Compact Modeling of Interconnect Circuits over Wide Frequency Band by Adaptive Complex-Valued Sampling Method

HAI WANG, SHELDON X.-D. TAN, and RYAN RAKIB, University of California, Riverside

In this article, we propose a new model order-reduction method for compact modeling of interconnect circuits over wide frequency band using a novel complex-valued adaptive sampling and error estimation scheme. We address the outstanding error control problems in the existing sampling-based reduction framework over a frequency band. Our new method, *WBMOR*, explicitly and efficiently computes the exact residual errors to guide the sampling process. We show by sampling along the imaginary axis and performing a new complex-valued reduction that the reduced model will match exactly with the original model at the sample points. Additionally, we show in theory that the proposed method can achieve the error bound over a given frequency range. In practice, the new algorithm can help designers choose the best order of the reduced model for the given frequency range and error bound via the adaptive sampling scheme. In addition, *WBMOR* can perform wideband accurate reductions of interconnect circuits for analog and RF applications where model accuracy needs to be maintained over a wide frequency range. We compare several sampling schemes such as Monte Carlo, logarithmic, recently proposed resampling, and ARMS methods. Experimental results on a number of RLC circuits show that *WBMOR* is much more efficient than all the other sampling methods, including the recently proposed resampling and ARMS schemes with the same reduction orders. Compared with the traditional real-valued sampling methods, the complex-valued sampling method is more accurate for the same computational cost.

Categories and Subject Descriptors: J.6 [**Computer Applications**]: Computer-Aided Engineering—*Computer-aided design*

General Terms: Design, Algorithms

Additional Key Words and Phrases: Model reduction, adaptive, complex-valued sampling

ACM Reference Format:
Wang, H., Tan, S. X.-D., and Rakib, R. 2011. Compact modeling of interconnect circuits over wide frequency band by adaptive complex-valued sampling method. ACM Trans. Des. Autom. Electron. Syst. 17, 1, Article 5 (January 2012), 22 pages.
DOI = 10.1145/2071356.2071361 http://doi.acm.org/10.1145/2071356.2071361

1. INTRODUCTION

Model order reduction (MOR) is an efficient technique for reducing the complexity of parasitic interconnect circuits. Existing approaches based on the Krylov subspace are very efficient [Antoulas 2005; Feldmann and Freund 1995; Odabasioglu et al. 1998; Silveira et al. 1996]. These methods, which perform implicit moment-matching by

This work is supported in part by NSF grants CCF-1116882, OISE-1130402, OISE-1051797 and in part by UC-MEXUSCONACyT Collaborative Research grant (2011-2013).
Some preliminary results of this paper appeared in the *Proceedings of the 15th Asia and South Pacific Design Automation Conference (ASPDAC)* [Wang et al. 2010].
Authors' address: H. Wang, S. X.-D. Tan, and R. Rakib, Department of Electrical Engineering, University of California, Riverside, CA 92521; email: {hawang, stan, rrakib}@ee.ucr.edu.
Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from the Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701, USA, fax +1 (212) 869-0481, or permissions@acm.org.
© 2012 ACM 1084-4309/2012/01-ART5 $10.00
DOI 10.1145/2071356.2071361 http://doi.acm.org/10.1145/2071356.2071361

projecting the original system onto a Krylov subspace, preserve the stability, passivity, and structural information for RLC interconnect circuits.

The Krylov subspace methods suffer from one long-standing problem: the lack of a global error bound. Due to this problem, designers can't predict the errors of the reduced model over a given frequency range before the reduction. For computing wire delay, coupled noise, and voltage noise of interconnects in the digital circuits, such lack of global errors is a lesser concern because matching a few moments is accurate enough. However, for analog and RF circuits, reduced models of the extracted RLCK circuits need to be accurate for a wide frequency range and in terms of the time domain waveforms (versus delay values). The existing Krylov subspace methods can't meet such a requirement, although some techniques like multipoint Krylov subspace can partially mitigate this problem at high computational cost.

Truncated balanced realization based methods (TBR) [Kamon et al. 2000; Moore 1981; Phillips et al. 2003; Yan et al. 2007] have a global error bound, but generally suffer from high computing costs, poor scalability, although some effort has been made to improve the computing efficiency [Zhou 2002; Wang and Balakrishnan 2005; Vasilyev and White 2005]. The high cost comes from computing the controllability and observability Gramians, upon which the truncation, and thus reduction of weak states, is performed. This problem was partially mitigated by the proposed Gramian approximation techniques [Phillips and Silveira 2005; Willcox and Peraire 2002], where Gramians are approximated in the frequency domain by computing system impulse responses at many frequency points (samples). However, how to efficiently perform the sampling (how many points and which points should be chosen) to control the errors of the reduced models still remains an open problem. Two methods were proposed recently to resolve this problem: a resampling scheme [Silveira and Phillips 2006] and the ARMS method [Villena and Silveira 2009]. In Silveira and Phillips [2006], a statistical resampling scheme is used. In each iteration, many reduced models are computed by resampling from a common pool of candidate sample points in a given frequency range, and the variations among all the reduced models are calculated at many frequency points, called search points. The search point with the largest variation will be taken as the sample point. The ARMS method uses residue-minimization technique to estimate the error (without building the reduced models) and add sample points where the largest residue exists. However, these methods may not find the best sample points or suffer from the large computational cost, as shown in our experimental results.

In this article, we address the error control issues in the sampling-based reduction framework. The new algorithm, *WBMOR*, consists of two important features: first, it uses the exact residual to estimate the errors of the reduced models and shows residual computation can be done very efficiently. Second, unlike the existing sampling-based reduction methods, *WBMOR* samples along the imaginary axis and performs new complex-valued reduction, leading to exact matching for the reduced model at the sampled frequency points. We show that the proposed method achieves the bounded error over a given frequency range. Practically, the new algorithm can help designers choose the optimal order of the reduced model for the given frequency range and error bound. In addition, *WBMOR* can perform wideband accurate reductions of interconnect circuits for analog and RF applications. Experimental results on a number of RLC circuits show that compared with the real-valued sampling methods, the complex-valued sampling method is more accurate for the same computational cost. *WBMOR* is much more efficient than the recently proposed resampling and ARMS schemes. We also compare several sampling schemes such as logarithmic and Monte Carlo sampling methods. The proposed *WBMOR* has been integrated into the reduction tool—UiMOR [Tan et al. 2010] (for UC Riverside Model Order Reduction Tool Suite) , which can be downloaded from UC Riverside MSLAB website [uim].

This article is organized as follows. In Section 2, we review the sampling-based reduction methods. In Section 3, we introduce our new wideband sampling-based reduction method. Experimental results are reported in Section 4 to demonstrate the effectiveness of our proposed method. Section 5 concludes.

2. REVIEW OF BALANCED TRUNCATION METHODS

2.1 The Standard TBR-Based Reduction Method

We first review the standard TBR method for the completeness of this article.

Consider a linear dynamic system in a standard state-space form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t). \end{aligned} \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{p\times n}$, and $u(t) \in \mathbb{R}^p$. The controllable Gramian X and the observable Gramian Y are the unique symmetric positive definite solutions to the Lyapunov equations.

$$\begin{aligned} AX + XA^T + BB^T &= 0 \\ A^TY + YA + C^TC &= 0. \end{aligned} \tag{2}$$

Since the eigenvalues of the product XY are invariant under similarity transformation, we can perform a similarity transformation ($A_b = T^{-1}AT, B_b = T^{-1}B, C_b = CT$) to diagonalize the product XY such that

$$T^{-1}XYT = \Sigma = diag({\sigma_1}^2, {\sigma_2}^2, \ldots, {\sigma_n}^2) \tag{3}$$

where the Hankel singular values of the system (σ_k) are arranged in a descending order. If we partition the matrices as

$$\begin{bmatrix} W_1^T \\ W_2^T \end{bmatrix} XY \begin{bmatrix} V_1 & V_2 \end{bmatrix} = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \tag{4}$$

where $\Sigma_1 = diag({\sigma_1}^2, {\sigma_2}^2, \ldots, {\sigma_r}^2)$ are the first r largest eigenvalues of Gramian product XY and W_1 and V_1 are corresponding eigenvectors. A reduced model can be obtained as follows:

$$\begin{aligned} \dot{x}(t) &= A_r x(t) + B_r u(t) \\ y(t) &= C_r x(t) \end{aligned} \tag{5}$$

where $A_r = W_1^T A V_1$, $B_r = W_1^T B$, $C_r = CV_1$. The error in the transfer function of the order r approximation is bounded by $2\sum_{i=r+1}^{N} \sigma_k$. In the TBR procedure, the computational cost is dominated by solving Lyapunov equations $O(n^3)$, which makes it too expensive to apply to integrated circuit problems, and thus an efficient Gramian approximation technique is highly appreciated.

2.2 The Sampling-Based Reduction Framework

To mitigate the high computational cost of the standard TBR method, a fast TBR method was proposed (called PMTBR) [Phillips and Silveira 2005], where the Gramians are approximated using the Monte-Carlo sampling approach. Specifically, if we look at the controllability Gramian, we need to solve the following Lyapunov equation:

$$AX + XA^T + BB^T = 0. \tag{6}$$

Alternatively, the Gramian X can also be represented in frequency domain as

$$X = \int_{-\infty}^{+\infty} (j\omega I - A)^{-1} B B^T (j\omega I - A)^{-H} d\omega \tag{7}$$

where superscript H denotes a Hermitian transpose. As a result, computing Gramian X boils down to evaluating the definite integral in (7) [Iserles 1996]. This can be done using numerical quadrature methods.

For an integral function $f(x)$, numerical quadrature methods try to approximate it as

$$\int_a^b f(x)dx \approx \sum_{k=1}^{m} w_k f(x_k) \tag{8}$$

where w_k, $k = 1, 2, ..., m$, are referred to as the quadrature point weights, while the interpolation points x_k, $k = 1, 2, ..., m$, are called quadrature points. The selections of w_k and x_k depend on the quadrature methods such as Newton-Cotes, Gaussian quadrature rules [Iserles 1996].

For the sampling-based reduction, our goal is not just computing the Gramian X, but the dominant eigenspace to form the projection matrix. As a result, let s_k be the kth sample point, if we define

$$z_k = z(s_k) = (s_k I - A)^{-1} B, \tag{9}$$

which is called kth snapshot of the system in (1) in the frequency domain. Then X can be approximated as

$$\hat{X} = \sum_{k=1}^{m} w_k z_k z_k^H = Z W^2 Z^H \tag{10}$$

where

$$Z = [z_1, z_2, \ldots, z_m] \tag{11}$$

and W is a diagonal matrix with diagonal entries $w_{kk} = \sqrt{w_k}$. w_k comes from a specific quadrature method. If we perform the singular value decomposition (SVD) on ZW and obtain

$$ZW = V \Sigma U \tag{12}$$

then V, which gives the dominant eigenspace of X, is used as the projection matrix.

The main advantage of the sampling-based TBR methods over the Krylov subspace methods is that they are globally more accurate, since they sample in a wide frequency range. However, given a large original model and error bound, how to choose the sample points efficiently to generate the reduced model as compact as possible with bounded error still remains an open problem.

3. NEW WIDEBAND SAMPLING-BASED REDUCTION METHOD

In this section, we address the efficient sampling problem of the sampling-based reduction methods. In Section 3.1, we first present the problem and three challenges in solving the problem. Then, we conquer the first one with the complex-valued reduction and present some of its important properties. Next, in Section 3.2, we conquer the second challenge and show how error is estimated in the proposed method in

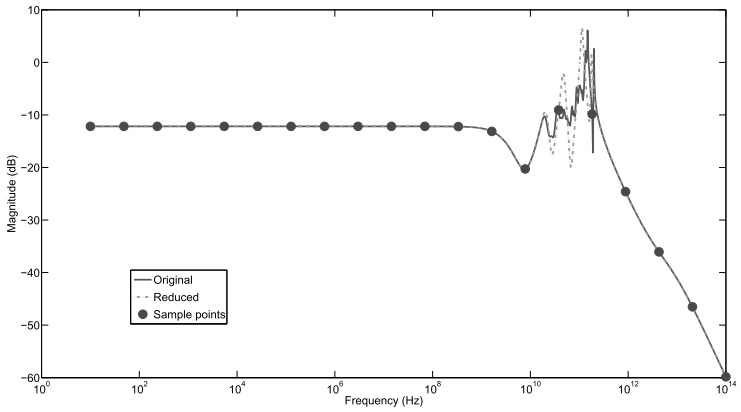

Fig. 1. Real axis sampling with 20 sample points.

Section 3.3. The third challenge is solved in Section 3.4 using an adaptive sampling scheme. Finally, the whole algorithm flow and analysis are presented in Section 3.5.

3.1 Challenges of Efficient Sampling

Typically, in the frequency band where the frequency response of a system is smooth, a sample point (its corresponding subspace) can cover[1] a wide frequency range. While in the spiky district, where many system poles exist, a sample point will only cover a small interval. As a result, pure uniform or random Monte Carlo sampling schemes may over-sample at the smooth frequency bands and suffer accuracy loss at the spiky bands. Let's illustrate this via an example of a RLC circuit of size 640. 20 uniformly distributed sample points are used to generate a reduced model of size 20. Obviously, as shown in Figure 1, the accuracy for this reduction is not satisfactory—the approximation is good for the smooth frequency bands, but it fails to match the area from 10^{10}Hz to 10^{12}Hz, which has a lot of peaks.

One way to mitigate this problem is by a static sampling scheme which simply adds more sample points at the spiky frequency bands. Unfortunately, knowing the frequency response of a large system as *a priori* is not feasible, since it requires solving the large system over the wide frequency range of interest. Moreover, we will not know how many samples will lead to an acceptable reduced model unless we perform the reduction many times.

A better way to perform reducion efficiently is via adaptive sampling. In the iterative adaptive sampling method, new sample points are added at the current iteration based on the information of the last iteration. However, there are three challenges we need to solve in this scheme. First, we desire the reduced system to have zero errors at the sample points and thus guarantee the convergence as more samples are added. This seems simple, but as will be shown later, existing real axis sampling does not have this property. Second, in order to guide the adaptive sampling at the next iteration and to indicate the convergence of the algorithm, we need to know the errors of the reduced system at the current iteration over the frequency range. However, calculating the errors explicitly requires expensive solving of the original model at many points, which should be avoided as much as possible. As a result, an error estimator with a cheaper computational cost is preferred. Third, assume we have solved the first

[1]*Cover* here means that the reduced model generated by the projection matrix matches the original system in the specified frequency band.

two problems, the sample points still need to be placed carefully. For example, how can we accelerate the convergence of the iteration? How can we prevent missing the potential sample points? The following sections will provide our solutions to all three challenges.

3.2 New Complex-Valued Sampling-Based Reduction Method

As shown in Figure 1, the reduced model may fail to match the original model exactly even at the sample points, such as at the fifth sample points from right to left. In this section, we show through sampling along the imaginary axis that we will obtain the exact match at the sample points.

Existing reduction techniques based on the Krylov subspace method are mainly expanded at $s = 0$, while multipoint expansion methods and fast TBR methods are expanded at multiple frequency points, but along the real axis. In other words, $s_k = \sigma_k$ instead of $s_k = j\omega_k$ is used in (9).[2] The reason is that if $s = j\omega$, the moments in the Krylov subspace methods and snapshots in the sampling-based approaches will become complex, causing the reduced matrices $\hat{G}$, $\hat{C}$, $\hat{B}$ and $\hat{L}$ to be complex matrices. This complex reduced system does not exist in the real world, and is hard to be realized into a reduced RLC circuit. We notice that expanding along the imaginary axis is not an issue for an explicit moment-matching method like the AWE and CFH methods [Chiprout and Nakhla 1995], as complex Padé approximation can be carried out to compute the poles and residues.

In this article, we propose a new sampling scheme called Complex-Valued Sampling-based Truncated Balanced Realization (*CVSTBR*) as a part of our *WBMOR* algorithm. *CVSTBR* samples along the imaginary axis $s = j\omega$ to preserve the physical meaning of the Gramian approximation. In addition, the new *CVSTBR* method leads to a *real* reduced system. We show that the resulting reduced system matches exactly with the original system at the sample points, which is not the case for sampling along the real axis (except for $s = 0$).

3.2.1 The Projection Framework of Sampling-Based Model Reduction. First, we present the projection framework which plays an important role in the sampling-based model reduction method.

For an interconnect circuit modeled as a RLC dynamic system with n states and p ports, the system Eq. (1) can be formulated in descriptor form in the Laplace domain as

$$\begin{aligned} Gx(s) + sCx(s) &= Bu(s), \\ y(s) &= Lx(s) \end{aligned} \tag{13}$$

where $G \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{n\times n}$ contain element information such as conductance, capacitance, and inductance. B and L matrices describe the positions of inputs and outputs in the network, respectively. And typically there is $B = L^T \in \mathbb{R}^{n\times p}$, which means the inputs and outputs are identical. $x(s) \in \mathbb{C}^n$ is the state variable vector that represents

[2]We remark that in Phillips and Silveira [2005], the authors mentioned that PMTBR can sample in the whole complex plane, but no further detail was provided. Here we claim that the imaginary axis should be the only place to sample.

node voltages and branch currents, where $s = \sigma + j\omega \in \mathbb{C}$. Then, for convenience, we rewrite (9), (10), (11), and (12):

$$z_k = z(s_k) = (s_k C + G)^{-1} B \tag{14}$$

$$\hat{X} = \sum_{k=1}^{m} w_k z_k z_k^H = Z W^2 Z^H \tag{15}$$

$$Z = [z_1, z_2, \ldots, z_m] \tag{16}$$

$$Z W = V \Sigma U \tag{17}$$

After the reduction, we will have a reduced model

$$\begin{aligned} \hat{G} x_r(s) + s\hat{C} x_r(s) &= \hat{B} u(s), \\ y_r(s) &= \hat{L} x_r(s) \end{aligned} \tag{18}$$

where $\hat{G} = V^T G V$, $\hat{C} = V^T C V$, $\hat{B} = V^T B$, $\hat{L} = LV$. V is computed from (17) and $V \in \mathbb{R}^{n \times q}$. $q \ll n$ is the dimension of the reduced system. $x_r(s) \in \mathbb{C}^q$ is the reduced state vector in the reduced system.

Assume we have p impulse inputs $u(s) = [e_1, e_2, ..., e_p]$ applied to the system in (13), where e_i is a $p \times 1$ vector whose ith position is 1 and the rest is 0. Define the impluse response matrix as

$$z(s) = (G + sC)^{-1} B = [x^1(s), x^2(s), ..., x^p(s)] \in \mathbb{C}^{n \times p} \tag{19}$$

where $x^i(s)$ is the state response according to e_i. Correspondingly, we have

$$z_r(s) = (\hat{G} + s\hat{C})^{-1} \hat{B} = [x_r^1(s), x_r^2(s), ..., x_r^p(s)] \in \mathbb{C}^{q \times p} \tag{20}$$

in the reduced system. From now on, we will use $z(s)$ and $z_r(s)$ instead of $x(s)$ and $x_r(s)$ for the identity $u(s)$ case.

The approximate impulse response $\tilde{z}(s) \in \mathbb{C}^{n \times p}$, which is the approximation of the original impulse response $z(s)$, can be recovered from the reduced impulse response $z_r(s)$ as

$$\tilde{z}(s) = V z_r(s) \approx z(s). \tag{21}$$

It can be also written as

$$\tilde{z}(s) = V(V^T Q(s) V)^{-1} V^T Q(s) z(s) \tag{22}$$

where $Q(s) = (G + sC)$ for simplicity [Galántai 2004] and, for short, we will use Q later.

Let's denote $P = V(V^T Q V)^{-1} V^T Q$. Since $P^2 = P$, it is clear that the matrix P is a projector along $Q^T V$ onto V. The range space of P is the subspace spanned by V and the orthogonal complement of its null space is spanned by $Q^T V$. Also notice that P is a function of s. Then, combining (21) and (22), we have

$$\tilde{z}(s) = V z_r(s) = P z(s) \approx z(s). \tag{23}$$

We have shown the relationship between the impulse responses of the original system and the reduced system in (23). It is obvious that the quality of the approximation in (23) is determined by the projector P. Next, we show how to obtain a good approximation by choosing the appropriate projector.

It is well known that if the subspace spanned by a matrix F is inside the range space of a projector P, F remains invariant after the projection by P, that is, $P = PF$. In our case, in order to find a reduced system without accuracy lost, V, which spans the range space of P, should include all the subspaces spanned by $z(s)$. This results in $\tilde{z}(s) = z(s)$, which looks like an excellent result. However, this "perfect" projector

usually leads to a "reduced system" with the same size as the original system. Notice that $z(s) \in \mathbb{C}^{n\times p}$ is a function of s and for a fixed value $s = s_i$, $z(s_i)$ spans a subspace of dimension p. For all values of s, the infinite number of p dimensional subspaces usually fill the whole n dimensional space. In order to include all these subspaces, V needs to have the dimension n leading to a $n \times n$ "reduced" system, which has the same size as the original system and is unacceptable for model reduction.

Although it is usually impossible to generate a reduced system that is accurate for all values of s, as just discussed, it is possible to generate a reduced system that is accurate for certain values of s. If we have generated $V \in \mathbb{C}^{n\times mp}$ to include m subspaces spanned by m values of s, then the resulting $\tilde{z}(s)$ matches $z(s)$ at all these values of s. Since $V \in \mathbb{C}^{n\times mp}$, the reduced system has an order of mp. In order to form this special V matrix, we have to sample the original system at the m values of s using (14), and the V matrix has the following property:

$$\text{span}(V) = \text{span}(z(s_1), z(s_2), \ldots, z(s_m)). \tag{24}$$

More details of formulating the V matrix are presented in Section 3.2.4.

We summarize the preceding discussions above in the following theory.

THEOREM 1. *If s_k is sampled , then $\tilde{z}(s_k) = z(s_k)$ for $k = 1, 2, ..., m$, where m is the number of sample points.*

PROOF. According to the definition of projection, P, which is a projector onto V, is the identity operator on the space spanned by V, that is, $\forall x \in V : Px = x$. According to (24), because $z(s_k) \in \text{span}(z(s_1), z(s_2), \ldots, z(s_m)) = \text{span}(V)$,

$$\tilde{z}(s_k) = Pz(s_k) = z(s_k),\ k = 1, 2, ..., m. \qquad \square$$

3.2.2 Accuracy of the Imaginary Axis Sampling. The previous section shows that if we sample the original system at m values of s and form the V matrix, the approximated impulse response matrix $\tilde{z}(s)$ equals to the original $z(s)$ at these sampled points. In this section, we discuss the sample point should be limited to the imaginary axis to capture the frequency response of the original system.

Assume (13) is a stable causal linear system, the system poles are on the left-hand side of the complex plane. On the imaginary axis $s = j\omega$, the Laplace transform converges and is equivalent to the Fourier transform. The frequency response of the system is

$$H(j\omega) = L(G + j\omega C)^{-1}B \tag{25}$$

and the impulse response of the system is the inverse Fourier transform of the frequency response

$$h(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(j\omega)e^{j\omega t}\,d\omega. \tag{26}$$

In order to have the same frequency behavior as the original system at a frequency point ω_k, the reduced system needs to have the same frequency response at both $-j\omega_k$ and $j\omega_k$. It is important because, for the original system, $H(j\omega)$ is conjugate symmetric, that is, $H(-j\omega) = H^*(j\omega)$. As a result, for the time domian waveforms—for example, the impulse response (26), the contributions by the $\omega \in (-\infty, 0)$ (i.e., $\frac{1}{2\pi}\int_{-\infty}^{0} H(j\omega)e^{j\omega t}\,d\omega$) and $\omega \in (0, \infty)$ (i.e., $\frac{1}{2\pi}\int_{0}^{\infty} H(j\omega)e^{j\omega t}\,d\omega$) are conjugate to each other such that their imaginary parts are eliminated, leading to the real-time domain waveform. Failing to retain this symmetric property in the reduced system is unacceptable. In addition, as shown in Section 3.2.4, it is even impossible to generate a real system if the sampling is not performed symmetrically on the imaginary axis.

We introduce V_c to denote the V matrix in the imaginary sampling case. According to (24), the V_c matrix of the imaginary sampling is generated as

$$\text{span}(V_c) = \text{span}(z(j\omega_1), z(j\omega_2), \ldots, z(j\omega_m), z(-j\omega_1), z(-j\omega_2), \ldots, z(-j\omega_m)) \tag{27}$$

for the frequency points $\omega_1, \omega_2, \ldots, \omega_m$ which we would like to match exactly. The practical way of generating the real-valued V_c matrix is presented in Section 3.2.4.

We are ready to present the following result as the extension of Theorem 1.

COROLLARY 3.1. *The reduced system has exact frequency response at the frequency ω_k if $j\omega_k$ and $-j\omega_k$ on the imaginary axis are sampled.*

PROOF. Consider the reduced frequency response $\hat{H}(j\omega)$. If $j\omega_k$ and $-j\omega_k$ are sampled, according to Theorem 1, there are $\tilde{z}(j\omega_k) = z(j\omega_k)$ and $\tilde{z}(-j\omega_k) = z(-j\omega_k)$, then we have

$$\hat{H}(j\omega_k) = L\tilde{z}(j\omega_k) = Lz(j\omega_k) = H(j\omega_k)$$

$$\hat{H}(-j\omega_k) = L\tilde{z}(-j\omega_k) = Lz(-j\omega_k) = H(-j\omega_k). \qquad \square$$

3.2.3 Accuracy of the Real Axis Sampling. For real axis sampling, according to (24), the V matrix is

$$\text{span}(V) = \text{span}(z(\sigma_1), z(\sigma_2), \ldots, z(\sigma_m)). \tag{28}$$

Similar to the imaginary axis sampling case, we have $\tilde{z}(\sigma_k) = z(\sigma_k)$. However, there is only $H(\sigma_k) = \hat{H}(\sigma_k)$. Generally, for all frequencies except for $\omega = 0$, there is no $H(j\omega) = \hat{H}(j\omega)$ guaranteed. In other words, the real axis sampling is not guaranteed to be accurate for any frequency except for the DC point.

3.2.4 Sampling Along the Imaginary Axis. Previously, we have shown by sampling along the imaginary axis that the frequency response of the reduced system matches that of the original system exactly. In this section, we present how the imaginary sampling is performed in a practical way.

Specifically, by performing the sampling along the imaginary axis, Z, as defined in (16), is a complex matrix. As explained in Section 3.2.2, for every sampling at $j\omega_k$, we also need to sample $-j\omega_k$. Actually, it is not necessary to perform the sampling at both points. Due to the conjugate property, if we have sampled $j\omega_k$, $k = 1, 2, \ldots, m$ and generated the corresponding Z, we only need to introduce Z^* as the conjugate of Z and form a new complex subspace $Z_c = [Z \;\; Z^*]$. Also assume W is an identity matrix, without loss of generality, we have new versions of (15) and (17) as

$$\hat{X}_c = Z_c Z_c^H \tag{29}$$

and

$$Z_c = V_c S_c U_c. \tag{30}$$

As a result, there is

$$\hat{X}_c = V_c S_c U_c U_c^H S_c V_c^H = V_c S_c^2 V_c^H, \tag{31}$$

which is the eigendecomposition of the approximated Gramian $\hat{X}_c$. In contrast to the complex $\hat{X}$ in (15), $\hat{X}_c$ here is a symmetric real matrix, even though Z_c is complex. According to the properties of the eigendecomposition, V_c should be a real unitary matrix, as it is the eigenspace of a symmetric real matrix, that is, $V_c^H = V_c^T$, which will generate the real reduced systems, even if we sample along the imaginary axis.

ALGORITHM 1. CVSTBR algorithm

Input: Circuit: G, C, B, L; sample points: ω_k, $k = 1, 2, ..., m$; the reduced order: q.
Output: Reduced system matrices: $\hat{G}$, $\hat{C}$, $\hat{B}$, $\hat{L}$.

1. Solve $z_k = (G + j\omega_k C)^{-1}B$ for $k = 1, 2, ..., m$.
2. Combine all the solved z_k to form Z.
3. Construct the complex Gramian subspace $Z_c = [Z, Z^*]$.
4. Perform economic SVD on Z_c and obtain the left singular matrix V_c and singular value matrix S. Keep only q dominant columns of V_c.
5. Rotate V_c to the real axis through multiplying the kth column of V_c with $exp(-j\phi_k)$, where ϕ_k is the phase of the kth column in the complex plane.
6. Build the reduced model $\hat{G}$, $\hat{C}$, $\hat{B}$, $\hat{L}$ using the projection matrix V_c and V_c^T.

Note that we do not perform the eigendecomposition to compute V_c, instead we use SVD on Z_c in order to save the computation time. However, V_c is still not a real matrix in general. This is because the singular vectors of the SVD are not unique in the complex plane, although the singular values are uniquely computed [Trefethen and Bau 1997]. For instance, for SVD of matrix $\mathcal{A} \in \mathbb{C}^{n\times n}$, $\mathcal{A} = VSU^T = \sum_{i=1}^{n} \sigma_i v_i u_i^T = \sum_{i=1}^{n} \sigma_i (v_i e^{j\phi})(u_i^T e^{-j\phi})$. In other words, singular vector v_i and u_j can rotate with the same angle but in the opposite directions without changing the subspace. Specifically, for the kth column v_k in V_c, multiplying with $exp(-j\phi_k)$ will rotate it to the real axis. The phase ϕ_k is the angle between v_k and the real axis.

The new complex-valued sampling-based reduction method, called *CVSTBR*, is summarized in Algorithm 1.

3.3 Residual-Based Error Estimator and Its Relationship with Imaginary Axis Sampling

We have introduced the new imaginary axis sampling scheme *CVSTBR* which matches the frequency response of the original system at the sampled frequencies. However, *CVSTBR* works in a static way with the provided sample frequencies, and an adaptive sample point selecting scheme is demanded. In order to perform the adaptive sampling, we need an error indicator to guide the adaptive sample point picking process. Calculating the errors explicitly requires expensive solving of the original model at many points, which should be avoided if possible. In this section, we present a good error estimator based on the residual of the reduced system. Combined with the imaginary sampling scheme, the residual error estimator is zero at the sampled frequencies and is good for guiding the adaptive sampling process.

Remined in (23), we have the reduced state approximation $\tilde{z}(s) = Vz_r(s) \approx z$. Plugging it into the original system with identity input

$$Gz(s) + sCz(s) = B \tag{32}$$

leads to the error matrix, which is called the *residual* matrix in this article:

$$\begin{aligned} R(s) &= GVz_r(s) + sCVz_r(s) - B \\ &= G\tilde{z}(s) + sC\tilde{z}(s) - B \end{aligned} \tag{33}$$

where $R(s) \in \mathbb{C}^{n\times p}$. Notice that if $\tilde{z}(s) = Vz_r(s)$ is very close to $z(s)$, the residual should be very small. As a result, the norm of $R(s)$, $||R(s)||$ can serve as a good error indicator for the reduced model.

We remark that $R(s)$ is not a dimensionless quantity for each element in it. Each element in it actually represents either node voltage or branch current residual. And each column of $R(s)$ can be viewed as normalized error at all the nodes/branches of the

system (18) stimulated by the impulse response applied at one port. As a result, the residual matrix $R(s)$ can be used as the input-normalized error indicator.

From Theorem 1, we have the following corollary.

COROLLARY 3.2. *The reduced system has zero residual at the sample points in the complex plane.*

PROOF. If s_k is sampled, from Theorem 1, there is $\tilde{z}(s_k) = z(s_k)$. By plugging it into (33), we have

$$\begin{aligned} R(s_k) &= G\tilde{z}(s_k) + s_k C\tilde{z}(s_k) - B \\ &= Gz(s_k) + s_k Cz(s_k) - B \\ &= 0. \end{aligned}$$

□

Furthermore, we would like to show how the residual is related to the imaginary sampling. If the sampling is performed on the imaginary axis, the residual in (33) becomes

$$\begin{aligned} R(j\omega) &= GV_c z_r(j\omega) + j\omega CV_c z_r(j\omega) - B \\ &= G\tilde{z}(j\omega) + j\omega C\tilde{z}(j\omega) - B, \end{aligned} \tag{34}$$

and there is a corollary similar to Corollary 3.2.

COROLLARY 3.3. *The reduced system residual is zero at the sampled frequency points if the sampling is performed on the imaginary axis.*

PROOF. If we have sampled the frequency ω_k, there are $\tilde{z}(j\omega_k) = z(j\omega_k)$ and $\tilde{z}(-j\omega_k) = z(-j\omega_k)$. As a special case of sampling in the complex plane, we have

$$\begin{aligned} R(j\omega_k) &= G\tilde{z}(j\omega_k) + j\omega_k C\tilde{z}(j\omega_k) - B \\ &= Gz(j\omega_k) + j\omega_k Cz(j\omega_k) - B \\ &= 0 \\ R(-j\omega_k) &= G\tilde{z}(-j\omega_k) - j\omega_k C\tilde{z}(-j\omega_k) - B \\ &= Gz(-j\omega_k) - j\omega_k Cz(-j\omega_k) - B \\ &= 0. \end{aligned}$$

□

Besides showing small values around the sampled frequencies, the residual should also reveal the errors at the unsampled frequencies to guide the adaptive sampling process. We now discuss why sampling using residual as the error estimator leads to good results. Assume we have sampled m points $\{\omega_1, \omega_2, \ldots, \omega_m\}$ of the n dimenstional system, the resulting reduced model has the order $2mp$. We have already shown at the sampled frequency point, for example, ω_k, the approximate impulse response $\tilde{z}(\pm j\omega_k)$ equals the original impulse response $z(\pm j\omega_k)$ such that the reduced system has the exact frequency response as the original system. Consider a frequency far away from the sampled frequencies, for example at ω_l, the approximate impluse response $\tilde{z}(\pm j\omega_l)$ lays inside the subspace $\text{span}(V_c) = \text{span}(z(\pm j\omega_1), z(\pm j\omega_2), \ldots, z(\pm j\omega_m))$, which is a $2mp$ dimensional subspace inside the n dimensional space. Since ω_l is not sampled, the original impulse response $z(\pm j\omega_l)$ generally does not equal $\tilde{z}(\pm j\omega_l)$ in this case. It can be decomposed into two components, one equals $\tilde{z}(\pm j\omega_l)$ inside $\text{span}(V_c)$, the other one equals $(z(\pm j\omega_l) - \tilde{z}(\pm j\omega_l))$ inside the $n - 2mp$ dimensional left null space of V_c. The component lays inside the left nullspace causes the error of the reduced system, and is revealed through the residual as $R(\pm j\omega_l) = (G \pm j\omega_l C)(z(\pm j\omega_l) - \tilde{z}(\pm j\omega_l))$. If $||R(\pm j\omega_l)||$ is a local peak, it means $||z(\pm j\omega_l) - \tilde{z}(\pm j\omega_l)||$ is relatively large. In this case, we need to sample the frequency ω_l to make the new subspace $\text{span}(V_c)$ include the component of $z(\pm j\omega_l)$ inside the left null space of the previous V_c.

3.4 Adaptive Sample Point Placement

We now demonstrate how to place the sample points along the imaginary axis in an adaptive way with the residual-based error estimation.

Before presenting the new method, we show some theoretical results regarding the residual errors and the complex-valued sampling.

THEOREM 2. *Given sufficient sample points in an imaginary axis, the residual function defined in* (34) *will become monotone-decreasing.*

PROOF. Based on Corollary 3.1, it is known that at the sample points, the responses of the reduced model exactly match those of the original model. Their residuals are all zero. The sampling approach can be viewed as the multipoint complex Krylov subspace method where only the zero-th moment is matched at every frequency point. For any frequency point $j\omega_k$, the reduced state response in (18) can be written in the moment form as

$$z_r(j\omega_k + \delta) = (I + \delta M_k + \delta^2 M_k^2 + ...)\mathbf{r}_k \tag{35}$$

where $M_k = -(\hat{G} + j\omega_k\hat{C})^{-1}\hat{C}$ and $\mathbf{r}_k = (\hat{G} + j\omega_k\hat{C})^{-1}\hat{B}$, $\delta \in \mathbb{C}$ is a small complex value change from $j\omega_k$.

If we sample sufficiently, δ, which can be viewed as the half distance between two adjacent points, will be small enough. As a result, the error between $z_r(j\omega_k)$ and $z_r(j\omega_k + \delta)$ can be approximated as $\delta M_k\mathbf{r}_k$. So the error will go down monotonically with δ when δ is small enough.

We now look at the residual $R(j\omega_k + \delta)$. Notice that we have $R(j\omega_k) = 0$ according to Corollary 3.3. If we ignore the second-order terms, we have

$$\begin{aligned} R(j\omega_k + \delta) &= (G + (j\omega_k + \delta)C)V_c z_r(j\omega_k + \delta) - B \\ &= R(j\omega_k) + \delta(GV_cM_k + j\omega_k CV_cM_k + CV_c)\mathbf{r}_k \\ &= \delta(GV_cM_k + j\omega_k CV_cM_k + CV_c)\mathbf{r}_k. \end{aligned}$$

As a result, the residual $R(j\omega_k + \delta)$ will decrease monotonically with δ when δ is small enough. □

Theorem 2 states if a point is getting close enough to a sampled point, its residual error becomes controllable. As a result, the proposed method can always achieve the error bound by sufficient sampling. But practically, it is not necessary to densely sample over all frequencies. Instead, we develop a new adaptive sampling algorithm in which sample points are added dynamically to frequencies having more poles near the imaginary axis, and thus more likely to show large residuals.

The new algorithm as demonstrated in Figure 2 has two iterative steps. At step (a) in Figure 2, *WBMOR* first builds a reduced model using the minimum sample points picked in the previous iterations. Then it tests all the candidate points, which are initially evenly distributed in the frequency range (in log scale). Residuals are computed at all the candidate points (they need to be computed or evaluated only once). At step (b), *WBMOR* drops all the satisfied candidate points whose residuals are smaller than the threshold from the candidate set. And for each unsatisfied point, two of its adjacent middle points are inserted to the candidate set. The step of inserting more samples is optional if the initial candidate points are sufficiently dense. At the same time, *WBMOR* takes the *peak points*, which are the unsatisfied candidate points and are local maximums in residual values, as the sample points. The peak points are dropped from the candidate set, as they have guaranteed zero residuals in the next iteration. The whole process continues until there are no unsatisfied candidates left.

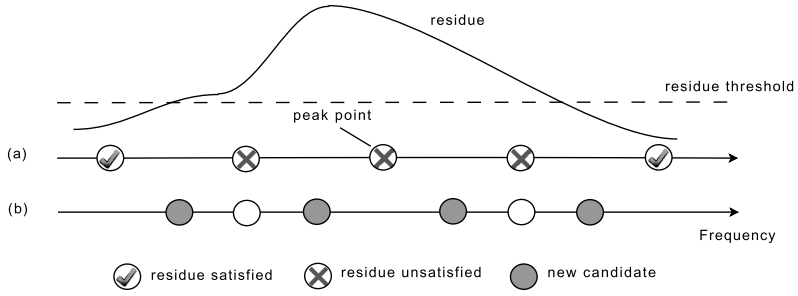

Fig. 2. An illustrative example for the *WBMOR* adaptive sampling scheme.

For most of the time, there are several peak points in one iteration. Thus the number of sample points selected in one iteration is not limited to one. This greatly enhances the convergence speed. Moreover, the peak points are usually not close to each other, which means they are relatively independent. This enables us to sample them all at the same time without causing over-sampling problems.

The dynamic sample insertion in step (b) guarantees *WBMOR* can generate a good model even when the initial candidate points are sparsely or badly distributed. This is important because sometimes there is too little information about the original model to guide the placement of the initial candidate points. Although the dynamic sample insertion in step (b) can be enabled to avoid missing the potential range with large errors, we notice that if the initial candidates are placed densely enough (around 100 points per decade is enough in our experiments), *WBMOR* will still obtain very good results with cheap computational cost without dynamic sample insertion.

In order to achieve a more efficient reduction, after the complex SVD on Z_c in (30), we select the dominant singular values and corresponding singular vectors in V_c based on a user set SVD threshold th_{svd} (th_{svd} is defined as the threshold for the weight of the trivial singular values).

3.5 WBMOR Algorithm Flow and Analysis

The whole *WBMOR* flow is shown in Algorithm 2.

The further truncation based on the singular values is performed after the iteration. As a result, SVD is only needed in the last iteration, and we use QR factorization instead of SVD in the loop to accelerate the algorithm. In addition, incremental QR algorithms, which only orthonormalize the newly sampled columns, can also be used to further save CPU time.

We now take a look at the computational cost of *WBMOR*. Most CPU time is spent on the residual matrix computation (34), since it is required for each candidate point in every iteration. GV_c and CV_c in (34) is calculated only once in an iteration because they are the same for every candidate and only vary over iterations. Also, $z_r(s)$ can be solved cheaply from the reduced system at each candidate. Thus, the cost of (34) at one candidate point is mainly the multiplication cost of $GV_c \cdot z_r$ and $CV_c \cdot z_r$, which is $O(npq)$ where n is the size of the original model, p is the port number, and q is the reduced model size at the current iteration. The cost for each candidate increases as the reduced model size q increases during the iterations. However, the number of candidate points will decrease as all the candidates smaller than the threshold are dropped. Thus, the later iteration will cost less as the algorithm proceeds. Moreover, *WBMOR* takes peak candidate points as sample points, whose number is usually larger than one. This greatly improved the convergence speed, and usually makes *WBMOR* stop in less than 10 iterations. In summary, although more expensive than the

ALGORITHM 2. WBMOR algorithm

Input: Circuit: G, C, B, L; frequency range: ω_{min}, ω_{max}; the residue threshold: th_{res} and the SVD threshold: th_{svd}.

Output: Reduced system matrices: $\hat{G}$, $\hat{C}$, $\hat{B}$, $\hat{L}$.

1. Place initial candidate points at a reasonable density in the given frequency range $[\omega_{min}, \omega_{max}]$ evenly in log scale.
2. Obtain the initial reduced model by sampling at ω_{min} and ω_{max} using the *CVSTBR* algorithm.
3. While max residue $> th_{res}$, do
4. Calculate the residues at the candidate points using (34).
5. Scan all the candidate points. If residue norm $< th_{res}$, drop this point; otherwise (optional), add two middle points between the current point and two adjacent points as new candidate points.
6. Add the candidates which are peak points with excessive residues as the sample points. Use the *CVSTBR* algorithm to compute the new reduced model with all the selected sample points.
7. Take the singular subspace V_c from the *CVSTBR* in the last iteration. Only keep its dominant singular vectors according to th_{svd}.
8. Build the real reduced model $\hat{G}$, $\hat{C}$, $\hat{B}$, $\hat{L}$ using the projection matrix V_c and V_c^T.

static sampling-based methods, *WBMOR* still has a relatively low computational cost considering its high accuracy.

4. EXPERIMENTAL RESULTS

4.1 Implementation and Settings

The proposed method, *WBMOR*, together with the resampling method [Silveira and Phillips 2006], and the recently proposed adaptive sampling-based reduction method ARMS [Villena and Silveira 2009], as well as other mentioned sampling methods have been implemented in Matlab 7.0. All the experimental results are collected on a Linux workstation with Intel Quadcore Xeon CPU with 2.99Ghz and 16GB memory. The *WBMOR* method has been integrated into the *UC Riverside Model Order Reduction Tool Suite (UiMOR)* [Tan et al. 2010], which is available for download online [uim].

For the proposed *WBMOR* method, the default setting is as follows: the residual threshold is chosen to be 0.1 and the number of initial points per decade is 100, and we only drop candidate points without adding new ones. The SVD threshold th_{svd} is 10^{-7}.

For the resampling method, the number of points in the pool is 20, the number of reduced models used in each iteration is 10, the size of the reduced model is 20. We also used 20 search points, 1/3 of which will be replaced. One point in the pool will be substituted in every iteration. We also implemented the speedup techniques such as efficient construction of projectors and heuristic search [Silveira and Phillips 2006].

For the ARMS method, however, many implementation details were not given in Villena and Silveira [2009]. In our implementation and comparison, it shares the same initial sampling candidate set with *WBMOR*. We apply QR factorization to do the orthogonalization to generate the residuals. We may argue that the incremental QR algorithm is more efficient. However, incremental QR needs to store the large and dense orthogonalized basis (the Q matrix in QR factorization) for every candidate point, and the resulting algorithm will soon become memory-limited.

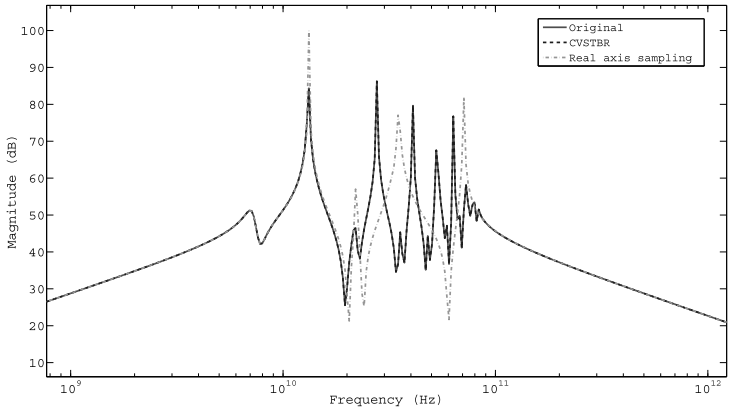

Fig. 3. Accuracy comparison of imaginary axis sampling and real axis sampling methods on TL.

Since our comparison includes both statical sampling methods and adaptive methods with different stop criteria, for a fair comparison, we have to keep the same or similar size of the final reduced order for all methods for an accurate comparison.

The benchmark circuits are published benchmarks such as the transmission line (TL) model, the PEEC model from Chahlaoui and Dooren [2002], which are available online for download. Some additional RLC circuits are also used for further accuracy and scalability comparisons.

We first show that the results of the static complex-valued sampling-based reduction method on the transmission line model example. We then show the results of the adaptive *WBMOR* method on the TL, PEEC models. Finally, the runtime and maximum error of each method are evaluated using more RLC circuits.

4.2 Comparison of Complex-Valued Sampling Scheme and Real-Valued Sampling Scheme

First, we would like to show the accuracy of *CVSTBR*. Note that in order to generate the same dimension of Z_c, the real sampling method has to use twice the sample points. In the example, the real sampling method uses 100 sample points that are uniformly distributed in the given frequency range, while the *CVSTBR* picks only half of these points, that is, 50 points. After SVD, both of the two methods keep 50 dominant columns of the left singular matrix V_c as the projection matrix, and thus the size of the two final reduced models is 50. All the samples are generated randomly so that the real-valued sampling method is similar to the proposed PMTBR method [Phillips and Silveira 2005]. Figure 3 shows the results of the comparison. It is clear that the model generated by the *CVSTBR* reaches a frequency response that cannot be distinguished from the original one, while the real sampling method has very large errors.

For the computational cost, since both schemes generate the same dimension of the Z_c matrix, they have the same SVD cost of Z_c. However, the real axis sampling scheme samples twice more than the *CVSTBR* to reach the same size of Z_c, and thus it is *more expensive*. To make the computational costs similar, the real axis sampling method will sample half of the points (both methods sample 50 points). Obviously, it will become more inaccurate. As a result, we can clearly observe that given the same computing cost, the complex-valued method *CVSTBR* is more accurate than the real axis sampling methods like PMTBR [Phillips and Silveira 2005]. Actually, this is the case for all other benchmarks we test for a reasonable number of samples.

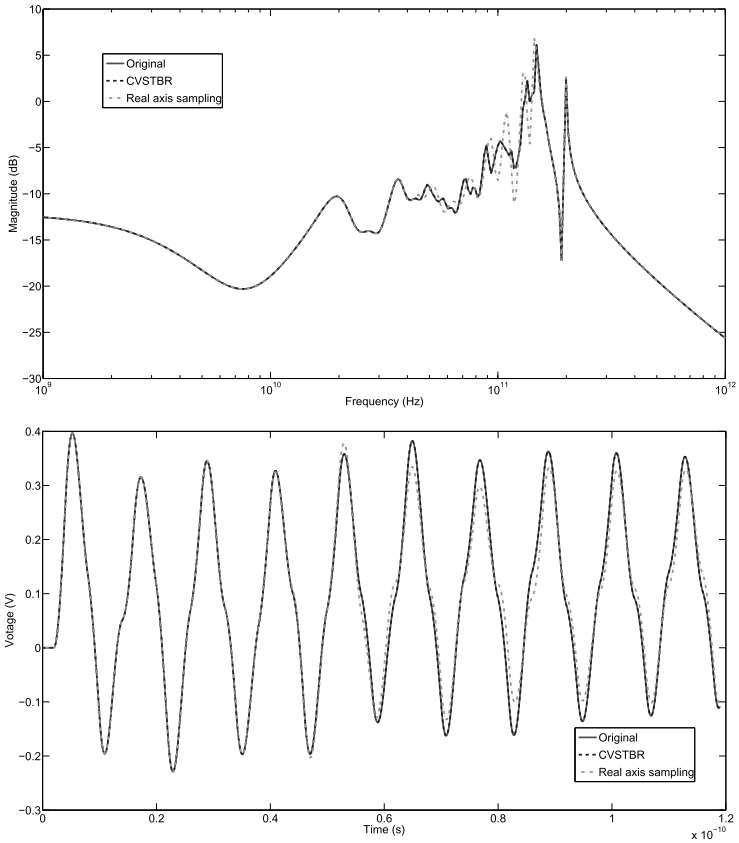

Fig. 4. Frequency response (top) and transient simulation (bottom) comparison of imaginary axis sampling and real axis sampling methods on rlc1.

Detailed analysis shows that instead of 50 sample points, only 13 points are enough to produce the 50 state reduced model in the ideal case (recall that every point generates 4 states in the final reduced model without the SVD process). This ideal case happens when we sample at only the critical points, such that, in the SVD process, all the singular vectors are important and cannot be truncated. In the next section, we show how we insert new samples around those critical regions more intelligently in the *WBMOR* algorithm.

A better accuracy comparison is in the time domain for the reduced models. We test on a RLC circuit, *rlc1*, which has a 640 states and 1 port. It has many peaks around 10^{11}Hz and is very hard to approximate. 100 sample points are used for *CVSTBR*, while 200 points for the real axis sampling method. The frequency range of interest is chosen to be $[10^9, 10^{12}]$, which covers the whole spiky area. The final size for both reduced models is kept manually at 100. A periodic 1Amp pulse current with 2×10^{-12}s rise time and 4×10^{-12}s hold time is used as the input signal at the port. The output voltage is measured from the same port. The frequency response and transient simulation results are shown in Figure 4. It can be seen from the time domain simulation, for the model generated by real-value sampling, that there obvious discrepancies at some peak and valley values in the transient response wave.

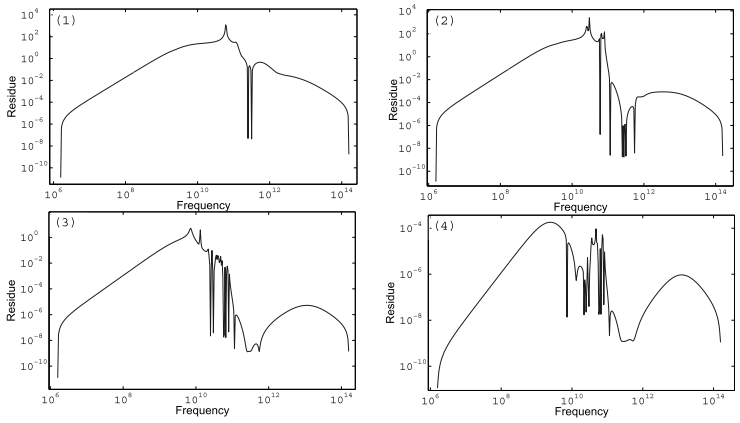

Fig. 5. The residual convergence process of *WBMOR* for four iterations.

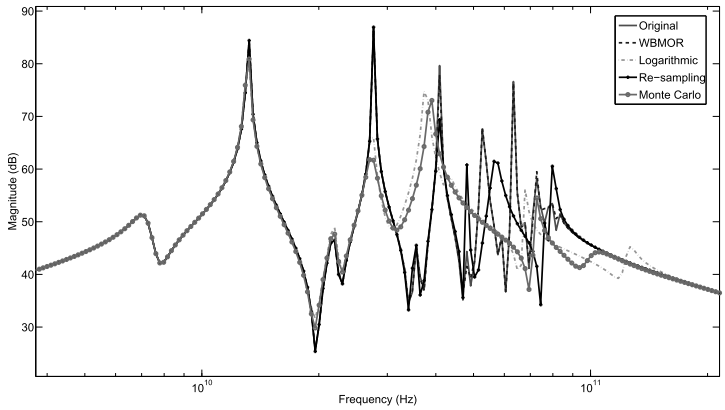

Fig. 6. Comparison with resampling method for the transmission line example.

4.3 Adaptive Process and Comparison with the Resampling Scheme

We now compare the new adaptive *WBMOR* method with several methods based on real axis sampling, including the recently proposed resampling method [Silveira and Phillips 2006], and the simple logarithmic and Monte Carlo sampling methods. In order to reach a fair comparison, we set all of the four methods with the same dimension of Z_c and the same size of final reduced system after SVD.

The first example is the transmission line model used in Section 4.2. In this case, *WBMOR* uses only 14 sample points and generates a reduced model of dimension 37. Figure 5 shows the residual convergence process of our *WBMOR* during four iterations. The maximum residual of the reduced system drops very fast from the value, as large as 10^4 to below 10^{-3} after four iterations. From Figure 6, it is clear that the new *WBMOR* method produces a reduced model as accurate as the the static *CVSTBR* (Figure 3) with many fewer sample points and a more compact size. Even with twice the number of samples and the same size of the final reduced model, none of them capture the fine details of the original system frequency response. Among these three real sampling methods, the resampling method has the best results, which, however, gradually become less accurate after the third large peak.

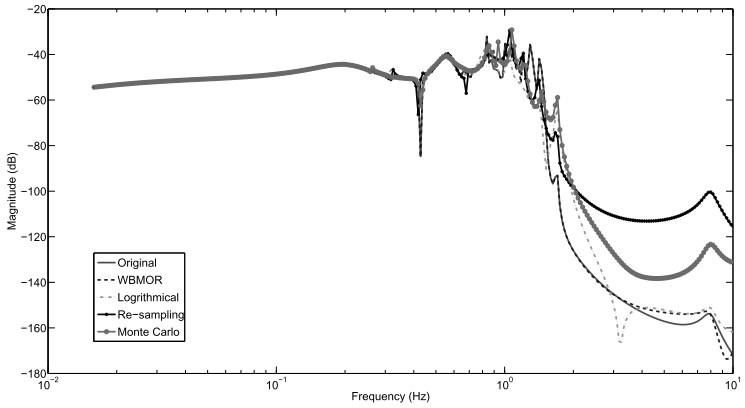

Fig. 7. Comparison with resampling method for the PEEC example.

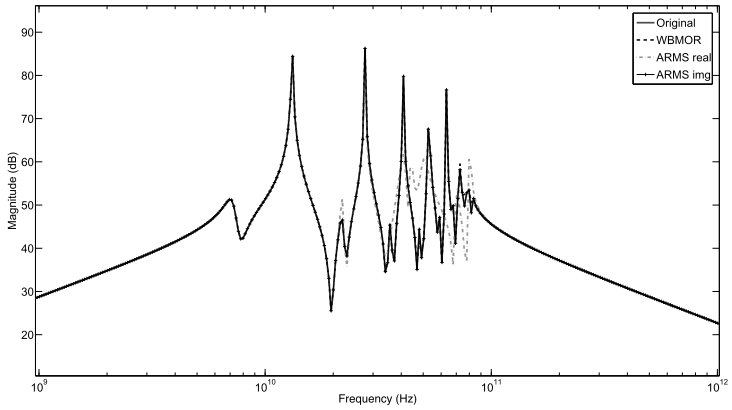

Fig. 8. Comparison with ARMS for the TL example.

Figure 7 shows the results of the widely used PEEC model. With 42 sampling points, *WBMOR* shows good results in most of the frequency bands, and only a little off around $10Hz$. This is because the responses at these frequencies are quite small (around -160dB) and are very hard to match due to numerical errors.

4.4 Comparison with the ARMS Method

The results compared with-recently proposed adaptive sampling method the ARMS, is shown in this section. Although the ARMS's implementation is based on real-valued sampling, we modify it by using the proposed imaginary axis sampling. In this way, we can see more clearly the performance of the two methods for their adaptive schemes. Specifically, in order to perform complex-valued reduction, an additional conjugate column should be put into the Gramian matrix Z in ARMS. In addition, SVD is used in the last iteration in order to generate the real-valued system.

The results on the TL model is shown in Figure 8. The order of the reduced models is 37, except for the imaginary sampling-based ARMS, which has a order of 36. Obviously, the real sampling-based ARMS method failed to find the critical samples, while the imaginary axis sampling-based ARMS does a relatively better job, although it still lacks some accuracy at certain frequency bands.

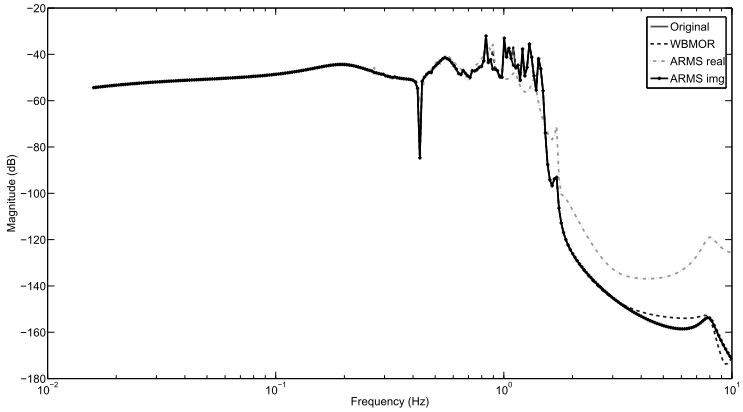

Fig. 9. Comparison with ARMS for the PEEC example.

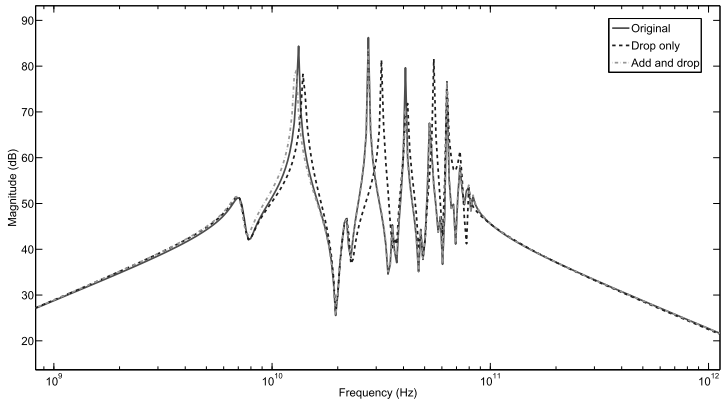

Fig. 10. Demonstrating adding the candidates feature of *WBMOR*.

Also, the comparison on the PEEC model is presented in Figure 9.

Until now, all the results of *WBMOR* were obtained by only dropping the candidates during iterations. In other words, we do not add any new candidate points. This generates good results, as shown in the previous experiments. However, at times, setting the initial set of candidate points becomes a tricky problem. We may start with a sparse candidate set, which results in a poor approximation, since the initial set may fail to cover some critical regions with large errors. This can be avoided by inserting candidate points in a dynamic and adaptive way. Take the TL model again, for example. We now start with only 10 candidates per decade instead of the 100 used previously. Figure 10 shows the results. By detailed inspection, it is clear that most of the important potential samples, which would have been missed, are correctly located through the dynamic point insertion process. The resulting *WBMOR* successfully reaches a good approximation even with the sparse initial candidate set.

4.5 CPU Runtime and Error Comparison

Finally, we report the runtime and the maximum relative errors of the examples in Section 4.3, together with several additional RLC circuits in Table I. The comparison

Table I. Scalability Comparison of Runtime and Relative Errors for *WBMOR* and Resampling Method

1	2	3	4	5	6	7	8	9	10	11	12
Ckt	Node	Port	Order	WBMOR		Re-sampling		Monte Carlo		Logarithmical	
				Time(s)	Max error	Time(s)	Max error	Time(s) Time(s)	Max error	Time(s) Time(s)	Max error
TL	256	2	37	0.9	0.3	1.4	21.6	0.09	8.5	0.09	17.7
PEEC	480	1	79	1.9	1.5	2.4	2000.1	0.71	33.1	0.70	32.9
rlc1	640	1	62	0.7	0.1	1.2	1.1	0.14	1.7	0.12	2.08
rlc2	1180	2	160	4.8	0.1	4.4	0.5	0.36	0.9	0.34	1.6
rlc3	2680	3	192	12.1	0.1	8.8	1.2	0.8	2.1	0.8	1.5
rlc4	10960	1	58	11.2	0.2	6.3	1.0	2.7	2.8	2.7	2.1

Table II. Scalability Comparison of Runtime and Relative Errors for *WBMOR* and ARMS Method

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ckt	Node	Port	WBMOR				ARMS real				ARMS img			
			order	iter	Time(s)	Max error	order	iter	Time(s)	Max error	order	iter	Time(s)	Max error
TL	256	2	37	4	0.9	0.3	38	19	5.3	11.6	40	10	10.7	2.8
PEEC	480	1	79	10	1.9	1.5	79	79	46.7	81.0	80	40	56.9	1.7
rlc1	640	1	62	11	0.7	0.1	62	62	135	1.2	62	31	126	0.1
rlc2	1180	2	160	9	4.8	0.1	160	80	1099	0.5	160	40	2402	0.1
rlc3	2680	3	192	7	12.1	0.1	192	64	2918	0.4	192	32	3705	0.05
rlc4	10960	1	58	9	11.2	0.2	58	58	2737	1.2	58	29	3273	0.2

with the ARMS method is presented in Table II. The relative errors used in the two tables are computed as

$$error(j\omega) = \frac{||H(j\omega) - \hat{H}(j\omega)||}{||H(j\omega)||}$$

at all the simulation frequency points.

We observe that the proposed method is much more accurate than all the real axis sampling-based methods. The modified imaginary axis-based ARMS has almost the same accuracy as *WBMOR*. We notice that the ARMS computing costs are much higher than the other two methods. The main computing costs are the QR operations. We need to do this for every remaining candidate point in each iteration, which is expensive for a large candidate pool. We can argue that incremental QR will be better off in this case. We actually implemented incremental QR in ARMS. But we need to store all the previous Q_i and R_i for each remaining candidate point. The sizes of Q_i and R_i grow during iterations, and their numbers are large due to the large number of initial candidate points (100 in our implementation). As a result, we soon run out of memory for some test cases.

5. CONCLUSION

We have proposed a novel model order-reduction method, *WBMOR*, for the wide frequency band modeling of interconnect circuits. *WBMOR* explicitly computes the exact residual errors to guide the sampling process in an adaptive way. We showed that by sampling along the imaginary axis and performing a new complex-valued sampling-based reduction, the reduced model will match exactly with the original model at the sample points. Theoretically, the proposed method can achieve the error bound over a given frequency range with sufficient sampling. Practically, we designed an adaptive

scheme to help designers choose the best order of the reduced model for the given frequency range and error bound. We compared several sampling schemes such as Monte Carlo, logarithmic, and the recently proposed resampling methods. Experimental results on a number of RLC circuits show that *WBMOR* is much more efficient than all the other sampling methods, including the recently proposed resampling and ARMS schemes with the same reduction orders. Compared with the real-valued sampling methods, the complex-valued sampling method is more accurate for the same computational cost.

REFERENCES

ANTOULAS, A. C. 2005. *Approximation of Large-Scale Dynamical Systems*. SIAM.

CHAHLAOUI, Y. AND DOOREN, P. V. 2002. A collection of benchmark examples for model reduction of linear time invariant dynamical systems. SLICOT working note 2002-2.

CHIPROUT, E. AND NAKHLA, M. S. 1995. Analysis of interconnect networks using complex frequency hopping. *IEEE Trans. Comput.-Aid. Des. Integrat. Circuits Syst. 14*, 2, 186–200.

FELDMANN, P. AND FREUND, R. W. 1995. Efficient linear circuit analysis by Pade approximation via the Lanczos process. *IEEE Trans. Comput.-Aid. Des. Integrat. Circuits Syst. 14*, 5, 639–649.

GALÁNTAI, A. 2004. *Projectors and Projection Methods*. Kluwer Academic, Amsterdam.

ISERLES, A. 1996. *A First Course in the Numerical Analysis of Differential Equations* 3rd Ed., Cambridge University Press.

KAMON, M., WANG, F., AND WHITE, J. 2000. Generating nearly optimally compact models from Krylov-subspace based reduced-order models. *IEEE Trans. Comput.-Aid. Des. Integrat. Circuits Syst. 47*, 4, 239–248.

MOORE, B. 1981. Principal component analysis in linear systems: Controllability, and observability, and model reduction. *IEEE Trans. Autom. Control 26*, 1, 17–32.

ODABASIOGLU, A., CELIK, M., AND PILEGGI, L. 1998. PRIMA: Passive reduced-order interconnect macromodeling algorithm. *IEEE Trans. Computer-Aid. Des. Integrat. Circuits Syst.* 645–654.

PHILLIPS, J. R. AND SILVEIRA, L. M. 2005. Poorman's TBR: A simple model reduction scheme. *IEEE Trans. Comput.-Aid. Des. Integrat. Circuits Syst. 24*, 1, 43–55.

PHILLIPS, J. R., DANIEL, L., AND SILVEIRA, L. M. 2003. Guaranteed passive balancing transformation for model order reduction. *IEEE Trans. Comput.-Aid. Des. Integrat. Circuits Syst. 22*, 8, 1027–1041.

SILVEIRA, L. M. AND PHILLIPS, J. R. 2006. Resampling plans for sample point selection in multipoint model-order reduction. *IEEE Trans. Comput.-Aid. Des. Integrat. Circuits Syst. 25*, 12, 2775–2783.

SILVEIRA, L. M., KAMON, M., ELFADEL, I., AND WHITE, J. 1996. A coordinate-transformed Arnoldi algorithm for generating guaranteed stable reduced-order models of RLC circuits. In *Proceedings of the International Conference on Computer Aided Design (ICCAD)*. 288–294.

TAN, S. X.-D., WANG, H., AND YAN, B. 2010. UiMOR – UC Riverside model order reduction tool for post-layout wideband interconnect modeling. In *Proceedings of the International Conference on Solid State and Integrated Circuit Technology (ICSICT)*.

TREFETHEN, L. N. AND BAU III, D. 1997. *Numerical Linear Algebra*. (SIAM).

UIMOR. UC Riverside Model Order Reduction Tool Suite. http://www.ee.ucr.edu/stan/project/uimor/uimor_main.htm.

VASILYEV, D. AND WHITE, J. 2005. A more reliable reduction algorithm for behavioral model extraction. In *Proceedings of the International Conference on Computer Aided Design (ICCAD)*. 813–820.

VILLENA, J. AND SILVEIRA, L. 2009. Arms–Automatic residue-minimization based sampling for multi-point modeling techniques. In *Proceedings of the IEEE/ACM Design Automation Conference (DAC)*. ACM, New York.

WANG, H., TAN, S. X.-D., AND CHEN, G. 2010. Wideband reduced modeling of interconnect circuits by adaptive complex-valued sampling method. In *Proceedings of the Asia South Pacific Design Automation Conference (ASPDAC)*. 31–36.

WANG, N. AND BALAKRISHNAN, V. 2005. Fast balanced stochastic truncation via a quadratic extension of the alternating direction implicit iteration. In *Proceedings of the International Conference on Computer Aided Design (ICCAD)*. 801–805.

WILLCOX, K. AND PERAIRE, J. 2002. Balanced model reduction via the proper orthogonal decomposition. *AIAA Journal*.

YAN, B., TAN, S. X.-D., LIU, P., AND MCGAUGHY, B. 2007. SBPOR: Second-order balanced truncation for passive model order reduction of RLC circuits. In *Proceedings of the IEEE/ACM Design Automation Conference (DAC)*. 158–161.

ZHOU, Y. 2002. Numerical methods for large scale matrix equations with applications in LTI system model reduction. Ph.D. thesis, Rice University.

Received February 2010; revised January 2011; accepted August 2011