A versatile harmonic balance method in a parallel framework

Andy Huang
Sandia National Laboratories
Albuquerque, NM USA
ahuang@sandia.gov

Xujiao Gao
Sandia National Laboratories
Albuquerque, NM USA
xngao@sandia.gov

Roger Pawlowski
Sandia National Laboratories
Albuquerque, NM USA
rppawlo@sandia.gov

Jason Gates
Sandia National Laboratories
Albuquerque, NM USA
jmgate@sandia.gov

Lawrence Musson
Sandia National Laboratories
Albuquerque, NM USA
lcmusso@sandia.gov

Gary Hennigan
Sandia National Laboratories
Albuquerque, NM USA
glhenni@sandia.gov

Mihai Negoita
Sandia National Laboratories
Albuquerque, NM USA
mnegoit@sandia.gov

***Abstract*—In this paper, we present a parallelized and versatile harmonic balance approach for modeling the small-signal and large-signal frequency-domain response of the coupled semiconductor drift-diffusion equations used in TCAD device simulations. Our approach begins with a time-domain TCAD code, and we describe the process to adapt the system into the frequency domain so that the transformation can be parallelized. Both small-signal and large-signal analyses are easily simultaneously incorporated. Furthermore, we introduce the Isofrequency Remapping Scheme, so that an arbitrary number of high frequencies can be analyzed without introducing a prohibitive expense. Results obtained by our small-signal and large signal harmonic balance methods are shown to capture the same response for a linear device, as expected. Further results use our harmonic balance method to explore a prohibitively expensive time-domain problem: a large-signal, two-tone simulation too costly for a time-domain analysis, for which we are able to produce the expected response with intermodulation.**

***Index Terms*—TCAD, harmonic balance method, frequency-domain analysis, frequency mapping method, parallelization**

I. Introduction

Efficient large-signal multi-tone simulation of radio-frequency (RF) and microwave semicondutor devices requires the use of a frequency-domain method, e.g., a harmonic balance (HB) method, to solve the physics-based semiconductor transport equations. The standard HB method [1][2] for semiconductor devices is often independently implemented from the time-domain (TD) analysis of the transport equations in a TCAD code. Furthermore, small-signal and large-signal frequency-domain (FD) analyses are also typically separated. This leads to significant redundant effort in simultaneously supporting TD and FD simulation capabilities. In particular, any physical model introduced to the TD model needs to be translated and incorporated into the FD code, and vice-versa.

In this work, we present a versatile multi-tone HB method such that the frequency-domain model is directly produced from the TD analysis of the transport equations. We achieve this by abstracting the assembly of the frequency-domain HB residual equations so that any formulation of the TD equations (including spatial discretization methods and material property models) automatically contributes to the corresponding frequency-domain equivalent via the Fourier transform. We also note that our method allows us to easily perform a small-signal linear analysis in addition to the large-signal analysis with no modification to the code; the abstract assembly captures the large-signal fully non-linear HB system by default and be toggled for the small-signal linearization.

The proposed HB method has been implemented in our MPI-parallel TCAD code, Charon [3]. We remark that our HB method is not particular to any spatial discretization method, and can be adapted to other discretization schemes. Charon was used in this work for comparison of TD simulations with our frequency-domain HB simulation results.

II. Method

For brevity, we outline our HB method only for the electron drift-diffusion equation. In its entirety, we can solve this equation coupled to the hole drift-diffusion equation, the Poisson equation, and/or the lattice temperature equation. We express the electron drift-diffusion equation by

$$\frac{\partial n}{\partial t} + \mathcal{F}_n(n, p, \phi) = 0, \tag{1}$$

where $n(t)$, $p(t)$, and $\phi(t)$ is the electron density, hole density, and electric potential, respectively. The quantity $\mathcal{F}_n$ represents the terms that are not explicitly time-dependent. Using a spatial basis function $\Lambda(\vec{x})$ on a spatial discretization element V, we form a residual equation

$$\int_V \frac{\partial n}{\partial t}\Lambda(\vec{x})\, d\vec{x} + \mathcal{R}_n^{\Lambda}(n(t), p(t), \phi(t)) = 0. \tag{2}$$

Note that the $\mathcal{R}_n^{\Lambda}(t)$ term denotes the spatial residual of the steady-state equation. The term $\frac{\partial n}{\partial t}$ is approximated using a time discretization method (e.g., backward-Euler) for a TD transient simulation, or else omitted for a steady-state simulation. We retain this in the HB method, but transform it into the frequency domain.

978-1-5386-6790-3/18/$31.00 ©2018 IEEE

The HB method (in both the small-signal and large-signal modes) has a solution ansatz whose form depends on the applied contact voltage frequencies ω_1 Hz, ω_2 Hz, $\ldots, \omega_\ell$ Hz, called the fundamental frequencies of a simulation. Without loss of generality, we assume that $\omega_1 < \omega_2 < \cdots < \omega_\ell$, and write $\vec{\omega} = (\omega_1, \ldots, \omega_\ell)$ for conciseness. A choice is made for the degree of intermodulation frequencies to be captured by the HB method, and the collection of those intermodulation frequencies is called a *truncation scheme*. Truncation schemes $\mathcal{T}$ are recorded as integer lattice points $\vec{k} = (k_1, k_2, \ldots, k_\ell) \in \mathbb{Z}^\ell$ which correspond to a linear combination $\vec{\omega} \cdot \vec{k}$ of the fundamental frequencies. Thus, the HB solution ansatz takes the form

$$n(\vec{x}, t) = N_0(\vec{x}) + \sum_{\vec{k} \in \mathcal{T}} \Big[N^c_{\vec{k}}(\vec{x}) \cos\left(2\pi \vec{\omega} \cdot \vec{k} t\right) + N^s_{\vec{k}}(\vec{x}) \sin\left(2\pi \vec{\omega} \cdot \vec{k} t\right) \Big] \tag{3}$$

Note that the true solution contains more frequencies than those fundamental frequencies applied, and the ansatz captures this [9]. Common truncation schemes interpolate between the Box and Diamond truncation schemes [1].

To arrive at the HB equations, we first multiply (2) by a Fourier basis function, i.e., $\cos(2\pi\vec{\omega} \cdot \vec{k}t)$ or $\sin(2\pi\vec{\omega} \cdot \vec{k}t)$ for $\vec{k} \in \mathcal{T}$, and integrate over a period. This results in $2|\mathcal{T}| + 1$ residual equations, disregarding $\sin(2\pi 0t)$. In the following, we explicitly describe the procedure involving $\cos(2\pi\omega_i t)$ because all other cases are similar. We next integrate the time-derivative summand analytically, and the time-independent summand numerically using a trapezoidal rule (which converges exponentially for periodic functions [10]). Thus, we obtain the $\cos(2\pi\omega_i t)$ HB residual equation:

$$0 = \pi\omega_i \int_V N^s_{\omega_i}(\vec{x}) \Lambda(\vec{x}) \, d\vec{x} + \sum_{m=0}^{L} w^m_{\omega_i} \mathcal{R}^\Lambda_n(n(t_m), p(t_m), \phi(t_m)) \tag{4}$$

involving time collocation points and quadrature weights

$$t_m = \frac{m}{L} \quad \text{and} \quad w^m_{\omega_i} \equiv \frac{2 - \delta_{0m} - \delta_{mL}}{2} \cos(2\pi\omega_i t_m)$$

where $\delta_{ab} = 1$ when $a = b$ and $\delta_{ab} = 0$ otherwise, and $L = 2\omega_\ell$ by the Nyquist Sampling Theorem.

We note here that the arguments $n(t_m)$, $p(t_m)$, and $\phi(t_m)$ of $\mathcal{R}^\Lambda_n$ are evaluated via the ansatz expression (3). For this evaluation, the small-signal and large-signal formulations differ slightly: for the small-signal analysis, the summation is restricted to only $\vec{k}$ which yields $\vec{\omega} \cdot \vec{k} = \omega_i$; for the large-signal analysis, the summation ranges over all $\vec{k} \in \mathcal{T}$. This has the following analytic interpretation: the small-signal response at the ω_i frequency is only influenced by the contact voltage amplitudes at 0 and ω_i hertz; whereas the large-signal response at any frequency is influenced by the contact voltage amplitudes of all frequencies and their intermodulations.

We see that the HB residual equation (4) is assembled by constructing $\mathcal{R}^\Lambda_n$, the steady-state TD model, $L + 1$ times and substituting the HB solution ansatz equation (3) into the TD model. Clearly, any discretization methods and physical models that are implemented for the TD analysis will be carried over to the HB residual because they are incorporated into $\mathcal{R}^\Lambda_n$. The parallelization we use to speed up the transform comes at two stages: we exploit the summations appearing in the ansatz (3) and the residual (4).

The Nyquist Sampling Theorem recommends a minimum of $2\omega_\ell + 1$ time collocation points, which is too great for large frequencies. This makes the transform (4) prohibitively expensive. To address this, we have developed an efficient *minimal iso-frequency remapping* as an alternative to various frequency mapping methods in the literature [5].

III. Isofrequency remapping scheme

In constructing the HB residual (4), we notice that replacing ω_i with a smaller frequency η_i would result in a similar residual equation. In some cases, this result will be exactly the same as that obtained by using ω_i, with the benefit that only $2\eta_i + 1$ summands are required instead of $2\omega_i + 1$. However, we cannot choose an arbitrary set of frequencies without risk of changing the HB equations because of non-linear interactions. Ideally, all frequencies $\vec{\omega} \cdot \vec{k}$ should be replaced by smaller frequencies while preserving the HB equations.

Frequency remapping schemes have been introduced before. A two-tone remapping scheme was first introduced in [4]. For circuit modelling, but also applicable for device modelling, the two-tone remapping scheme [4] was generalized to a uniformly-spaced spectral remapping scheme [5]. However, these schemes may not be appropriate because we require more than two fundamental frequencies, and the frequencies may not be uniformly spaced.

An efficient frequency remapping scheme should:

1) Minimize the magnitude of the largest frequency in the remapped truncation scheme.
2) Faithfully preserve coincident linear combinations of the fundamental frequencies.

The second property ensures that the remapped system is equivalent to the original, i.e., preserves $\mathbb{Z}$-dependence.

We emphasize the importance of the second point with an example. Suppose that $(\omega_1, \omega_2) = (1 \times 10^6, 1.2 \times 10^6)$. Then the indices $(-5, 5)$ and $(10, 0)$ both correspond to 10 MHz. In this case, if $\mathcal{T}$ includes both of these indices, then we cannot always use the commonly used two-tone remapping scheme

$$\begin{pmatrix} 1 \text{ MHz} \\ 1.2 \text{ MHz} \end{pmatrix} \mapsto \begin{pmatrix} 1 \text{ Hz} \\ (p+1) \text{ Hz} \end{pmatrix}$$

where p is the order of the truncation scheme. This is because the two-tone remapping assumes all linear combinations $\vec{k}$ to correspond to unique harmonics $\vec{k} \cdot \vec{\omega}$; in particular, if $p \neq 2$, then $5p$ Hz $= (-5, 5) \cdot (1 \text{ Hz}, p + 1 \text{ Hz})$ is not equal to 10 Hz $= (10, 0) \cdot (1 \text{ Hz}, p + 1 \text{ Hz})$. Thus, the two-tone remapping scheme [4][5] does not preserve the non-linear interactions of the true HB equations. It is desirable for a remapping scheme to preserve all non-linear interactions; we call such a scheme an *isofrequency* remapping scheme.

The motivating heuristic for our HB scheme is the observation that an isofrequency remapping can be realized as the solution of an integer linear programming problem. Possible remapping candidates $\vec{\eta} = (\eta_1, \eta_2, \ldots, \eta_\ell) \in \mathbb{Z}^\ell$ for the remapped frequencies satisfy $\eta_i < \eta_{i+1}$ for all i. Thus, they form an integer cone C, as depicted in Fig. 1. Furthermore,

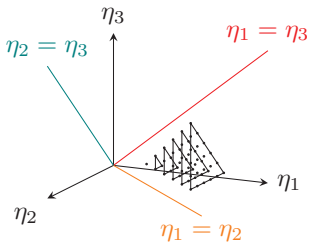


Fig. 1. Cone of remapping candidates

coincident linear combinations of the fundamental frequencies $\vec{\omega}$ determine linear constraints. In particular, if $\vec{\alpha}, \vec{\beta} \in \mathcal{T}$ such that $\vec{\omega}\cdot\vec{\alpha} = \vec{\omega}\cdot\vec{\beta}$, then an isofrequency remapping $\vec{\omega} \mapsto \vec{\eta}$ would yield $\vec{\eta} \cdot \vec{\alpha} = \vec{\beta} \cdot \vec{\eta}$; the constraint is determined by requiring $\vec{\eta} \cdot (\vec{\alpha} - \vec{\beta}) = 0$. Finally, non-incident linear combinations of the fundamental frequencies determine half-plane constraints. In particular, if $\vec{\gamma}, \vec{\delta} \in \mathcal{T}$ such that $\vec{\gamma} \cdot \vec{\omega} > \vec{\delta} \cdot \vec{\omega}$, then we demand that $\vec{\eta} \cdot (\vec{\gamma} - \vec{\delta}) > 0$. Our remapping scheme is thus the mapping $\vec{\omega} \mapsto \vec{\eta}$, where η minimizes $||\eta||_{\ell^\infty}$ among all $\eta \in C$. Our algorithm is summarized in Fig. 2.

An example application of our remapping scheme for an order $p = 10$ box truncation yields the map:

$$\begin{pmatrix} 1 \text{ MHz} + 0 \text{ Hz} \\ 1 \text{ MHz} + 0 \text{ Hz} \\ 1 \text{ MHz} + 0 \text{ Hz} \end{pmatrix} \mapsto \begin{pmatrix} 1 \text{ Hz} \\ 26 \text{ Hz} \\ 51 \text{ Hz} \end{pmatrix}$$

Note that the frequencies of this box truncation scheme are not equally spaced, and the difference in magnitude between the intermodulation frequencies are great.

Algorithm 1

1: Determine the annihilators A.
2: Reduce A to a maximal $\mathbb{Z}$-linearly independent set.
3: Determine the cone C of remapping candidates
4: Define the linear program:
Domain: cone C of remapping candidates
Unknown: seek $\eta \in C$
Constraints: require $\nu \cdot \eta > 0$ for all $\nu \in N$
5: Solve the linear program for η to obtain the map $\omega \mapsto \eta$

Fig. 2. Isofrequency Remapping Algorithm pseudocode

IV. Simulation Results and Discussion

We now show some simulations by our HB method alongside some transient simulations performed by Charon on a model symmetric PN diode. In the following, the cathode is held at 0V while a periodic voltage is applied to the anode.

A. Linear responses

The diode has a quasi-linear response for voltages beyond 1 volt. Its current-voltage characteristic curve is shown in Fig. 3.

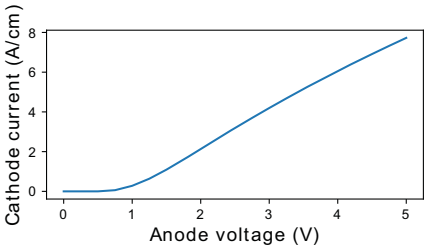


Fig. 3. PN diode current-voltage curve.

To compare our HB methods, we performed a transient simulation with an applied sinusoidal voltage of

$$(4 + 2\sin(2\pi \cdot 2 \cdot t)) \text{ volts}$$

at the anode, for which the voltage varies completely within the diode's linear response regime. We then performed our large-signal and small-signal HB methods. Plots of all results are shown in Fig. 4. We clearly see that the linear response is adequately captured by both of our HB methods.

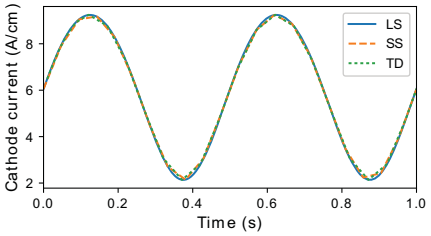


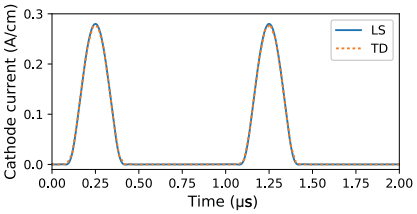
Fig. 4. Responses captured by our large-signal and small-signal harmonic balance methods, transformed into the time domain for comparison with the time-domain simulation.

B. Non-linear responses

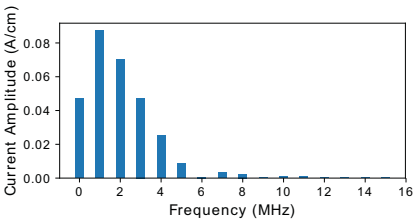
The small-signal analysis assumption is not appropriate for the diode's non-linear regime. For low voltages, we compared our large-signal HB method to the TD method. In Fig. 5, we show the results of a large-scale HB simulation with

$$\sin(2\pi \cdot 10^6 \cdot t) \text{ volts}$$

applied to the anode. The rectifying behavior of the diode is fully captured by our harmonic balance method. The frequency response spectrum includes more frequencies than the stimulus 1 MHz frequency, though it is the dominant response mode.



(a)



(b)

Fig. 5. (a) PN diode cathode harmonic balance and time-domain current responses from 1 MHz applied voltage. (b) The response spectrum, obtained by harmonic balance.

C. Multi-tone responses

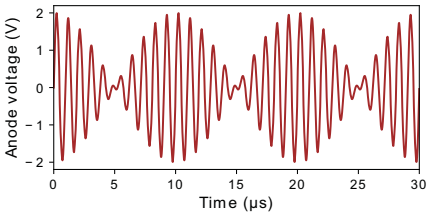
The advantage of a HB method over a TD method is that it can capture multi-tone responses accurately, while using a comparably short simulation time. The TD simulation is extremely computationally expensive for problems involving great input frequencies. Our HB method handles an arbitrary number of tones due to the Isofrequency Remapping Scheme. To demonstrate our HB method's multi-tone capability, we show an example in Fig. 6 of a two-tone simulation with

$$\left(\sin(2\pi \cdot 1.0 \times 10^6 \cdot t) + \sin(2\pi \cdot 1.1 \times 10^6 \cdot t)\right) \text{ volts}$$

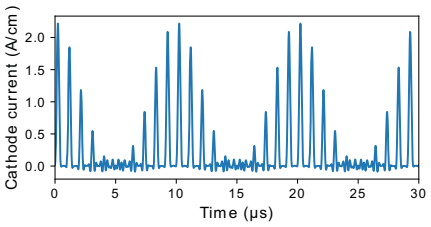
applied to the anode. Note that 0.1 MHz is a non-linear intermodulation frequency, as it is the difference between the two applied anode voltage frequencies. In Fig. 6(a), we clearly see an envelope with this frequency in the time domain applied anode voltage. The same 0.1 MHz intermodulation frequency appears in the current response, Fig. 6(b). We remark that the small oscillations about 0 V are due to a Gibbs phenomenon, an artifact of the order of the harmonic balance truncation scheme. These oscillations are reduced when a higher order truncation scheme is used. The frequency response spectrum, Fig. 6(c), clearly shows many additional intermodulation frequencies due to non-linear interactions. For this problem, our HB simulation took on the order of minutes, while the TD simulation was expensive so we omitted it.

D. High-frequency stimuli

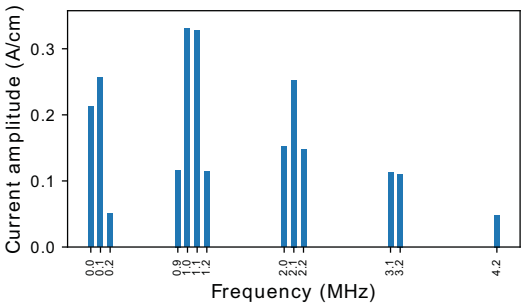
Note that at frequencies up to 100 MHz, the diode rectifies the negative half of the sinusoidal input. From 100 MHz onward, the behavior is more complex - not completely rectifying nor linear. In Fig. 7, we plot responses obtained by HB for frequencies in this range. Our results are consistent with compact models for diodes: a resistor and capacitor in series. At high frequencies, the capacitor's complex impedance drops and the response more resembles that of a linear device.



(a)



(b)



(c)

Fig. 6. (a) A two-tone applied voltage at the anode, supported at 1.0MHz and 1.1MHz. Note the 10μs window coming from the 0.1MHz envelope. (b) Large-signal harmonic balance response transformed into the time domain, depicted over three 10μs periods. (c) Frequency response spectrum.

This contributes to the leading phase shift of the sinusoidal-like profile in the right panel of Fig. 7.

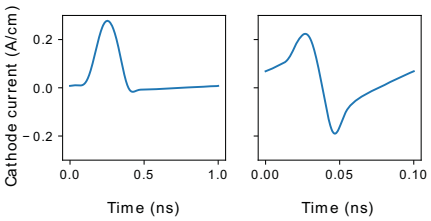


Fig. 7. Cathode current responses at 1 GHz and 10 GHz.

V. Conclusion

In summary, we have shown that our proposed HB method, as implemented in Charon, allows us to efficiently construct the HB residual equations in parallel while utilizing our Isofrequency Remapping Scheme. Furthermore, we have described its versatility and investigated it on a class of linear and non-linear problems. Because of the remapping scheme, we avoid the prohibitive expense required by a TD simulation, and we are able to obtain their frequency response spectra.

ACKNOWLEDGMENTS

This work is funded by the Advanced Scientific Computing (ASC) program at Sandia National Laboratories. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

The views expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

REFERENCES

[1] B. Troyanovsky, "Frequency Domain Algorithms for Simulating Large Signal Distortion in Semiconductor Devices," Ph.D. Thesis, Nov. 1997.

[2] B. Troyanovsky, Z. Yu, and R.W. Dutton, "Large Signal Frequency Domain Device Analysis Via the Harmonic Balance Technique," Simulation of Semiconductor Devices and Processes, Erlangen, Germany, Sep. 4-6, 1995

[3] X. Gao, D. Mamaluy, P. R. Mickel, and M. Marinella, "Three-Dimensional Fully-Coupled Electrical and Thermal Transport Model of Dynamic Switching in Oxide Memristors", ECS Transactions, vol. 69 (5), pp.183-193, Oct. 2015

[4] D. Hente and R.H. Jansen, "Frequency domain continuation method for the analysis and stability investigation of nonlinear microwave circuits," IEEE Proceedings, part H, vol. 133, no.5, pp. 351-362, Oct. 1986

[5] J. C. Pedro and N. B. Carvalho, "Efficient Harmonic Balance Computation of Microwave Circuits' Response to Multi-Tone Spectra", 29th European Microwave Conference Proc., vol. I, pp.103-106, Munich, Oct. 1999

[6] F. M. Rotella, G. Ma, Z. Yu, R. W. Dutton, "Modeling and Simulation of an RF LDMOS Device Using Harmonic Balance PISCES", TECHCON 1998

[7] B. Troyanovsky, Rotella, Yu, Dutton, Sato-Iwanaga, "Large Signal Analysis of RF/Microwave Devices with Parasitics Using Harmonic Balance Device Simulation", SASIMI 1996

[8] Sato-Iwanaga, Fujimoto, Masato, Ota, Inoue, Troyanovsky, Yu, and Dutton, "Distortion Analysis of GaAs MESFETs Based on Physical Model using PISCES-HB", IEDM 1996

[9] Seidman, Thomas I., "Time-dependent Solutions of a Nonlinear System Arising in Semiconductor Theorem, II: Boundedness and Periodicity", IMA Preprint Series 80, July 1984

[10] A. Kurganov and J. Rauch, "The Order of Accuracy of Quadrature Formulae for Periodic Functions", Advances in Phase Space Analysis of Partial Differential Equations: In Honor of Ferruccio Colombini's 60th Birthday, 2009