Fast Electromagnetics-Based Co-Simulation of Linear Network and Nonlinear Circuits for the Analysis of High-Speed Integrated Circuits

Qing He and Dan Jiao, Senior Member, IEEE

Abstract—A fast electromagnetic simulator is developed to co-simulate the linear network and nonlinear circuits in an integrated circuit system. In this simulator, the physical layout of a large-scale linear network is rigorously reduced to a single surface or a few surfaces where the nonlinear circuits are located. The reduction is done analytically, and, hence, the computational overhead is minimal. The reduced system is then split into a linear system and a nonlinear system so that both systems can be solved efficiently. The linear system of equations is solved rapidly by the time-domain layered finite-element reduction-recovery method. The nonlinear system of equations is solved by developing an efficient method. This method renders the contribution from the linear network to the nonlinear system a diagonal matrix in the Jacobian matrix, hence significantly speeding up the nonlinear solution. After the reduced system is solved, the unknowns elsewhere in the computational domain are recovered efficiently by the time-domain layered finite-element reduction-recovery method. The proposed simulator has been applied to co-simulate on-chip interconnects and CMOS transistors. Numerical results have demonstrated its accuracy and efficiency. The proposed simulator is capable of capturing the global electrical interaction between integrated circuit interconnects, package, RF/analog components, substrates, and nonlinear drivers/receivers across the full electromagnetic spectrum. In addition, it bypasses the extraction of the linear network, preserves the passivity and stability of the linear network, and captures the interaction between the linear network and nonlinear devices.

Index Terms—Co-simulation, electromagnetic simulation, fast solvers, finite-element method (FEM), integrated circuits (ICs), nonlinear circuits, time domain.

I. INTRODUCTION

The prevailing circuit simulation paradigm is heavily dominated by Simulation Program with Integrated Circuit Emphasis (SPICE) [1] and its variants. In such a paradigm, the linear network is extracted first so that their RLC-based, transmission-line-based, or scattering-parameter-based representations can be obtained. This step is known as extraction. The extracted circuit model is then stitched with nonlinear devices to simulate the entire circuit. This step is known as simulation. Due to the requirement of SPICE, even though the simulation is the end goal, the extraction still needs to be performed so that the circuit model of the interconnect and passive components, i.e., the linear network, can be generated for the use of a SPICE-based simulation.

The extraction itself is computationally intensive when the linear network is large. Furthermore, the extraction-based circuit simulation flow introduces a large number of redundant computations. For example, the linear network has to be computed for a large number of right-hand sides, i.e., excitations, for the circuit model extraction. However, in reality, only a small number of excitations could be encountered in the simulation of a large-scale integrated circuit (IC). Hence, a lot of computations involved in the extraction stage are redundant. Moreover, the passivity and stability of the circuit models obtained by extraction cannot be guaranteed. A passivity checking and enforcement procedure has to be incorporated into prevailing circuit simulation flow to ensure a stable circuit simulation, leading to both additional computational overhead and model inaccuracy. In addition, since the extraction of the linear network is performed alone, the coupling between the interconnect and nonlinear devices is not captured in the prevailing circuit simulation paradigm. As the density and complexity of interconnects increase, and the interaction between nonlinear devices and interconnects in ICs intensifies, the drawbacks inherent in an extraction-based circuit simulation flow will be exacerbated even more.

An electromagnetics-based co-simulation of the linear network and nonlinear circuits naturally bypasses the extraction because the linear network is co-simulated with the nonlinear devices directly in time domain. In addition, in an electromagnetics-based co-simulation, the redundant computation due to extra right-hand sides is avoided, the passivity and stability of the linear network is naturally preserved, and the coupling between the linear network and nonlinear devices is captured. Moreover, the electromagnetics-based co-simulation is accurate across the full electromagnetic spectrum and, hence, can be used to guide the design of digital, analog, mixed-signal, and RF ICs at very high frequencies. However, for an electromagnetics-based co-simulation to gain widespread acceptance in the very large-scale integration (VLSI) community, it has to be fast enough to be put into practical use.

The electromagnetics-based co-simulation of the linear network and nonlinear circuits has been studied in the past

Manuscript received December 31, 2009; accepted August 24, 2010. Date of publication November 11, 2010; date of current version December 10, 2010. This work was supported in part by the Office of Naval Research under Grant N00014-10-1-0482 and by the National Science Foundation under Grant 0747578 and Grant 0802178.

The authors are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: qhe@purdue.edu; djiao@purdue.edu).

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/TMTT.2010.2086590
[2]–[15]. Approaches to coupling both the first-order Maxwell’s
equations and the second-order vector wave equations with the
lumped circuit models have been developed. The field-circuit
coupling algorithm has been explored in the framework of the
finite-difference time-domain method [2]–[4], the
time-domain finite-element method (FEM) [5]–[11], and the
time-domain integral equation method [12], [13]. However,
many of these algorithms were developed for the simulation of
microwave and millimeter-wave ICs or package and board
problems. They often have not been found to be amenable to
on-chip VLSI design because of unique on-chip modeling
challenges such as conductor loss, strong nonuniformity, large
number of conductors, large aspect ratio, and large number
of nonlinear devices [16]. In addition, although a variety of
coupling algorithms have been developed, the research on
fast co-simulation methods is still in its infancy.

Recently, a family of time-domain finite-element reduc-
tion-recovery (FE-RR) methods were developed for solving
large-scale ICs and package problems [17]–[21]. These
methods can reduce the system matrix of $O(N)$ rigorously
to that of $O(M)$ for any multilayered structure, where $N$ is
the number of unknowns in the entire 3-D structure and $M$
is the number of unknowns on a single surface. Furthermore,
the reduction from $O(N)$ to $O(M)$ was achieved either via
analytical means or with negligible computational cost. The
reduced system is solved in optimal complexity, i.e., linear
complexity. The rest of the unknowns are then recovered, also
in linear complexity. The hierarchical FE-RR method [20] is
applicable to any Manhattan-type multilayered structure; the
orthogonal FE-RR method [21] further advances the method to
solve any irregularly shaped multilayer structure. The FE-RR
methods have been successfully applied to solve large-scale
IC and package design problems. For example, the matrix
resulting from a large-scale combined die-package problem
involving 333 million unknowns can be solved in 200 s on a
single 2.66-GHz Intel Xeon 5300 processor by the orthogonal
FE-RR method [21]. Despite its high capacity and efficiency,
the FE-RR methods have not taken the simulation of nonlinear
circuits into consideration yet.

In this work, we develop a fast electromagnetics-based
nonlinear–linear co-simulation algorithm to capture the global
electrical interaction between integrated circuit interconnects,
packages, RF/analog components, substrates, and nonlinear
drivers/receivers. In this method, we first reduce an arbitrarily
shaped 3-D multilayer structure to a single surface or a few
surfaces where the nonlinear circuits are located. The reduction
is rigorous without involving any theoretical approximation.
Furthermore, the reduction is achieved without any computa-
tional cost via analytical means by the use of the layered FE-RR
method [17]. We then separate the linear equations from the
nonlinear ones in the reduced system so that both can be solved
efficiently. The linear system is solved rapidly by the layered
FE-RR method. The nonlinear system is solved by developing
an efficient method. In this method, the contribution of the
linear system to the nonlinear system in the Jacobian matrix is
a diagonal matrix. Compared with other FEM-based co-simu-
lation methods which generate a dense submatrix in the Jacobian
matrix, the proposed method is much more computationally
efficient. After the reduced system is solved, the unknowns on
the other surfaces and in the volume are recovered efficiently by
the FE-RR methods.

II. DERIVATION OF THE SYSTEM OF EQUATIONS FOR
CO-SIMULATION

Consider an integrated system consisting of a linear network
and nonlinear circuits shown in Fig. 1. All of the interconnects,
packages, and passive components belong to the linear network.
All of the nonlinear devices are in the nonlinear block. Here, we
derive the system of equations that governs the co-simulation of
the linear network and nonlinear circuits.

A. System of Equations of the Linear Network

The physical phenomena in the linear network are governed by
Maxwell’s Equations, which suggest
\[
\nabla \times \left[ \mu_r^{-1} \nabla \times \mathbf{E}(\mathbf{r}, t) \right] + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} + \mu_0 \sigma \frac{\partial \mathbf{J}(\mathbf{r}, t)}{\partial t} = -\mu_0 \frac{\partial \mathbf{H}(\mathbf{r}, t)}{\partial t} \quad (1)
\]
where $\mathbf{E}$ is the electric field, $\mu_0$ is free-space permeability, $\mu_r$
is relative permeability, $\varepsilon_0$ is permittivity, $\sigma$ is conductivity, $\mathbf{J}$ is
current density, and $\mathbf{H}$ denotes a point in a 3-D space.

A time-domain finite-element solution of (1) and (1)’s
boundary conditions results in the following system of ordinary
differential equations [22]:
\[
T \frac{du}{dt} + R \frac{du}{dt} + S u + w = \frac{d\mathbf{I}}{dt} \quad (2)
\]
in which $T$, $R$, and $S$ are square matrices, $u$ is the unknown field
vector, $w$ is a vector related to the absorbing boundary condition,
and $\mathbf{I}$ is the vector of the currents injected into the linear system.
The elements of the matrices $T$, $R$, and $S$ are given by
\[
T_{ij} = \mu_0 \varepsilon_\nu (\mathbf{N}_i, \mathbf{N}_j)_\nu V,
R_{ij} = \mu_0 \sigma (\mathbf{N}_i, \mathbf{N}_j)_\nu V,
S_{ij} = \mu_r^{-1} (\nabla \times \mathbf{N}_i, \nabla \times \mathbf{N}_j)_\nu V \quad (3)
\]
where $\mathbf{N}_i$ and $\mathbf{N}_j$ are the vector basis functions used to expand
unknown field $\mathbf{E}$ and $(\cdot, \cdot)_\nu V$ denotes a volume integration. The
elements of the current vector $\mathbf{I}$ are given by
\[
\mathbf{I}_i = -\mu_0 (\mathbf{N}_i, \mathbf{J})_\nu V. \quad (4)
\]
At the 4th edge, if a current source of magnitude $I_i$ is attached,
(4) can be evaluated as
\[
\tilde{I}_i = -\mu_0 I_i l_i \quad (5)
\]
where \( L_i \) is the length of the \( i \)th edge. In deriving (5), we assume that the vector basis \( N_i \) is normalized, and \( N_i \) is oriented along the same direction as the current source \( I_i \). (If the basis is not normalized and its direction is not aligned with the current source, (4) can be evaluated in a similar fashion.) From Fig. 1, it can be seen clearly that the current flowing into the linear system has two components: \( I_s \) which is a supply current, and \( I_{\text{inf}} \) which is injected from the nonlinear network. Hence, (5) can be written as

\[
\tilde{I}_i = -\mu_0(I_{\text{si}} + I_{\text{ni}})k_i.
\]

(6)

The voltage across the \( i \)th edge, \( V_i \), can be evaluated from \( u \) after (2) is solved. For example, if \( V_i \) is evaluated along the direction that is opposite to the \( i \)th normalized vector basis, then

\[
V_i = -I_i u_i.
\]

(7)

B. System of Equations for the Nonlinear Circuits

The nonlinear circuit shown in Fig. 1 can be modeled by

\[
I_c = f(V_c, t)
\]

(8)

where \( f \) is a nonlinear function, \( t \) is time, \( I_c \) is current, and \( V_c \) is voltage. If the nonlinear circuit is a network that consists of a number of nonlinear components, it can be analyzed by the Modified Nodal Analysis [23], which results in the following nonlinear system of equations:

\[
\tilde{G}(x) \dot{x} + \tilde{C}(x) \frac{dx}{dt} = b
\]

(9)

where the unknown vector \( x = [V_c, i]^T \), in which \( V_c \) is a vector of node voltages and \( i \) is a vector of branch currents flowing through inductors and voltage sources. In (9), \( \tilde{G} \) denotes a nonlinear mapping from \( x \) to \( b \), and \( \tilde{C} \) denotes a nonlinear mapping from \( dx/dt \) to \( b \). Both \( \tilde{G} \) and \( \tilde{C} \) can be time-dependent. The nonlinear model (8) can be viewed as a special case of (9).

C. System of Equations at the Interface Between the Linear Network and Nonlinear Circuits

From Fig. 1, it can be seen clearly that, at the interface between the linear network and nonlinear circuits, the following system of equations satisfies

\[
I_{\text{nl}} + I_c = 0.
\]

(10)

D. Combined System of Equations for Co-Simulation

With equations developed above in the three subsections, we complete the system of equations that governs the integrated nonlinear–linear system shown in Fig. 1. To accurately obtain the transient response of such a system, we need to co-simulate (2), (9), and (10).

It is worth mentioning that, if the function \( f \) in (8) is linear and time-independent, the co-simulation of (2), (9), which is reduced to (8), and (10) is straightforward in an FEM-based solver. To explain, the functions \( f \) of a constant and linear resistor of resistance \( R \), an inductor of inductance \( L \), and a capacitor of capacitance \( C \) are

\[
f_R = \frac{V_c}{R}, \quad f_L = \frac{1}{L} \int V_c dt, \quad f_C = C \frac{dV_c}{dt}
\]

(11)

respectively. By substituting (10) and (11) into (6), and employing (7), it can be readily derived from (2) that, if the lumped \( R \), \( L \), and \( C \) are attached to the \( i \)th edge in an FEM-based mesh, they only contribute to the \( i \)th diagonal element of \( \mathbf{R}, \mathbf{S}, \) and \( \mathbf{T} \), which amounts to adding \( \mu_0 I_i^2 / R \) to \( \mathbf{R}_{ii} \), \( \mu_0 I_i^2 / L \) to \( \mathbf{S}_{ii} \), and \( \mu_0 I_i^2 C \) to \( \mathbf{T}_{ii} \) respectively.

If the circuit connected to the linear network is nonlinear, the aforementioned approach becomes highly computationally expensive because the entire system matrix has to be factorized and solved at each time step, as the system is time-dependent and nonlinear. In the following section, we propose an efficient algorithm to co-simulate the linear network and nonlinear circuits, i.e., the combined system of (2), (9), and (10).

III. EFFICIENT CO-SIMULATION ALGORITHM

Compared with the number of edges present in an FEM-based discretization of a 3-D computational domain, in general, the edges that are attached to the nonlinear circuits are orders of magnitude smaller. For example, in a state-of-the-art integrated circuit, the transistors are only attached to the bottommost metal layer. Therefore, it becomes meaningful to first reduce the 3-D system to a single surface or a few surfaces where the nonlinear circuits are attached. The reduced system can be orders of magnitude smaller than that of the original 3-D system because the number of nonlinear circuits typically is much smaller than that of the linear network. To minimize the reduction cost, we will invoke the time-domain layered FE-RR method [17] to analytically perform the reduction. The detailed procedure of the reduction is given in Section III-A.

After we obtain a reduced system, we divide the unknowns into two groups: one is associated with the linear network, and the other is associated with the nonlinear circuit. This procedure is illustrated in Section III-B. We then solve the linear equations by the layered FE-RR method, and the nonlinear equations by an efficient method which is detailed in Section III-D. After the solution of the reduced system is obtained, we recover the rest of the solutions efficiently by the layered FE-RR method.

A. Reduction of the Physical Layout of a Linear Network to the Surfaces Where Nonlinear Circuits Are Located

Discretizing (2) by a central difference scheme, we obtain

\[
\mathbf{P} u^{n+1} = (2 \mathbf{T} - \Delta t^2 \mathbf{S}) u^n + [0.5 \Delta t^2 \mathbf{R} - \mathbf{T}] u^{n-1}
\]

\[
-\Delta t^2 u^n + \Delta t^2 \left[ \frac{dI}{dt} \right]^n
\]

(12)

in which

\[
\mathbf{P} = \mathbf{T} + 0.5 \Delta t^2 \mathbf{R}
\]

(13)

and \( \Delta t \) represents the time step used in the electromagnetics-based simulation of the linear network. The field value at the
The computational domain of a combined linear–nonlinear system is discretized into layers of triangular prism elements as shown in Fig. 2(a). Since existing integrated circuits are multi-layered structures in general, the triangular prism element based discretization is indeed natural for choice. The unknowns are then ordered layer by layer. In each layer, the unknowns associated with the solid edges are surface unknowns, and the unknowns associated with the dashed edges are volume unknowns.

Without loss of generality, assuming that the nonlinear circuits are attached to the $i$th surface. We will formulate a reduced system that only involves unknowns on the $i$th surface by using the layered finite-element reduction-recovery method [17]. The procedure is as follows. The 3-D layered system matrix is first reduced to a 2-D layered one as shown in Fig. 2(b). The 2-D layered system matrix is then reduced to a single-surface one as shown in Fig. 2(c). Once the unknowns on a single surface are solved, the unknowns on other surfaces and the unknowns in the volume can be recovered as shown in [17]–[21].

In Fig. 3, we show the system matrix for the 3-D layered system, 2-D layered system, single-layered system, and single-surface system respectively. The 3-D layered system matrix shown in Fig. 3(a) is matrix $\mathbf{P}$ shown in (12) and (13). In Fig. 3(b), matrices $\mathbf{M}_l$ ($l = 1, 2, \ldots, L$) and $\mathbf{K}_l$ ($l = 1, 2, \ldots, L$) are the submatrices formed for surface unknowns in $\mathbf{P}$. In the single-layered system shown in Fig. 3(c), matrices $\mathbf{M}''_{id}$ and $\mathbf{M}''_{iu}$ can be obtained analytically with minimal computational overhead from

\[
\begin{align*}
\mathbf{M}_2'' &= \mathbf{M}_1'' + \mathbf{M}_2'' - \mathbf{K}_1\mathbf{M}_1''^{-1}\mathbf{K}_1 \\
\mathbf{M}_3'' &= \mathbf{M}_2'' + \mathbf{M}_3'' - \mathbf{K}_2\mathbf{M}_2''^{-1}\mathbf{K}_2 \\
&\vdots \\
\mathbf{M}_{id}'' &= \mathbf{M}_{i-1}'' + \mathbf{M}_i'' - \mathbf{K}_{i-1}\mathbf{M}_{i-1}''^{-1}\mathbf{M}_{i-1} \\
\mathbf{M}_{iu}'' &= \mathbf{M}_i'' + \mathbf{M}_{i+1}'' - \mathbf{K}_{i+1}\mathbf{M}_{i+1}''^{-1}\mathbf{M}_{i+1}'' 
\end{align*}
\]
In (14) and (15), there is no need to compute the matrix inverse and matrix–matrix multiplication. This is because, in the layered FE-RR method, $M_l, \ l = 1, 2, \ldots, L$, is constructed to be linearly proportional to each other, $K_l$ ($l = 1, 2, \ldots, L$) is made linearly proportional to each other, and $M_l, l = 1, 2, \ldots, L$, is also made linearly proportional to $K_l, l = 1, 2, \ldots, L$. Therefore, $M_l, l = 1, 2, \ldots, L$, of primed quantities are also linearly proportional to $K_l, l = 1, 2, \ldots, L$. As a result, the reduction shown in Fig. 3, which is mathematically represented by (14) and (15), is achieved without any numerical calculation.

From $M_{li}$ and $M'_{li}$, the single-interface system shown in Fig. 3(d) can be obtained as follows:

$$M''_i = M'_{li} - K_i (M'_{ii})^{-1} K_i. \quad (16)$$

Again, there is no need to compute matrix inversion and matrix–matrix multiplications because $K_i$ and $M'_{ii}$ are linearly proportional to each other. From the reduction process, it can also be seen that, due to the linear proportionality between matrices $M$ and $K$, $M''_i$ essentially can be obtained by scaling the $M_i$ in the original system matrix by a coefficient. Hence, the reduced single-surface matrix preserves the same sparsity as that in the original system matrix, enabling an efficient computation.

In addition to the left-hand matrix reduction from $P$ to $M''_i$, the right-hand vector of the original system (12) is also reduced to that of the $i$th surface, following the procedure given in [17]. Since the nonlinear circuits are connected to the linear network via the current excitation vector $d\bar{I}/dt$, in the following, we show how $d\bar{I}/dt$ is reduced.

From (6), the excitation vector $d\bar{I}/dt$ in the right-hand side of (12) can be written as

$$\frac{d\bar{I}}{dt} = \frac{d\bar{I}_s}{dt} + \frac{d\bar{I}_{nt}}{dt} \quad (17)$$

which can be further written as

$$\frac{d\bar{I}}{dt} = \frac{d\bar{I}_s}{dt} - \frac{d\bar{I}_c}{dt} \quad (18)$$

by using (10). Since $\bar{I}_c$, only exists on the $i$th surface, in other words, the entries of $I_c$ are all zero except for those on the $i$th surface, the reduction of the vector $d\bar{I}_c/dt$ from that of the original 3-D problem to that of the $i$th surface does not change the values in $d\bar{I}_c/dt$. Therefore, the right-hand side of the single-surface system shown in Fig. 3(d) can be explicitly written as

$$b''_i = b''_{is} + \Delta t_e \frac{d\bar{I}_c}{dt} \quad (19)$$

where $b''_{is}$ represents what is reduced from all the terms on the right-hand side of (12) except for $d\bar{I}_c/dt$. Following the derivation of (6), the entries of $I_c$ can be derived as

$$\bar{I}_{cj} = -\mu_0 (I_{nt,j}) l_j = \mu_0 (J_{cj}) l_j. \quad (20)$$

Now, the computing task is to solve a reduced system

$$M''_i u = b''_{is} - \Delta t_e \frac{d\bar{I}_c}{dt}. \quad (21)$$

Since, in general, not all of the edges on the $i$th surface are attached to the nonlinear circuits, (21) is mixed with both linear and nonlinear equations. In Section III-B, we separate the linear equations from nonlinear ones so that both can be solved efficiently.

### B. Separate Linear Equations From Nonlinear Ones

We divide unknowns in the reduced system into two groups: one is associated with the linear network, and the other is attached by nonlinear circuits. The former group is denoted by $u_e$, and the latter is denoted by $u_c$. Then we recast the reduced system into the following system:

$$\begin{bmatrix} M''_{iee} & M''_{iec} \\ M''_{eic} & M''_{ecc} \end{bmatrix} \begin{bmatrix} u_e \\ u_c \end{bmatrix} = \begin{bmatrix} b''_{ie} \\ b''_{ic} \end{bmatrix} + \begin{bmatrix} 0 \\ -\Delta t_e \frac{d\bar{I}_c}{dt} \end{bmatrix} \quad (22)$$

in which subscript $c$ denotes quantities associated with the linear network, and $c$ denotes those associated with the nonlinear circuit.

### C. Efficient Solution of the Linear System of Equations

The system of equations represented by the first row in (22) is a linear system. It can be solved efficiently by a general sparse matrix solver since the dimension of $M''_{ee}$ is already reduced to that of a single surface, and the reduction preserves the sparsity of $M''_{ee}$. If the reduced single-surface system is still large, the hierarchical FE-RR method [20] or the orthogonal FE-RR method [21] can be employed to achieve a linear-complexity solution of $M''_{ee}$. Since (22) is a reduced system that only involves the unknowns on the $i$th surface, after (22) is solved, we recover the $u_e$ on the other surfaces and in the volume from the solution on the $i$th surface by the layered FE-RR method [17]. Again, if the hierarchical FE-RR method or the orthogonal FE-RR method is employed, the recovery complexity is also linear.

In the following, we focus on efficient solutions to the nonlinear system represented by the second row of (22).

### D. Efficient Solution of the Nonlinear System of Equations

Here, we present two methods for solving the nonlinear system of equations. The first method is useful when the number of nonlinear circuits is small. The second method is efficient for solving a large number of nonlinear circuits.

#### 1) Method I: Equation (22) can be rewritten as

$$M''_{iee} u_e + M''_{iec} u_c = b''_{ie} \quad (23)$$

$$M''_{eic} u_e + M''_{ecc} u_c + \Delta t_e \frac{d\bar{I}_c}{dt} = b''_{ic}. \quad (24)$$

Substituting (23) into (24) with $u_e$ yields

$$M_{eq} u_c + \Delta t_e \frac{d\bar{I}_c}{dt} = b_{eq} \quad (25)$$

where

$$M_{eq} = M''_{iee} - M''_{iec} M''_{ec}^{-1} M''_{iee}$$

$$b_{eq} = b''_{ie} - M''_{iec} M''_{ec}^{-1} b''_{ic}. \quad (26)$$
Assuming there are $m$ nonlinear circuits with each circuit connecting to $n$ serial edges in a finite element mesh, the dimension of $\mathbf{M}_{eq}$ is $mn \times mn$. In addition, $\mathbf{M}_{eq}$ is a dense matrix due to the coupling from the linear network to the nonlinear circuit, which is represented by the $\mathbf{M}_{eq}^{c} \mathbf{M}_{eq}^{c} \mathbf{M}_{eq}^{c}$ term in (26). In (25), the unknown vector $\mathbf{u}_c$ and $d\dot{\mathbf{I}}_c/dt$ can be written as

$$
\mathbf{u}_c = (u_{c,1,1} \cdots u_{c,1,n} \ u_{c,2,1} \cdots u_{c,2,n} \ \cdots \ u_{c,m,1} \cdots u_{c,m,n})^T
$$

$$
d\dot{\mathbf{I}}_c/dt = \left( \frac{d\dot{I}_{c,1}}{dt} \cdots \frac{d\dot{I}_{c,1}}{dt} \ \frac{d\dot{I}_{c,2}}{dt} \cdots \frac{d\dot{I}_{c,2}}{dt} \ \cdots \ \frac{d\dot{I}_{c,m}}{dt} \ \cdots \ \frac{d\dot{I}_{c,m}}{dt} \right)^T
$$

(28)

where, in $\mathbf{u}_c$, the first subscript denotes the circuit index, and the second subscript denotes the edge index; in $d\dot{\mathbf{I}}_c/dt$, since the $i$th circuit is attached across $n$ edges in serial, the $n$ edges share the same current $I_{c,i}$.

Since the current $d\dot{\mathbf{I}}_c/dt$ is a nonlinear function of time due to the connection with nonlinear and time-variant circuits, apparently, all of the equations in (25) have to be solved in a nonlinear fashion. In fact, in (25), only $m$ equations need to be solved in a nonlinear fashion. This is because each circuit is attached to $n$ edges, where there are $n$ equations for each nonlinear circuit in (25). Since the $n$ equations all share the same $d\dot{\mathbf{I}}_c/dt$ as can be seen from (28), a linear combination of these $n$ equations will zero out the $d\dot{\mathbf{I}}_c/dt$ term in $n-1$ equations, leaving only one equation to be associated with the nonlinear term $d\dot{\mathbf{I}}_c/dt$. As a result, for $m$ nonlinear circuits, among the $mn$ equations in (25), only $m$ equations need to be solved in a nonlinear fashion. There are several classical methods that can be used to find the solution of a nonlinear equation. In this work, we use Newton’s method [or Newton–Raphson method] [24], which offers a computationally efficient and numerically stable approach to solve for a nonlinear equation.

To complete the co-simulation, we solve (25) together with (9) by the Newton’s method. Using the approach described above, the linear network only contributes a dense matrix of size $m$ by $m$ to the Jacobian matrix used in the Newton iteration rather than a dense matrix of size $m \times n$ by $m \times n$, which reduces the cost greatly when $m$ and $n$ are large.

After (25) together with (9) is solved, $\mathbf{u}_c$ is obtained. By substituting $\mathbf{u}_c$ into (23), we can solve for $\mathbf{u}_e$ on the $i$th surface. The $\mathbf{u}_e$ on the other surfaces and in the volume can then be recovered from the field solution on the $i$th surface.

2) Method II: In Method I, although the matrix size involved in the nonlinear simulation is reduced from $mn \times mn$ to $m \times m$, the matrix $\mathbf{M}_{eq}$ is dense. This makes the nonlinear simulation very time consuming when the number of nonlinear circuits, $m$, is large. To overcome this problem, we develop the second method as follows.

We stagger $\mathbf{u}_c$ and $\mathbf{u}_e$ to solve (22). This leads to

$$
\mathbf{M}_{eq}^{c} \mathbf{u}_{c,k+1} + b_{te}^{c} - \mathbf{M}_{eq}^{c} \mathbf{u}_{c,k} = 0
$$

$$
\mathbf{M}_{eq}^{e} \mathbf{u}_{e,k+1} + \Delta \mathbf{G}_{e}^{c} \frac{d\dot{\mathbf{I}}_{c,k+1}}{dt} = b_{te}^{c} - \mathbf{M}_{eq}^{c} \mathbf{u}_{c,k+1}
$$

(29)

where the subscript $k$ denotes the iteration index. We perform a few iterations of (29) until it converges. Clearly, when (29) converges, the solution of (29) is the same as that of (25). By doing so, we keep the sparsity of the original matrices $\mathbf{M}_{eq}^{c}$ and $\mathbf{M}_{eq}^{e}$. Moreover, $\mathbf{M}_{eq}^{c}$ is a diagonal matrix because the edges which are connected to circuits generally do not belong to the same finite element. As a result, the contribution from the linear network to the Jacobian matrix is a diagonal matrix when solving the second row of (29) together with (9) by the Newton’s method. Hence, the resultant nonlinear solution is efficient. The computational overhead is a number of iterations of (29). Since the number of iterations is small, less than 15 on average in our simulation of realistic on-chip circuits, the computational overhead is negligible. To help better understand the fast convergence of (29), we point out that the staggered marching shown in (29) is equivalent to using the diagonal blocks of the system matrix of (22), $\mathbf{M}_{eq}^{c}$ and $\mathbf{M}_{eq}^{e}$, as the preconditioner to solve (22). Since the off-diagonal blocks of (22) are extremely sparse, the diagonal-block-based preconditioner is very effective.

E. Performance Analysis

The numerical procedure of the proposed method for co-simulation can be summarized as follows.

Step 1) Reduce the matrix equation in (12) to the $i$th surface where the nonlinear circuit is located by the scheme described in Section III-A.

Step 2) Recast the resultant $i$th surface matrix equation (21) to system (22).

Beginning of the Time Marching Process

With the value of $\mathbf{u}_c$ and $\mathbf{u}_e$ at time step $n$ and $n-1$ known, perform the following steps.

Step 3) When the number of nonlinear circuits is small, solve the nonlinear algebraic differential equation (25) and (9) for $\mathbf{u}_{e,n+1}^{c}$ by the Newton–Raphson method, then solve (23) for $\mathbf{u}_{e,n+1}^{e}$. The detailed procedure is given in Section III-D1. When the number of nonlinear circuits is large, solve (29) subject to (9) by the staggered marching iteration to obtain $\mathbf{u}_{e,n+1}^{c}$ and $\mathbf{u}_{e,n+1}^{e}$. The detailed procedure is given in Section III-D2.

Step 4) Recover $\mathbf{u}_{e,n+1}^{c}$ on the other surfaces and in the volume by the approach discussed in Section III-C.

Step 5) Construct the new right-hand side of (12) for the next time step.

Go back to Step 3).

End of the Time Marching Process

The cost of Step 1) is negligible because the reduction is performed analytically by the layered FE-RR method. Step 2) does not involve any cost because it is an analytical recasting. In Step 3), if Method I described in Section III-D1 is used for solving the nonlinear system, the computational cost is to form the Schur complement matrix $\mathbf{M}_{eq} = \mathbf{M}_{eq}^{c} - \mathbf{M}_{eq}^{e} \mathbf{M}_{eq}^{c}^{-1} \mathbf{M}_{eq}^{c}$. Since the dimension of the $\mathbf{M}_{eq}^{c}$ is already reduced to that of a single surface, compared with other time-domain FEMs..
that deal with the entire 3-D problem size for co-simulating with nonlinear devices [5]–[10], the dimension of the proposed scheme is orders of magnitude smaller. Furthermore, by the FE-RR methods [17]–[21], each matrix solve of $M'_{\text{Diff}}$ has a linear complexity. Hence, the $M'_{\text{Diff}}M_{\text{Jac}}^{-1}M_{\text{Diff}}$ can be obtained in $O(m^2N_u)$ complexity, where $m'$ is the number of circuit unknowns, and $N_u$ is the number of field unknowns on a single surface. The system matrix contributed by the linear network to the Jacobian matrix for solving the nonlinear system is a dense matrix of $m^2$ dimensions, where $m$ is the number of nonlinear circuits. When $m$ is large, Method II given in Section III-D2 should be used for the simulation of the nonlinear system. In this method, there is no cost for forming an additional dense matrix $M_{\text{Jac}}$. In addition, the system contributed by the linear network to the Jacobian matrix for solving the nonlinear system, $M_{\text{Jac}}^0$ is a diagonal matrix. Although a few iterations are needed in the staggered marching scheme, the iteration number is small. In Step 4), the recovery has a linear complexity based on the FE-RR methods [17]–[21]. In Step 5), the right-hand side of (12) only involves sparse matrix-vector multiplication and, hence, can also be obtained efficiently.

F. Stability Analysis

The time step required by the electromagnetics-based simulation of the linear network for maintaining stability can be different from that required by the Newton-method-based simulation of nonlinear circuits. We denote the former by $\Delta t_e$ and the latter by $\Delta t_c$ and discuss their choices in the following.

A time-domain FEM-based analysis is stable if the following condition is satisfied:

$$\Delta t_e \leq \frac{\sqrt{\lambda_{\text{max}}}}{\sqrt{\rho(\Gamma)}}$$

in which $\lambda_{\text{max}}$ is the value at which the roots of a characteristic equation start to leave the unit circle [25], and $\rho(\cdot)$ denotes the spectral radius of matrix (·). As discussed in [18], a typical time step for on-chip integrated circuit simulation is at a level of $10^{-16}$ s to maintain the stability of a central difference-based scheme. For the nonlinear circuit simulation, we utilize the SPICE-based criterion to choose the time step [26], [27]. A typical $\Delta t_c$ for state-of-the-art nonlinear devices is from the order of $10^{-9}$ s to the order of $10^{-12}$ s, which is much larger than that of the central difference-based time-domain FEM simulation of the linear on-chip system. As a result, we can use a large and variable time step for the nonlinear circuit simulation and a small fixed time step for the time-domain FEM-based simulation of the linear network to avoid unnecessary updates of the circuit simulation. For the numerical examples simulated in this paper, since the cost of the nonlinear circuit part accounts for less than 2% of the total computational cost, we used the time step dictated by the time-domain FEM as the time step for the co-simulation.

IV. NUMERICAL RESULTS

Here, we give numerical results to demonstrate the performance of the proposed co-simulation method. In the first example, we validated the proposed method on a lossless parallel-plate waveguide structure, for which the numerical results can be benchmarked with those of SPICE. According to typical on-chip circuit dimensions, the waveguide width (along $y$) was set as 1 $\mu$m, the height (along $x$) was set as 0.1 $\mu$m, and the waveguide length (along $z$) was set as 4 mm. The spatial discretization was chosen as $\Delta x = 0.1 \mu$m, $\Delta y = 0.25 \mu$m, and $\Delta z = 1 \mu$m. The dominant TEM mode was launched on the incident plane at $z = 0$. The incident wave was a sinusoidal source oscillating at 10 GHz. The first-order absorbing boundary condition was placed at the two ends of the waveguide along the length. The time step was chosen as $10^{-16}$ s. A resistor ($R = 37.7 \Omega$), an inductor (L = 1 nH), a capacitor (C = 1 pF), and a diode ($i_D(t) = I_0[e^{-\alpha t}/\sqrt{1 + \alpha^2}e^{-\alpha t}/\sqrt{1 + \alpha^2}]$, $I_0 = 10^{-14}$ A, $V_0 = 0.026$ V) were added at the far end of the waveguide ($z = 4$ mm), respectively. The voltage $V(t)$ across the lumped circuit was simulated by the proposed method and compared with that obtained by SPICE [5], [28], which employed a lossless transmission-line model to represent the waveguide. As shown by Figs. 4 and 5, the proposed method for co-simulation exhibits an excellent agreement with SPICE. The maximum
number of iterations is four in the Newton–Raphson scheme for solving the nonlinear diode during the time marching.

The second example is a bipolar junction transistor (BJT) that terminates a good conducting microstrip line as described in [2]. The strip is discretized into \(6\Delta y\) along the width, \(50\Delta x\) along the length, and \(3\Delta x\) along the height with \(\Delta x = 0.265\) mm, \(\Delta y = 0.39\) mm, and \(\Delta z = 0.4\) mm. The relative permittivity is chosen as 2.2 such that the characteristic impedance is 50 \(\Omega\). The BJT described by the Ebers–Moll model [29] with \(T = 300^\circ K\), \(I_0 = 10^{-16}\) amp, \(\alpha_R = 0.5\), and \(\alpha_F = 0.9001\) is loaded at the far end of the microstrip line. A unit exponential source (\(V(t) = 1 - e^{-t/\tau}\)) with \(\tau = 190\) ps is launched as the incident field at the near end of the microstrip line. The time step used is \(10^{-13}\) s. The maximum number of Newton iterations is eight. The results for both active (\(R_C = 50\) \(\Omega\)) and saturated (\(R_C = 10\) \(\Omega\)) regions of operation are shown in Fig. 6, where an excellent agreement with SPICE can be observed.

The third example is an on-chip problem in which one and multiple CMOS inverters drive an interconnect structure as shown in Fig. 7. The MOS transistor is constructed using SPICE-like level 1 (Shichman–Hodges) model [30] by

\[
I_D = \begin{cases} 
0, & \text{when } V_{GS} - V_{TO} < 0 \text{ (cutoff region)} \\
\frac{W}{L} \frac{K}{2} (1 + \lambda_V) V_{DS} [2(V_{GS} - V_{TO}) - V_{DS}], & \text{when } 0 < V_{DS} < V_{GS} - V_{TO} \text{ (linear region)} \\
\frac{W}{L} \frac{K}{2} (1 + \lambda_V) (V_{GS} - V_{TO})^2, & \text{when } V_{DS} > V_{GS} - V_{TO} \text{ (saturation region)} 
\end{cases}
\]

\( (31) \)

and

\[
\begin{align*}
C_{GS} &= C_{GSO} W \\
C_{GD} &= C_{GDO} W \\
C_{GB} &= C_{GBO} W \\
C_{BD} &= \frac{C_{G} + C_{jw}}{1 - \frac{V_{BD}}{V_{th}}} , & V_{BD} \leq FC \times V_{BO} \\
C_{BS} &= \sqrt{\frac{V_{BD}}{V_{th}}} , & V_{BS} \leq FC \times V_{BO} \\
C_{BD} &= (C_{G0} + C_{jw0}) \frac{1 - 1.5FC \frac{V_{BD}}{V_{th}}}{(1 - FC)^{1.5}} , & V_{BD} > FC \times V_{BO} \\
C_{BS} &= (C_{G0} + C_{jw0}) \frac{1 - 1.5FC \frac{V_{BS}}{V_{th}}}{(1 - FC)^{1.5}} , & V_{BS} > FC \times V_{BO} .
\end{align*}
\]

\( (32) \)

The transistors in 0.25-\(\mu\)m technology are used in this simulation. The physical parameters are not given in detail in the reference. The detailed parameters used for (31) and (32) in our simulation are listed in Table I. The meaning of each parameter can be found from [31].

A falling edge of 0.25 ns is chosen as the input signal of the inverter. The characteristic input–output behaviors of the CMOS inverter simulated by the proposed simulator and those simulated by SPICE are shown in Fig. 8(a). An excellent agreement can be seen. We then co-simulate the combined transistor and interconnect system. The interconnect structure is a three-metal-layer test-chip interconnect structure [17]. The interconnect length (100 \(\mu\)m) along \(y\) is subdivided into five layers. Its two ends are both attached to an air layer, which is truncated by a first-order absorbing boundary condition. The top and bottom
boundaries along the z-direction are perfect electrically conducting (PEC) boundaries. The left and right boundaries along the x-direction are perfect magnetically conducting boundaries. The layer growth direction is chosen as y. Each layer is divided into 676 triangular prism elements with 1053 surface unknowns and 378 volume unknowns. The inverters are connected vertically between the center M2 wire and the ground plane. Each inverter is attached across three edges. For the two-inverter case, two center wires are located in the M2 layer symmetrically with each wire driven by one inverter. The five-inverter case has five M2 wires with each inverter connected to one wire. The meshing size is relatively large for the five-inverter case with 1453 surface unknowns and 518 volume unknowns in each layer. For the five-inverter example, (25) has $3 \times 5 = 15$ unknowns with a $15 \times 15$ dense matrix $M_{eq}$. By the approach described in Section III-D1, (25) is transformed to five nonlinear equations and ten linear equations. As a result, the contribution from the linear system to the Jacobian matrix in the nonlinear solver is a reduced $5 \times 5$ dense matrix.

In total, $5 \times 10^6$, $8 \times 10^6$, $7 \times 10^6$ time steps are run for one, two, and five inverters, respectively. The results are shown in Fig. 8(b)–(d) respectively. The input signals of the inverters and the output signals at the interface between the inverter and the interconnect structure are plotted. Due to the effect of the interconnect structure, the output signals of the inverters are delayed and attenuated.

The performance comparison of the co-simulation performed by the proposed method and that of the traditional time-domain FEM is given in Tables II and III respectively. In these two tables, Case I is a 100-μm-long interconnect structure discretized into five layers driven by one inverter; Case II involves two inverters, Case III involves five inverters, and Case IV is a 2000-μm-long interconnect structure discretized into 100 layers driven by one inverter. As can be seen from Tables II and III, compared with the traditional time-domain FEM-based co-simulation, the proposed methods are much more efficient. In addition, the larger the problem size is, the more significant the speedup is. We also compare the performances of the proposed two methods for the simulation of the nonlinear system. Method II exhibits an obvious advantage in matrix factorization compared with Method I, as can be seen from Table II. Because the nonlinear circuit unknowns are not that large in this example,
the CPU time at each time step cost by the proposed Method I is similar to that of Method II.

V. CONCLUSION

In this paper, an efficient electromagnetics-based co-simulation method is developed for co-simulating linear network and nonlinear circuits present in an integrated circuit. The original 3-D system is first reduced to an orders-of-magnitude smaller system where the nonlinear circuits are located. The reduction is rigorous and performed analytically, and, hence, the computational overhead is minimal. The reduced system is then separated into a linear system and a nonlinear system, both of which are solved efficiently. The remainder of the unknown solutions in the original 3-D system are then recovered by the FE-RR method. The proposed co-simulation algorithm is applicable to arbitrarily shaped multilayer structures embedded in inhomogeneous materials. Furthermore, it bypasses the extraction of the linear network, naturally preserves the passivity and stability of the passive linear network, and captures the coupling between the linear network and nonlinear circuits. In addition, the proposed co-simulation framework can be readily interfaced with the SPICE-based circuit simulator to accentuate the advantage of SPICE in simulating active and nonlinear devices and the advantage of the proposed electromagnetic simulator in simulating large-scale linear networks.

REFERENCES

Qing He received the B.S. degree in electronic and information engineering from Zhejiang University, Hangzhou, China, in 2006. He is currently working toward the Ph.D. degree at the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN.

He was a Research Assistant with the Center for Space Science and Applied Research, Chinese Academy of Sciences from 2006 to 2007. He is currently working with the On-Chip Electromagnetics Group, Purdue University, West Lafayette, IN. His current research interests are computational electromagnetics, high-performance very large-scale integrated computer-aided design, and fast and high-capacity numerical methods.

Dan Jiao (S’00–M’02–SM’06) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, in 2001. She was then with the Technology CAD Division, Intel Corporation, until September 2005 as a Senior CAD Engineer, Staff Engineer, and Senior Staff Engineer. In September 2005, she joined Purdue University, West Lafayette, IN, as an Assistant Professor with the School of Electrical and Computer Engineering. In 2009, she was promoted to Associate Professor with tenure. She has authored or coauthored two book chapters and over 100 papers in refereed journals and international conferences. Her current research interests include computational electromagnetics, high-frequency digital, analog, mixed-signal, and RF IC design and analysis, high-performance very large-scale integrated computer-aided design, modeling of micro- and nano-scale circuits, applied electromagnetics, fast and high-capacity numerical methods, fast time-domain analysis, scattering and antenna analysis, RF, microwave, and millimeter-wave circuits, wireless communication, and bio-electromagnetics.

Dr. Jiao has served as the reviewer for many IEEE journals and conferences. She was the recipient of the Ruth and Joel Spira Outstanding Teaching Award in 2010, the National Science Foundation CAREER Award in 2008, the Jack and Cathie Kozik Faculty Start up Award in 2006, which recognizes an outstanding new faculty member within the School of Electrical and Computer Engineering, Purdue University. She was also the recipient of an Office of Naval Research Award through the Young Investigator Program in 2006 and the Best Paper Award in 2004 from Intel Corporation’s annual corporate-wide technology conference (Design and Test Technology Conference) for her work on generic broadband model of high-speed circuits. In 2003, she was the recipient of the Intel Logic Technology Development (LTD) Divisional Achievement Award in recognition of her work on the industry-leading BroadSpice modeling/simulation capability for designing high-speed microprocessors, packages, and circuit boards and the Intel Technology CAD Divisional Achievement Award for the development of innovative full-wave solvers for high-frequency IC design. In 2002, she was the recipient of the Intel Hero Award by Intel Components Research (Intel-wide she was the tenth recipient) for the timely and accurate two- and three-dimensional full-wave simulations and the Intel LTD Team Quality Award for her outstanding contribution to the development of the measurement capability and simulation tools for high-frequency on-chip cross-talk. She was also the recipient of the 2000 Raj Mittra Outstanding Research Award by the University of Illinois at Urbana-Champaign.