance can be observed in projection basis size, accuracy of the error estimator, and reliability of the whole reduction process, compared to the GM-MOR technique. The improved quality of the error estimator has a significant effect on the size and accuracy of the reduced order models. Compared to GM-MOR, the new algorithm yields more compact and more accurate models, allowing wideband reduced-order models to be constructed automatically with an accuracy a few orders of magnitude higher than was previously possible.

II. Theory

The N-dimensional Finite Element discretization of a Helmholtz equation for a dielectric-loaded, lossy structure Ω excited through P ports with M_i modes at i-th port results in the following second-order input–output system of equations:

$$(\Gamma + sG + s^2C)E(s) = sBI, \quad U = B^T E(s), \tag{1}$$

where $\Gamma, G, C \in \mathbb{C}^{N\times N}$ are system matrices, $s = j\omega/c$, I, and U are the vectors of amplitude of the normalized currents and voltages, respectively, $B \in \mathbb{C}^{N\times M}$ denotes a normalized port selection matrix, $E(s) \in \mathbb{C}^{N\times M}$ is a matrix of unknown FE coefficients, and M is the total number of excitation modes. The reduced-order model, which approximates the properties of the FEM system (1) can be obtained by means of one of the standard MOR approaches [1], [5], [10]:

$$(\Gamma_r + sG_r + s^2C_r)E_r(s) = sB_rI, \quad \hat{U} = B_r^T E_r(s), \tag{2}$$

where $\Gamma_r = Q^T\Gamma Q$, $G_r = Q^TGQ$, $C_r = Q^TCQ$, $B_r = Q^TB$ are reduced system matrices, $Q \in \mathbb{C}^{N\times qM}$ is the orthonormal projection matrix, and q denotes the reduced order, where $qM \ll N$. If the reduced-order model is to be accurate over a wide frequency band and is to be created in an automated

The authors are with the Department of Electronics, Telecommunication, and Informatics, Gdańsk University of Technology: grzfotyg@pg.gda.pl, marczar2@pg.edu.pl, adam.lamecki@ieee.org, m.mrozowski@ieee.org. This work was supported by the TEAM-TECH project: EDISOn—Electromagnetic Design of flexIble SensOrs—operated within the Foundation for Polish Science TEAM-TECH Programme, co-financed by the European Regional Development Fund, Smart Growth Operational Programme 2014–2020.

and computationally efficient way, a greedy multipoint model-order reduction technique (GM-MOR) [8] is a suitable approach for generating a projection basis Q.

In GM-MOR, the following input parameters need to be specified: the lower and upper frequency limits f_{min}, f_{max}, the target error tolerance tol, the number of frequency subranges N_S, the subrange size $\Delta = (f_{\text{max}} - f_{\text{min}})/N_S$, the maximum number of block moments at a single frequency $N_{\text{max}}^{\text{moments}}$, and the maximum number of block moments in Q, q_{max}.

GM-MOR begins by generating the vectors of the projection basis Q at the initial frequency point $f = (f_{\text{max}} + f_{\text{min}})/2$. Subsequent block moments of the projection basis are generated by means of SAPOR [5]. This process is performed until the number of block moments reaches $N_{\text{max}}^{\text{moments}}$ or the maximum value of the estimated error reaches the value tol in the frequency subrange $[f - \Delta/2, f + \Delta/2]$, where the error estimator is defined as follows:

$$E_S(s) = max_{i,l}\{\eta_i[2s(b_l^i)^T b_l^i - (b_l^i)^T \Gamma Q E_r - s^2(b_l^i)^T C Q E_r - s(b_l^i)^T b_l^i (b_l^i)^T Q E_r - s(b_l^i)^T G Q E_r]/|2s\eta_i(b_l^i) b_l^i|\}, \quad (3)$$

where b_l^i is the l-th mode at the i-th port, η_i is the impedance at the i-th port (see [8] for details), and E_r is obtained by means of (2). Note that all computations in (3) are performed on low-order matrices, so the error is estimated rapidly.

Algorithm 1: RGM-MOR - the main loop

Require: f_{min}, f_{max}, tol, q_{max}, C, G, Γ and B
1: Set: $f = (f_{\text{min}} + f_{\text{max}})/2$, $\Delta = f_{\text{max}} - f_{\text{min}}$, $Q_G = []$, $j = 1$, $q = 0$, $N_{\text{max}}^{\text{moments}}$
2: **while** $E_{\text{max}} > tol$ AND no. columns of $Q_G < q_{\text{max}}$ **do**
3: $[Q_G, q, \cdot] = \text{SPARAMSAPOR}(q, f, \Delta, Q_G, tol, N_{\text{max}}^{\text{moments}}, \cdot)$
4: SVD(Q_G);
5: BASISCOMPRESSION(Q_G);
6: UPDATEMATRICES($Q_G, \cdot$)
7: $[E_{\text{max}}, f_{E_{\text{max}}}] = \text{ESTERROR}(f_{\text{min}}, f_{\text{max}}, \cdot)$;
8: Set the new expansion point $f = f_{E_{\text{max}}}$, $j = j + 1$
9: **if** $j == 1$ **then** $\Delta = \text{FINDSUBRANGESIZE}(tol, \cdot)$
10: **end while**
11: **return** C_r, G_r, Γ_r and B_r

Next, the reduction error is estimated for the whole frequency bandwidth by means of the error estimator (3). If the maximum value of the estimated error (E_{max}) is below tol, or if the number of block moments in Q_G reaches q_{max}, the reduction procedure halts. Otherwise, the frequency point at which the estimated error has its maximum value is chosen as a next expansion point for the projection basis generation. Finally, the redundant vectors are removed from the projection basis by means of singular value decomposition (SVD).

The main shortcoming of the GM-MOR algorithm is that the subsequent block moments added to the basis are not orthogonalized with respect to the previous blocks, computed in previous expansion points. This neglect leads almost always to a severe loss of orthogonality of the basis Q, especially when the high precision of the reduced model is required. In effect, the system of equations (2) is ill-conditioned and, because of this, the quality of the error estimator deteriorates and its prediction starts to differ significantly from the actual error—the error estimator predicts an error much larger than the actual value of the error. Since the automation of the GM-MOR relies on the estimated error, too high values of the error estimator result in an unnecessarily large projection basis being generated by GM-MOR. What is more, the values of the input parameters: N_S and $N_{\text{max}}^{\text{moments}}$ have to be selected for each case separately, based on the frequency bandwidth of interest and properties of the analyzed structure. As can be seen in [8], this selection has a significant impact on GM-MOR performance in terms of the reduction runtime and the reduced-model size. For wide-band frequency analysis of lossy structures it is rather recommended to perform a reduction process with a few block moments computed in each of the many frequency sub-ranges (high value of N_S and low value of $N_{\text{max}}^{\text{moments}}$), whereas for the narrow-band analysis rather a single-point reduction should be utilized ($N_S = 1$ and a high value of $N_{\text{max}}^{\text{moments}}$). However, the process of selecting N_S and $N_{\text{max}}^{\text{moments}}$ is not automated. Therefore, the GM-MOR should be regarded as an expert-tool, rather than a fully-automatic black-box reduction algorithm.

The above shortcomings of GM-MOR are addressed in the next subsection, resulting in the novel Reliable Greedy Multipoint Model-Order Reduction Algorithm (RGM-MOR).

Algorithm 2: SPARAMSAPOR: Single-point block SAPOR

Require: q, f, Δ, Q_G, tol, $N_{\text{max}}^{\text{moments}}$, B^TB, B^TCQ_G, B^TGQ_G, $B^T\Gamma Q_G$
1: For $s_0 = j2\pi f$ compute Q_1, P_1.
2: $q \leftarrow q + 1$, $Q_{G,q} = Q_1$
3: MODIFIEDGRAMSCHMIDT($Q_{G,q}$)
4: UPDATEMATRICES($Q_{G,q}, \cdot$)
5: $[E_{\text{max}}^{\text{loc}}, \cdot] = \text{ESTERROR}(f - \Delta/2, f + \Delta/2, \cdot)$
6: Set $i = 1$
7: **while** $E_{\text{max}}^{\text{loc}} > tol$ AND $i < N_{\text{max}}^{\text{moments}}$ **do**
8: For $s_0 = j2\pi f$ compute Q_{i+1}, P_{i+1}.
9: $q \leftarrow q + 1$, $Q_{G,q} = Q_{i+1}$
10: MODIFIEDGRAMSCHMIDT($Q_{G,q}$)
11: UPDATEMATRICES($Q_{G,q}, \cdot$)
12: $[E_{\text{max}}^{\text{loc}}, \cdot] = \text{ESTERROR}(f - \Delta/2, f + \Delta/2, \cdot)$
13: Set $i \leftarrow i + 1$
14: **end while**
15: **return** Q_G, q, B^TCQ_G, B^TGQ_G and $B^T\Gamma Q_G$

A. *Reliable Greedy Multipoint Model-Order Reduction*

The main steps involved in the Reliable Greedy Multipoint Model-Order Reduction (RGM-MOR) are summarized in the pseudocode presented in Algorithm 1 (the main loop) and Algorithm 2 (the SPARAMSAPOR procedure). Compared to GM-MOR [8] the new algorithm involves:

1) A global orthogonal projection basis: which is denoted as Q_G. Individual block moments ($Q_{G,q}$) are generated by means of the SAPOR approach executed for the specified expansion points (f). The global basis allows one to perform local orthogonalization.

2) Local orthogonalization: As opposed to GM-MOR, each block moment $Q_{G,q}$ is orthogonalized (steps 3 and 10 in Algorithm 2) with respect to the previous blocks $Q_{G,1} \ldots Q_{G,q-1}$ by means of the Modified Gram–Schmidt procedure [11]. The orthogonalization steps are followed by the UPDATEMATRICES procedure (steps 4 and 11), in which the matrices used in (2) and (3)—B^TCQ_G, B^TGQ_G, $B^T\Gamma Q_G$, G_r, C_r, Γ_r, B_r—are updated to account for the addition of the subsequent block moment $Q_{G,q}$. Next, in steps 5 and 12, the error is estimated in the frequency subrange $(f - \Delta/2, f + \Delta/2)$. Since the projection basis Q_G is orthonormal, the system of equations (2) is well-conditioned (unlike in GM-MOR) and the estimated error is well correlated with the actual error.

3) Global orthogonalization and basis compression: Since the global basis Q_G is composed of a few bases obtained in different expansion points, it may become very large and may contain vectors that are not needed to span the solution. In order to keep the size of the projection basis as small as possible and preserve the reduced-model from being ill-conditioned, we propose to orthogonalize and compress the basis in each iteration of the main loop, prior to the global error estimation process. In step 4 of the main loop, singular value decomposition (SVD) is performed on the projection basis (Q_G) and only the vectors, which correspond to the singular values greater than 10^{-12} are retained. In effect, the orthogonality of Q_G is close to the machine-precision level.

Next (step 5), the basis is compressed by means of Proper Orthogonal Decomposition of the reduced order model. To this end we construct the matrix, which contains all solution vectors (snapshots) of the reduced model in $s_1, s_2, \ldots, s_K$ and perform SVD, in order to remove the redundancy from the basis:

$$W_R = \mathrm{SVD}([E_r(s_1), E_r(s_2), \ldots, E_r(s_K)]). \quad (4)$$

The compressed basis is obtained as follows:

$$Q_G := Q_G W_R. \quad (5)$$

The global orthogonalization and basis compression process are performed prior to the global error estimator evaluation. Although this change slightly increases the computational time, as the projection of the FE matrices has to be performed anew (Algorithm 1, step 6), it guarantees that equation (2) used in the error estimator (step 7) is well-conditioned. As a result, the error estimator follows the actual error more accurately than in the case of GM-MOR. This well-conditioned error estimator is used both as a halting criterion and as an indicator for selecting the next expansion point for generating the projection basis.

4) Automatic selection of the width of the frequency subrange: In the first iteration of the main loop, when the number of block moments in Q_G reaches $N_{\max}^{\text{moments}}$, the error is estimated in the whole frequency band ($\Delta = f_{\max} - f_{\min}$). Next, the value of Δ for the subsequent expansion points is selected in step 9. It is set as a width of a frequency range around the expansion point, in which the error estimator is below the tol value. As the effect, the N_S parameter (used in GM-MOR) is no longer needed and the Δ parameter is chosen automatically.

III. Numerical results

In order to validate the accuracy and efficiency of the proposed reduction technique (RGM-MOR) with respect to GM-MOR [8] we have considered two numerical tests implemented in MATLAB. The first test deals with the dielectric resonant antenna (DRA) structure, whose geometry is described in Fig. 1a). FEM discretization resulted in a system of equations with 124,860 variables. In order to validate the multipoint reduction scheme, a wide frequency-band analysis from 2 to 10 GHz has been employed. The target accuracy has been set to $tol = 10^{-10}$, whereas the frequency subrange size has been determined automatically: $\Delta = 1.14$ GHz. Figure 2 shows a comparison of the maximum values of the actual and estimated S_{11} errors as a function of the reduced order q. In RGM-MOR, the estimated error is well correlated with the actual error, and it can thus be effectively used as a halting criterion. Based on the error estimate computed with a new algorithm (red curve: maximum estimated error), it is seen that for $tol = 10^{-10}$, the RGM-MOR process halts for $q = 152$, which took 947 s, with the RAM memory usage: 12.5 GB. On the other hand the maximum value of the error estimator in GM-MOR (black curve: maximum estimated error) never reaches the target accuracy ($tol = 10^{-10}$), due to the loss of orthogonality of Q_G. This means that, in this case, GM-MOR is unable to produce the reduced model with the error tolerance set to 10^{-10}. To assess the quality of the basis generated by each algorithm, we may also look at the actual error (this corresponds to a hypothetical situation in which we have access to the actual error—which in reality is not known a priori—and can use it to determine the size of the basis that causes the actual error to drop below a tol level). It can be seen that RGM-MOR generates much more compact reduced order model than GM-MOR—152 vectors suffice to obtain the actual error at the tol level with RGM-MOR, while 225

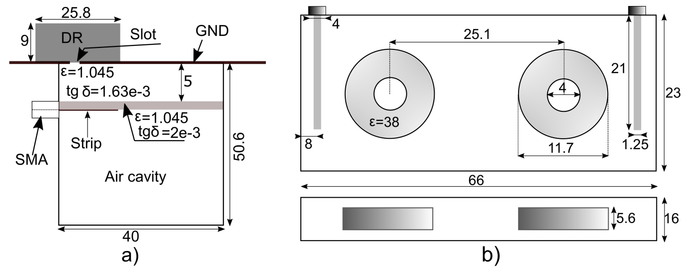

Fig. 1: a) Dielectric resonator antenna [12], b) Bandstop resonator filter [8]

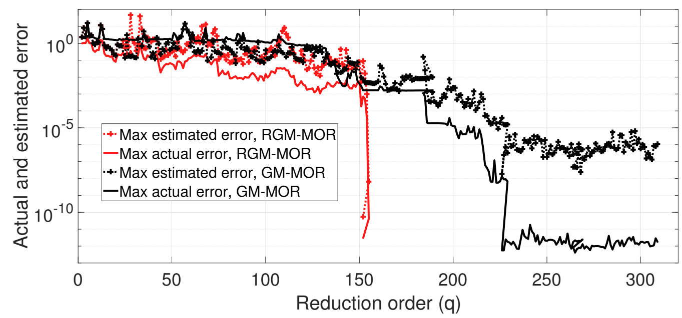

Fig. 2: Comparison of the maximum value of the actual and estimated S_{11} errors as a function of q for the DRA structure. Negative slope of the curves occurring when the error drops is due to basis compression (step 5 in Algorithm 1).

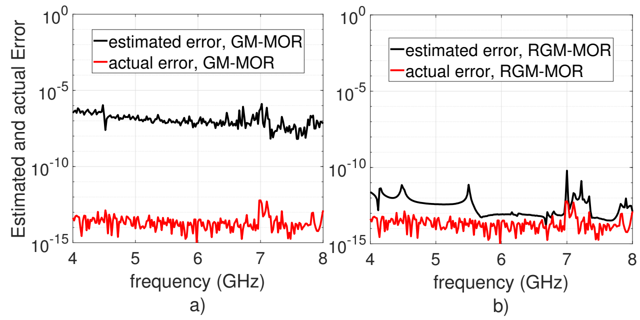

Fig. 3: Actual and estimated errors obtained by GM-MOR and RGM-MOR for $q = 42$ (filter case)

vectors are needed with GM-MOR.

The second test considers a bandstop dielectric resonator filter Fig. 1b). The scattering parameters (S_{11} and S_{21}) were computed at 400 frequency points in the range 4–8 GHz using an FEM system of equations with 101,264 variables. The error tolerance was set to: $tol = 10^{-10}$, whereas the frequency subrange size has been determined automatically: $\Delta = 1.33$ GHz. The RGM-MOR reduction process halts once the error estimator reaches the tol value ($q = 33$), which took 70 s, with the RAM memory usage: 3.75 GB. With GM-MOR, the error estimator does not reach the tol value, even for $q = 90$, due to the loss of orthogonality of the basis Q_G. As in the first test, we can also look at the size of the basis needed to create the reduced order model by considering the number of vectors that cause the actual error to drop below threshold. For RGM-MOR, only $q = 27$ is needed to obtain the actual error at the tol level, whereas $q = 32$ is needed with GM-MOR. This again confirms that more compact models are obtained with the proposed algorithm.

Figure 3 shows the actual errors obtained by GM-MOR and RGM-MOR, with the estimated errors for $q = 42$. In both cases the actual error is almost at the machine precision level, however in the first case the maximum value of the estimated error is above 10^{-6}, while in the second case the estimated error is well correlated with its actual value.

In order to investigate the performance of GM-MOR and RGM-MOR in terms of the computational time, we have analyzed the same two structures, however the tol level has been set to 10^{-4} (which is in range of the GM-MOR algorithm). In the first case (DRA antenna) the reduction process takes 679 s and 689 s, for GM-MOR and RGM-MOR, respectively, whereas for the bandstop filter it takes 49 s and 39 s. It can be seen that for the high level of the tol parameter, both approaches are comparable.

Finally, we have investigated the orthogonality of the projection basis obtained using GM-MOR and RGM-MOR. Figure 4 illustrates the measure $\|(Q_G^T \cdot Q_G - I)\|_2$ for each of the reduction orders q, where $\|\cdot\|_2$ denotes the 2-norm of a matrix. For the orthonormal basis, this measure should be at the machine precision level: ($\epsilon = 2.2 \times 10^{-16}$ in MATLAB). It is clear that, after the initial phase, the projection basis obtained using GM-MOR completely loses orthogonality, whereas with RGM-MOR, the projection basis remains orthogonal even for high values of q.

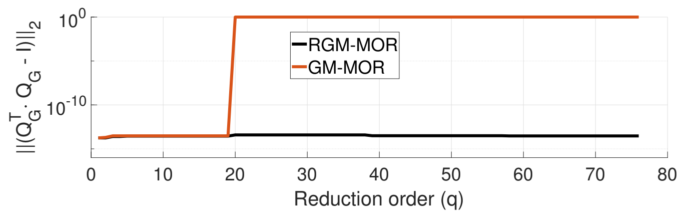

Fig. 4: Orthogonality of the projection basis obtained using GM-MOR and RGM-MOR

IV. Conclusion

This paper has proposed a fully automated reliable greedy multipoint model-order reduction (RGM-MOR) approach based on a moment-matching method applied at many expansion points in the frequency band of interest. The number of frequency sub-ranges is selected automatically. Subsequent block moments added to the projection basis are enforced to be orthogonal and compressed and this significantly improves the quality of the error estimator and in effect the proposed RGM-MOR process yields compact and highly accurate ROMs.

References

[1] V. De La Rubia, U. Razafison, and Y. Maday, “Reliable fast frequency sweep for microwave devices via the reduced-basis method,” *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 12, pp. 2923–2937, 2009.

[2] M. Hess and P. Benner, “Fast Evaluation of Time Harmonic Maxwell's Equations Using the Reduced Basis Method,” *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 6, pp. 2265–2274, June 2013.

[3] Y. Konkel, O. Farle, A. Sommer, S. Burgard, and R. Dyczij-Edlinger, “A posteriori error bounds for Krylov-based fast frequency sweeps of finite-element systems,” *IEEE Trans. Magn.*, vol. 50, no. 2, pp. 441–444, 2014.

[4] M. K. Sampath, A. Dounavis, and R. Khazaka, “Parameterized model order reduction techniques for FEM based full wave analysis,” *IEEE Trans. Adv. Packag.*, vol. 32, no. 1, pp. 2–12, 2009.

[5] Y. Su, J. Wang, X. Zeng, Z. Bai, C. Chiang, and D. Zhou, “SAPOR: second-order Arnoldi method for passive order reduction of RCS circuits,” in *Proceedings of the 2004 IEEE-ACM International conference on Computer-aided design*. IEEE Computer Society, 2004, pp. 74–79.

[6] L. Feng, J. G. Korvink, and P. Benner, “A fully adaptive scheme for model order reduction based on moment matching,” *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 5, no. 12, pp. 1872–1884, 2015.

[7] T.-S. Nguyen, T. Le Duc, T.-S. Tran, J.-M. Guichon, O. Chadebec, and G. Meunier, “Adaptive Multipoint Model Order Reduction Scheme for Large-Scale Inductive PEEC Circuits,” *IEEE Trans. Electromagn. Compat.*, vol. 59, no. 4, pp. 1143–1151, 2017.

[8] M. Rewieński, A. Lamecki, and M. Mrozowski, “Greedy Multipoint Model-Order Reduction Technique for Fast Computation of Scattering Parameters of Electromagnetic Systems,” *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 6, pp. 1681–1693, 2016.

[9] M. Rewienski, A. Lamecki, and M. Mrozowski, “A Goal-Oriented Error Estimator for Reduced Basis Method Modeling of Microwave Devices,” *IEEE Microw. Compon. Lett.*, vol. 25, no. 4, pp. 208–210, 2015.

[10] W. Wang, G. N. Paraschos, and M. N. Vouvakis, “Fast frequency sweep of FEM models via the balanced truncation proper orthogonal decomposition,” *IEEE Trans. Antennas Propag.*, vol. 59, no. 11, pp. 4142–4154, 2011.

[11] L. Giraud, J. Langou, and M. Rozloznik, “The loss of orthogonality in the Gram-Schmidt orthogonalization process,” *Computers & Mathematics with Applications*, vol. 50, no. 7, pp. 1069–1075, 2005.

[12] A. A. Kucharski and P. M. Slobodzian, “The application of macromodels to the analysis of a dielectric resonator antenna excited by a cavity backed slot,” in *Microwave Conference, 2008. EuMC 2008. 38th European*. IEEE, 2008, pp. 519–522.