This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Improved Well-Conditioned Model Order Reduction Method Based on Multilevel Krylov Subspaces

Neeraj Kumar, Kalarickaparambil J. Vinoy, *Senior Member, IEEE*, and Srinivasan Gopalakrishnan

***Abstract*—Reduced order models (ROMs) based on the asymptotic waveform evaluation enable fast and efficient parametric analysis of large-scale matrix systems exhibiting nonlinear dependence on certain desired parameter(s). However, they are known to be narrowband due to the inherently ill-conditioned moment generation process. While well-conditioned approaches exist that enforce the moment-matching criteria by introducing some correction terms, these are not optimal and are difficult to parallelize. This letter introduces a well-conditioned multilevel Krylov model order reduction (WMKMOR) technique which is accurate over a larger band and faster to set up than existing approaches. Also, the multiple levels of Krylov subspaces in WMKMOR are generated independently. The improvement in ROM bandwidth using this technique is demonstrated for a finite-element model of an electromagnetic scattering example.**

***Index Terms*—Asymptotic waveform evaluation (AWE), finite-element method (FEM), model order reduction (MOR), scattering.**

I. Introduction

MODEL order reduction (MOR) methods enable fast and efficient parametric analysis of large-scale computational models. Essentially, MOR techniques involve computing and projecting onto a lower dimensional subspace capturing the behavior of the original system in a specified band of parameter variations. This letter focuses on a class of MOR techniques based on moment matching known as asymptotic waveform evaluation (AWE).

AWE and its derivatives have been widely used in the computational electromagnetics (EMs) [1], [2] and the circuit simulation communities [3], [4]. However, the traditional AWE suffers from an inherently ill-conditioned moment generation process leading to error accumulation and stagnation as the reduced order model (ROM) size increases [2]. The well-conditioned AWE (WCAWE) [2], [5] tackled the stagnation by enforcing orthonormalization through some correction terms in its iterations. However, it lacks on the following aspects.

1) It is iterative and therefore inherently sequential.
2) It is relatively cumbersome to implement compared to the regular AWE.
3) It lacks any Krylov subspace structure.

To address these issues, we introduced the concept of multilevel Krylov subspaces [6], where the moment vector space is shown to be embedded inside a multilevel Krylov subspace structure. In this letter, we refine the idea further and employ the implicit orthonormalization of the constituent Krylov subspaces (without breaking the moment-matching property) to construct an ROM that is much more accurate and faster to set up. Although the idea can be applied to other engineering domains, for demonstration, we choose an EM scattering problem modeled using the finite-element method (FEM).

II. Well-Conditioned Multilevel Krylov Subspaces

EM FEM models incorporating absorbing boundaries and/or media with losses lead to matrix systems that depend nonlinearly on the frequency of excitation. Such systems can be represented as

$$\mathbf{A}(s)\mathbf{x}(s) = \mathbf{b}(s) \tag{1}$$

where $\mathbf{A}$ is a complex matrix, $\mathbf{x}$ is the solution vector, $\mathbf{b}$ is the excitation vector, and s is the frequency parameter. Such systems commonly arise from the FEM modeling of EM scattering problems with ABC boundaries [1].

Expanding (1) in a Taylor series about a center frequency s_0 followed by matching like powers on either side leads to the iteration

$$\begin{aligned}\mathbf{x}_0 &= \mathbf{A}_0^{-1}\mathbf{b}_0\\ \mathbf{x}_n &= \sum_{i=1}^{n}\mathbf{P}_i\mathbf{x}_{n-i} + \tau_n, \quad n \geq 1\end{aligned} \tag{2}$$

where $\mathbf{P}_n = -\mathbf{A}_0^{-1}\mathbf{A}_n$, $\tau_n = \mathbf{A}_0^{-1}\mathbf{b}_n$, $\mathbf{x}_n$ is the nth derivative of the solution vector $\mathbf{x}$ at s_0, and the subscripts denote the derivative order with respect to s.

The original system (1) is then projected onto the subspace spanned by these moment vectors, i.e.,$\mathbf{V}_n = \text{span}\,[\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{n-1}]$, leading to the ROM

$$\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}} \tag{3}$$

where $\tilde{\mathbf{A}} = \mathbf{V}_n^H\mathbf{A}\mathbf{V}_n$ and $\tilde{\mathbf{b}} = \mathbf{V}_n^H\mathbf{b}$.

Equation (3) can be solved quickly and efficiently to find $\tilde{\mathbf{x}}$, and the approximate solution to the full system (1) reconstructed as $\mathbf{x}_{\text{ROM}} = \mathbf{V}_n\tilde{\mathbf{x}}$. However, the process in (2) is ill-conditioned as it suffers from round-off errors. WCAWE attempts to construct a well-conditioned basis for $\mathbf{V}_n$ using some correction terms. However, WCAWE is cumbersome to implement and lacks any Krylov structure. Another technique, SAPOR [7], uses the second-order Krylov subspace but is based on linearization and limited to the second-order systems

Manuscript received September 4, 2018; accepted October 22, 2018. *(Corresponding author: Neeraj Kumar.)*

N. Kumar is with Huawei Technologies India Pvt. Ltd., Bengaluru 560037, India (e-mail: neerajkumar9583@gmail.com).

K. J. Vinoy is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru 560012, India (e-mail: kjvinoy@iisc.ac.in; kjvinoy@ieee.org).

S. Gopalakrishnan is with the Department of Aerospace Engineering, Indian Institute of Science, Bengaluru 560012, India (e-mail: krishnan@iisc.ac.in).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LMWC.2018.2878957

with excitation vector depending linearly on the frequency. We now introduce a multilevel Krylov subspace structure that is simple, parallelizable, and applicable to systems and excitations with arbitrary nonlinear dependence on frequency.

Rearranging (2) and introducing some notations, we get

$$\begin{aligned}
\mathbf{x}_0 &= \mathbf{f}_0\\
\mathbf{x}_1 &= \mathbf{f}_1 + \tau_1\\
\mathbf{x}_2 &= \mathbf{f}_2 + \mathbf{P}_1\mathbf{t}_1 + \tau_2\\
&\ \ \vdots\\
\mathbf{x}_n &= \mathbf{f}_n + \sum_{i=1}^{n-1}\mathbf{P}_i\mathbf{t}_{n-i} + \tau_n
\end{aligned}\tag{4}$$

where

$$\begin{aligned}
\mathbf{f}_0 &= \mathbf{A}_0^{-1}\mathbf{b}_0,\\
\mathbf{f}_n &= \sum_{i=1}^{n}\mathbf{P}_i\mathbf{f}_{n-i}, \quad n \geq 1\\
\mathbf{t}_1 &= \tau_1\\
\mathbf{t}_{n-1} &= \sum_{i=1}^{n-2}\mathbf{P}_i\mathbf{t}_{n-i-1} + \tau_{n-1}, \quad n \geq 3
\end{aligned}\tag{5}$$

The excitation derivatives τ_i can have significantly larger magnitudes than the other summands in (4). This results in significant round-off errors in $\mathbf{x}_i$ as i increases. To minimize these errors, we propose to isolate the influence of each τ_i by letting it to generate its own Krylov subspace. If we start with the following notation:

$$\begin{aligned}
\mathcal{X}_n &= [\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n]\\
\mathcal{K}_n^0 &= [\mathbf{f}_0, \mathbf{f}_1, \ldots, \mathbf{f}_n]\\
\mathcal{W}_1 &= [\tau_1]\\
\mathcal{W}_2 &= [\mathbf{P}_1\mathbf{t}_1 + \ \tau_2]\\
&\ \ \vdots\\
\mathcal{W}_n &= \left[\sum_{i=1}^{n-1}\mathbf{P}_i\mathbf{t}_{n-i} + \tau_n\right]
\end{aligned}\tag{6}$$

then (4) implies that

$$\mathcal{X}_n \subset \mathcal{K}_n^0 + \mathcal{T}_n^1 \tag{7}$$

where $\mathcal{T}_n^1 = \sum_{i=1}^{n}\mathcal{W}_i$ and the superscripts denote the space id (or level). An nth second-order Krylov subspace is defined as [8]

$$\mathcal{K}_n(\mathbf{A}, \mathbf{B}; \mathbf{v}) = \operatorname{span}\{\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{n-1}\} \tag{8}$$

where

$$\begin{aligned}
\mathbf{r}_0 &= \mathbf{v}, \quad \mathbf{r}_1 = \mathbf{A}\mathbf{r}_0\\
\mathbf{r}_j &= \mathbf{A}\mathbf{r}_{j-1} + \mathbf{B}\mathbf{r}_{j-2} \quad \text{for } j \geq 2
\end{aligned}\tag{9}$$

for square matrices $\mathbf{A}$ and $\mathbf{B}$ and the seed vector $\mathbf{v}$. Similarly, we define $\mathcal{K}_n^0 = \mathcal{K}_{n+1}(\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_p; \mathbf{f}_0)$ in (6) as an $(n+1)$th dimensional pth-order Krylov subspace seeded by the vector $\mathbf{f}_0$, where p is the maximum order of derivative possible for matrix A in (1). Clearly, the iterates of $\mathcal{K}_n^0$ are unaffected by $\mathbf{t}_i$ and τ_i. Thus, from (7), the AWE space $\mathcal{X}_n$ is embedded inside a bigger space formed by augmenting the generalized Krylov subspace $\mathcal{K}_n^0$ with $\mathcal{T}_n^1$. Kumar *et al.* [9] had demonstrated that this augmentation results in an ROM with enhanced accuracy. This can be attributed to the reduced round-off errors due to the separate and independent generation of the Krylov and the augmenting subspaces. To extend this idea further, note that $\mathcal{T}_n^1$ has the same structure as (2) and has dimension one less than that of $\mathcal{X}_n$. This makes it a candidate for further splitting. Therefore, (7) implies that

$$\mathcal{X}_n \subset \mathcal{K}_n^0 + \left(\mathcal{K}_n^1 + \mathcal{T}_n^2\right) \tag{10}$$

where $\mathcal{K}_n^1$ is an nth dimensional pth-order Krylov subspace seeded by the vector τ_1. Note that space-id (superscript) 1 in $\mathcal{T}_n^1$ corresponds to that in $\mathcal{K}_n^1$. Continuing with this approach, if b $(\leq n)$ is the maximum number of derivatives of the $\mathbf{b}$ vector to incorporate, we observe that

$$\begin{aligned}
\mathcal{X}_n \subset \mathcal{S} &= \mathcal{K}_n^0 + \left(\mathcal{K}_n^1 + \left(\mathcal{K}_n^2 + \ldots \left(\mathcal{K}_n^{b-1} + \mathcal{T}_n^b\right)\right)\right)\\
&= \sum_{i=0}^{b-1}\mathcal{K}_n^i + \mathcal{T}_n^b
\end{aligned}\tag{11}$$

where $\mathcal{K}_n^i$ is an $(n-i+1)$th dimensional pth-order Krylov subspace generated by the seed vector τ_i. If $b = n$, then $\mathcal{K}_n^n = \mathcal{T}_n^n = [\tau_n]$ and $\mathcal{S} = \sum_{i=0}^{n}\mathcal{K}_n^i$. Finally, each $\mathcal{K}_n^i$ can be orthonormalized using WCAWE by a thread i in a chosen parallelization setting. We name this procedure as well-conditioned multilevel Krylov MOR (WMKMOR) and are shown in Algorithm 1.

The symbol $\mathbf{e}_q$ represents the vector of zeros except in the position q where the value is 1, and $\mathbf{P}_{\mathbf{U}w}(n, m) = \prod_{t=w}^{m}\mathbf{U}^{-1}_{[t:n-m+t-1, t:n-m+t-1]}$ are the correction terms with $w = 1$or 2, as described in [2] and [5].

Algorithm 1 *The Well-Conditioned MKMOR*

for id $= 0, 1, 2, \ldots, \ (b-1)$ **do**
  $\hat{\mathbf{v}}_1^{\text{id}} = \mathbf{A}_0^{-1}\mathbf{b}_{\text{id}}$
  $\mathbf{U}_{[1,1]}^{\text{id}} = \left\|\hat{\mathbf{v}}_1^{\text{id}}\right\|$
  $\mathbf{v}_1^{\text{id}} = \hat{\mathbf{v}}_1^{\text{id}}\mathbf{U}_{[1,1]}^{\text{id},-1}$
  for $n = 2, 3, \ldots,$ (order $-$ id)**do**
    $\hat{\mathbf{v}}_n^{\text{id}} = \mathbf{P}_1\mathbf{v}_{n-1}^{\text{id}} - \sum_{m=2}^{\min(a_1, n-1)}\mathbf{P}_m\mathbf{V}_{n-m}^{\text{id}}\mathbf{P}_{\mathbf{U}2}(n, m)\,\mathbf{e}_{n-m}$
    for $\alpha = 1, 2, \ldots, n-1$do
      $\mathbf{U}_{[\alpha,n]}^{\text{id}} = \mathbf{v}_\alpha^{\text{id},*}\hat{\mathbf{v}}_n^{\text{id}}$
      $\hat{\mathbf{v}}_n^{\text{id}} = \hat{\mathbf{v}}_n^{\text{id}} - \mathbf{U}_{[\alpha,n]}^{\text{id}}\mathbf{v}_\alpha^{\text{id}}$
    endfor
    $\mathbf{U}_{[n,n]}^{\text{id}} = \left\|\hat{\mathbf{v}}_n^{\text{id}}\right\|$
    $\mathbf{v}_n^{\text{id}} = \hat{\mathbf{v}}_n^{\text{id}}\mathbf{U}_{[n,n]}^{\text{id},-1}$
  endfor
endfor
$WCAWE\left(\mathcal{T}_{order}^{b}\right)$

III. Numerical Example

To demonstrate the improvement in accuracy achieved using WMKMOR, we choose as an example a patch buried in a dielectric-filled cavity recessed in a perfect electric conductor (PEC) ground plane [10]. The full FEM model in (1) is based on the total field formulation, first-order ABC and has 80 138 degrees of freedom. The frequency band of interest is 4–6 GHz with the expansion point for the ROM chosen at the center. We define a residual $\mathbf{r} = (\mathbf{A}\mathbf{V}_n\tilde{\mathbf{x}} - \mathbf{b})$ and specify

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

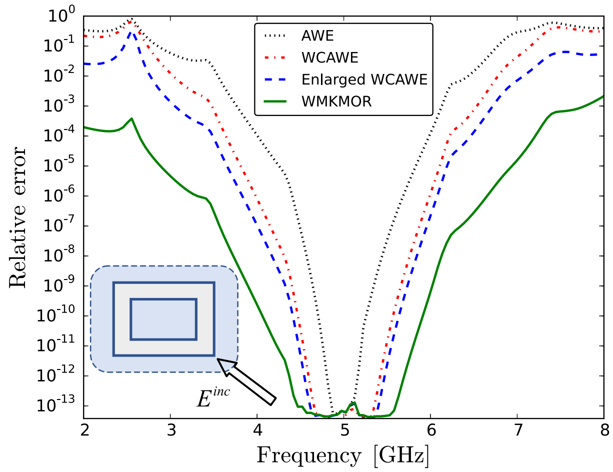

Fig. 1. Relative errors for ROMs based on WMKMOR, enlarged WCAWE with the same size (136) as WMKMOR, WCAWE, and AWE all using 15 excitation vector derivatives. Inset: top view of the geometry. The patch (3.66×2.8 cm^2) is buried halfway in a dielectric-filled cavity (7.32×5.2 cm^2) 0.316 cm deep recessed in the PEC ground. A plane wave is incident at $\phi^{\text{inc}} = 45°$ and $\theta^{\text{inc}} = 60°$. The dielectric has a relative permittivity $\varepsilon_r = 2.17$ and a loss tangent 0.001.

a tolerance $\xi = 1e - 5$ such that the ROM size can increase until

$$r_{\text{rel}} = \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}\bigg|_{s_{\text{edge}}} < \xi. \quad (12)$$

Both $\mathbf{r}$ and $\mathbf{b}$ are evaluated at $s_{\text{edge}} \in \{4, 6\}$ [11], [2]. This led to 16 vectors generated by WCAWE involving up to 15 excitation derivatives. This implies that 16 levels of Krylov subspaces be independently generated and orthonormalized for WMKMOR. These subspaces are then joined and the composite space reorthonormalized.

To compare the ROM with respect to the FEM model, we define

$$\text{Relative error} = \frac{\|\mathbf{x}_{\text{ROM}} - \mathbf{x}_{\text{FEM}}\|_2}{\|\mathbf{x}_{\text{FEM}}\|_2}. \quad (13)$$

Fig. 1 shows the comparison of the relative errors of WCAWE and WMKMOR. Clearly, the multilevel Krylov subspace structure of WMKMOR leads to about 2 orders of accuracy gain at the band edges than WCAWE. However, whereas WCAWE generates 16 vectors, WMKMOR ends up generating a total of 136 vectors due to its multiple levels. Therefore, for fair comparison, we let WCAWE iterations to continue further till its subspace size matches that of WMKMOR, both using 15th-order derivatives of the excitation vector $\mathbf{b}$. Clearly, this "enlarged" WCAWE subspace exhibits stagnation since it can no longer match moments if excitation derivatives higher than order 15 are not included.

To compare ROM speed with FEM, we define [2]

$$\text{Breakeven point} = \frac{\text{ROM setup time}}{\text{FEM solve time} - \text{ROM solve time}}. \quad (14)$$

Table I shows the breakeven points for the two. Again, a faster setup time for WMKMOR leads to lower (better) breakeven point in comparison with WCAWE and its expanded version. Also, WMKMOR captures 76% more bandwidth than that captured by WCAWE.

Note that for a chosen residual tolerance ξ, if n vectors are generated by WCAWE, WMKMOR generates about $(n)(n+1)/2$ vectors. Thus, although WMKMOR is fast and accurate over a wider band than WCAWE and its extended version, ξ dictates the memory footprint of the algorithm and has to be chosen judiciously. We recommend using multipoint expansions in larger bands.

TABLE I
BREAKEVEN POINTS FOR VARIOUS ROMS WITH RESPECT TO THE FEM MODEL

ROM Type	ROM setup time (s)	ROM solve time (s)	Break-even point	Normalized bandwidth w.r.t. WCAWE (for $\xi = 1e-5$)
AWE	15.72	0.13	6.52	0.67
WCAWE	22.93	0.13	9.52	1
Enlarged WCAWE	59.45	0.15	24.87	1.1
WMKMOR	16.61	0.15	6.95	1.76

IV. CONCLUSION

A novel approach has been introduced to improve the accuracy of AWE-based ROMs. We introduce the notion of multilevel pth-order Krylov subspaces and show that the AWE space is embedded inside this multilevel structure. These subspaces are generated independently and in parallel. While the first-order and second-order Krylov subspaces can be orthonormalized using the Arnoldi/Lanczos process and the linearization-based SAPOR method, respectively, orthonormalizing the pth-order Krylov subspaces in the multilevel structure requires a well-conditioned approach that maintains the moment-matching property. To achieve this, multiple instances of WCAWE orthonormalize the structure, in parallel. The resulting WMKMOR method is shown to be far more accurate in a given band and faster to set up than WCAWE ROM of similar size.

REFERENCES

[1] Y. Zhu and A. C. Cangellaris, Eds., *Multigrid Finite Element Methods for Electromagnetic Field Modeling*. Hoboken, NJ, USA: Wiley, 2006.

[2] R. D. Slone, "Fast frequency sweep model order reduction of polynomial matrix equations resulting from finite element discretizations," Ph.D. dissertation, ElectroSci. Lab., Ohio State Univ., Columbus, OH, USA, 2002.

[3] L. T. Pillage and R. A. Rohrer, "Asymptotic waveform evaluation for timing analysis," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 9, no. 4, pp. 352–366, Apr. 1990.

[4] E. Chiprout and M. S. Nakhla, *Asymptotic Waveform Evaluation and Moment Matching for Interconnect Analysis*. Norwell, MA, USA: Kluwer, 1994.

[5] R. D. Slone, R. Lee, and J.-F. Lee, "Well-conditioned asymptotic waveform evaluation for finite elements," *IEEE Trans. Antennas Propag.*, vol. 51, no. 9, pp. 2442–2447, Sep. 2003.

[6] N. Kumar, K. J. Vinoy, and S. Gopalakrishnan, "A reduced order model for electromagnetic scattering using multilevel Krylov subspace splitting," in *Proc. IEEE Int. Conf. Comput. Electromagn.*, Hong Kong, Feb. 2015, pp. 341–343.

[7] 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 *Proc. IEEE/ACM Int. Conf. Comput. Aided Des.*, Nov. 2004, pp. 74–79.

[8] Z. Bai and Y. Su, "SOAR: A second-order Arnoldi method for the solution of the quadratic eigenvalue problem," *SIAM J. Matrix Anal. Appl.*, vol. 26, no. 3, pp. 640–659, Mar. 2005.

[9] N. Kumar, K. J. Vinoy, and S. Gopalakrishnan, "Augmented Krylov model order reduction for finite element approximation of plane wave scattering problems," in *Proc. IEEE MTT-S Int. Microw. RF Conf.*, New Delhi, India, Dec. 2013, pp. 1–4.

[10] J.-M. Jin and J. L. Volakis, "A hybrid finite element method for scattering and radiation by microstrip path antennas and arrays residing in a cavity," *IEEE Trans. Antennas Propag.*, vol. 39, no. 11, pp. 1598–1604, Nov. 1991.

[11] X. M. Zhang and J.-F. Lee, "Application of the AWE method with the 3-D TVFEM to model spectral responses of passive microwave components," *IEEE Trans. Microw. Theory Techn.*, vol. 46, no. 11, pp. 1735–1741, Nov. 1998.