A Parallel-in-Time Circuit Simulator for Power Delivery Networks with Nonlinear Load Models

Chung-Kuan Cheng[†‡], Chia-Tung Ho[‡], Chao Jiao[+], Xinyuan Wang[‡], Zhiyu Zeng[+], Xin Zhan[+]
[†]CSE and [‡]ECE Departments, UC San Diego, La Jolla, CA, USA
[+] Cadence Design Systems Inc, 12301 Research Blvd Building V Suite 200, Austin, TX 78759
email: {ckcheng,c2ho}@ucsd.edu, cjiao@cadence.com, xiw193@eng.ucsd.edu, {zzeng, xinzhan}@cadence.com

***Abstract*—We apply the parallel-in-time method to the transient simulation of power delivery networks with nonlinear load models. With adaptive Newton-Raphson iterations, the parallel-in-time method achieves considerable speedup compared to the general sequential solver.**

***Index Terms*—PDN transient simulation, nonlinear load model, parallel-in-time**

I. Introduction

In modern VLSI designs, the power integrity becomes a critical issue to ensure the reliability and performance of designs. The challenges of power integrity analysis arise from the tighter noise margin with reducing power supply voltage, higher resistance on metal wires due to scaling, and strong coupling noise between the active devices.

The simulation of power integrity analysis encounters the problems from the increasing size of power delivery networks (PDN) as well as the accuracy of load models. Due to the increasing design complexity, the PDNs could be extremely huge and stiff, which makes the simulation a critical task. To simplify the system-level power integrity analysis, the on-chip macrocells are usually characterized as independent current sources with linear elements. However, the accuracy of power grid analysis is lost and the results could be far from the real cases. An efficient simulation framework is in high demand to handle the issues.

In this paper, we propose a nonlinear macrocell model to capture the dynamic behavior of PDNs and we take advantage of the recent progress in the parallel-in-time approach, such as Parareal (Parallel in Real time) [8] and MGRIT (Multigrid Reduction in Time) [3], and applied the idea to the PDN transient simulations. The main contributions of this paper are listed as follows,

- We adopt a nonlinear voltage-dependent macrocell model in the PDN simulation framework to characterize the dynamic behaviors of whole systems.
- We apply the parallel-in-time method to parallelize the conventional sequential time stepping of the PDN transient simulation.
- We use the adaptive Newton-Raphson (NR) method to solve the nonlinear system efficiently in the iterations of step integrations.

The rest of this paper is organized as follows. In the next section, we introduce the formulation of the PDN transient simulations and nonlinear macromodels. In Sec. IV, we propose the application of the parallel-in-time method to the PDN transient simulations with adaptive NR iterations. Finally, a group of PDNs is used to validate our method. The experimental results are shown in Sec. V.

II. Background

Given the circuit netlist and device models, the general formulation is shown as follows,

$$\frac{dq(x)}{dt} + f(x) = Bu(t), \tag{1}$$

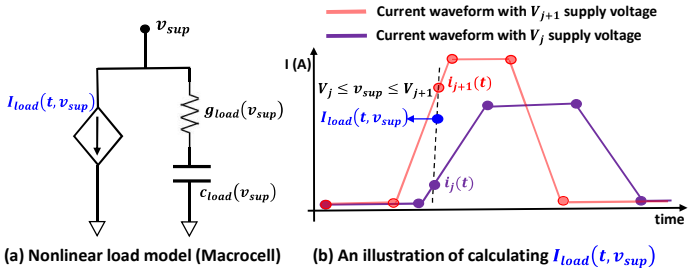


Figure 1: An example of nonlinear load model in PDN.

where $x \in \mathbb{R}^{n\times 1}$ is the vector of nodal voltages and branch currents. The charge/flux is represented by $q \in \mathbb{R}^{n\times 1}$ and $f \in \mathbb{R}^{n\times 1}$ contains the current/voltage terms. Vector $u(t)$ represents all the external excitations at time t and B inserts the signals to the system. If the element constitutive equations are linearized, we can represent Eq. 1 as the matrix form DAEs,

$$\mathcal{C}(x)\dot{x}(t) + \mathcal{G}(x)x(t) = Bu(x,t), \tag{2}$$

where matrices $\mathcal{C}(x) \in \mathbb{R}^{n\times n}$ consists capacitance/inductance and $\mathcal{G}(x) \in \mathbb{R}^{n\times n}$ represents the conductance/resistance, respectively. Vector $u(x,t)$ contains the linear and nonlinear input sources. The elements are functions of x.

With given initial state $x(t)$ and assumption that the system is unchanged in the step from t to $t+h$, the linear multi-step integration methods are widely used approximate the solution in Eq. 2, such as Forward Euler (FE), Backward Euler (BE), Trapezoidal (TR) and explicit Matrix Exponential [10]. To guarantee the stability of algorithms [2], [12], we mainly talk about the BE method in this work.

III. Nonlinear Load Models in PDNs

We propose a nonlinear macrocell load model to include the effects of Dynamic Voltage Drop (DvD) in PDNs [6], [7], [11]. A voltage dependent current source $I_{load}(t, v_{sup})$ with series RC, $R_{load}(v_{sup})$ and $C_{load}(v_{sup})$, are used to model the current fluctuation caused by DvD at the power supply node (i.e., v_{sup}), as shown in Fig. 1. Our nonlinear load model provides the fixed pivot points information, which enable us to determine the simulation time points in advance.

The nonlinear load models are generated at different supply voltages (i.e., V_j). During the transient simulation, the values of elements can be interpreted based on v_{sup} at t as

$$\begin{aligned} I_{load}(t, v_{sup}) &= i_j(t) + \frac{(i_{j+1}(t) - i_j(t))}{(V_{j+1} - V_j)}(v_{sup} - V_j) \\ g_{load}(v_{sup}) &= g_j + \frac{(g_{j+1} - g_j)}{(V_{j+1} - V_j)}(v_{sup} - V_j) \\ c_{load}(v_{sup}) &= c_j + \frac{(c_{j+1} - c_j)}{(V_{j+1} - V_j)}(v_{sup} - V_j) \end{aligned} \tag{3}$$

where v_{sup} lies between two supply voltages $[V_j, V_{j+1}]$ and the coefficients i, g, c represent the element values at each supply voltage in the macrocell model.

978-1-7281-6161-7/20/$31.00

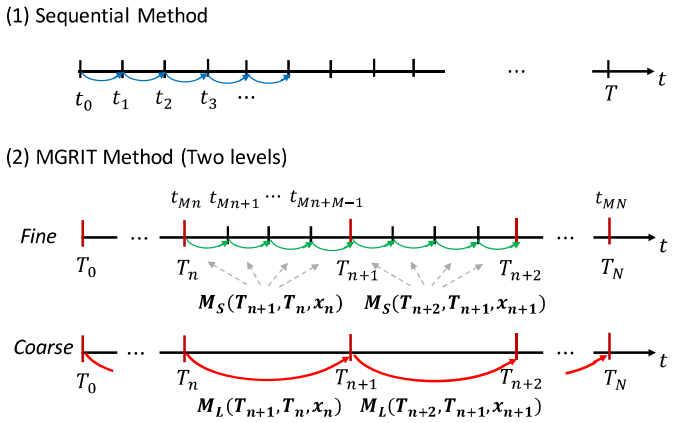


Figure 2: In PDN transient simulation, step integrators are applied to (1) general sequential method and (2) MGRIT method with two levels.

IV. A Parallel-in-Time method for PDN Transient Simulation

We propose a parallel-in-time method for nonlinear PDN transient simulations with the MGRIT method [3] and adaptive Newton-Raphson techniques, named as *MGRIT-AdapNR*. Firstly, we discuss the application of the MGRIT method to circuit simulation in Sec IV-A. Then we introduce the adaptive NR method to solve the nonlinear PDNs in Sec IV-B.

A. MGRIT method with Linear Step integrators

Parareal was first presented as a numerical method to solve evolution problems [8] and extended to PDEs with many follow ups. Consider the DAEs in Eq. 2, BE integration starts from

$$x(t+h) = x(t) + h\dot{x}(t+h), \tag{4}$$

which gives

$$\left(\frac{\mathcal{C}(x)}{h} + \mathcal{G}(x)\right)x(t+h) = \frac{\mathcal{C}(x)}{h}x(t) + Bu(x,t+h), \tag{5}$$

where we can define the operator $M = \left(\frac{\mathcal{C}(x)}{h} + \mathcal{G}(x)\right)^{-1}$ on the rhs of equation at t. The DAEs can be solved with linear step integration method with NR iterations.

Fig. 2 demonstrate (1) the general sequential integraton method and (2) the two level MGRIT method. We assume the fine time grids have uniform step size h. Each time interval in the coarse time grid equals $H = Mh$. Let $M_L(T_{n+1}, T_n, x_n)$ denote the long step integration on the coarse time grid from T_n to T_{n+1}, where $T_n = t_{Mn}$. Let $M_S(T_{n+1}, T_n, x_n)$ denote the short step integration on the fine grid which takes M steps from T_n to T_{n+1}. The MGRIT method performs k iterations and approximates the next approximation with the formulation:

$$x_{n+1}^{(k+1)} = M_L(T_{n+1}, T_n, x_n^{(k+1)}) + M_S(T_{n+1}, T_n, x_n^{(k)}) - M_L(T_{n+1}, T_n, x_n^{(k)}). \tag{6}$$

The first long step integration has to be performed sequentially in order to wait for x_n^{k+1}. The second and the third term only depends on results from the previous iteration, where the integrations between any time interval can be operated in parallel.

The interpretation of Parareal/MGRIT as a time multigrid method is well illustrated in previous work, detailed proof can be found in [4]. The iteration in MGRIT is consistent with the fine grid problem and the algorithm follows the linear convergence of multigrid methods [2], [4], [5].

B. Nonlinear Systems and Adaptive Newton-Raphson Iterations

For a nonlinear system, the implicit formulation Eq. 2 requires NR iterations to achieve a converged solution. We define the residual of the system at t as

$$r(x) \approx Bu(x,t) - \mathcal{C}(x)\dot{x}(t) - \mathcal{G}(x)x(t). \tag{7}$$

Based on the Taylor expansion around the current approximation $x^{(k)}$, the next approximation $x^{(k+1)}$ satisfies

$$0 = r(x^{(k+1)}) \approx r(x^{(k)}) + J(x^{(k)})(x^{(k+1)} - x^{(k)}), \tag{8}$$

where $J(x)$ is the Jacobian matrix with $J_{ij}(x) = \frac{\partial r_i}{\partial x_j}$. In practical circuit simulation, the $J(x)$ is given by the nonlinear elements and choice of multi-step method. The NR iterations follow the relation

$$x^{(k+1)} = x^{(k)} - J(x^{(k)})^{-1} r(x^{(k)}). \tag{9}$$

The corresponding Jacobian is updated at each iteration according to $x^{(k)}$. Either the residual $r(x^{(k+1)})$ is below given tolerance or the change of solution from $x^{(k)}$ to $x^{(k+1)}$ is small enough the iterations are terminated.

Unlike the traditional method where NR iterations are used at each step, adaptive NR (adap. NR) method skips the NR iterations if the change of x at $t+h$ satisfies

$$\|\Delta x^{(0)}\|_\infty \leq \Delta_{th}, \tag{10}$$

where $\Delta x^0 = x^{(0)}(t+h) - x(t)$ and Δ_{th} is the given threshold. Considering the nonlinear macrocell model is less sensitive to its voltage than transistors, we can set larger Δ_{th} to improve the performance.

V. Experimental Results

The *MGRIT-AdapNR* is implemented via the open source software library Xbraid [1] in C++. All experiments are performed on a 1.8GHz Intel Xeon 24-CPUs server.

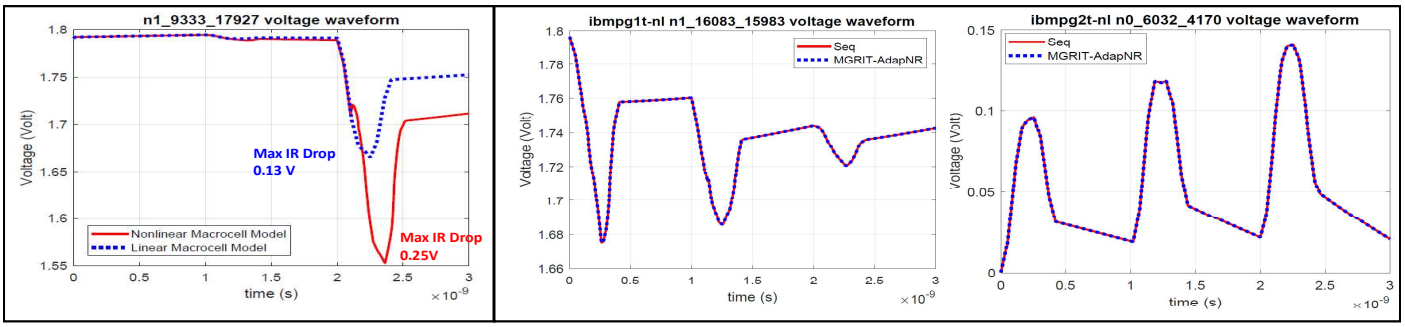
Table I shows the statistics of PDNs with size ranges from thousands to millions, where the design "genckt30" is created based on the specifications in [9] and used for optimum parameter exploration. For ibmpg1t-nl, ibmpg2t-nl, and ibmpg3t-nl, we extend the original power loads to nonlinear load models with the guidance from industry and use the original PDNs of ibmpg1t, ibmpg2t, and ibmpg3t [9]. The nonlinear load models are updated using Eq. 3 in the transient simulations. We compare *MGRIT-AdapNR* with Sequential solver (*Seq*) using NR iterations at each time step. The maximum absolute error e_{max} and average absolute error e_{avg} are calculated from the probing nodes of each design and reported in the following experiments. The runtime represents walltime.

Table I: Design specifications of PDNs

PDN	#R	#C	#L	#Loads	#Size	#Probing Nodes
genckt30	2.6K	1.4K	0	720	1.6K	90
ibmpg1t-nl	54K	11K	277	11K	40K	24
ibmpg2t-nl	245K	37K	330	37K	165K	20
ibmpg3t-nl	1.6M	201K	955	201K	1M	20

A. Study I: Linear vs Nonlinear Load Model

Fig. 3(a) shows the simulation results of a nodal waveform from ibmpg1t with linear load models and ibmpg1t-nl with nonlinear load models. The simulation time is 3ns with 900 time steps. The maximum IR drop with nonlinear load models is 92% larger, which is underestimated by the linear models. The nonlinear load model is essential for power integrity analysis.



(a) Linear vs Nonlinear Macrocell Model with ibmpg1t PDN

(b) Seq vs MGRIT-AdapNR (Table IV)

Figure 3: (a) Linear vs Nonlinear Macrocell Model; (b) Nodal waveforms of *Seq* and *MGRIT-AdapNR* (Table IV).

B. Study II: Optimum Parameter Exploration

We perform multiple experiments on #Cores, Coarsening Factor (CF), and Maximum Level (ML) of *MGRIT-AdapNR* to find the optimum settings in terms of runtime and accuracy.

#Cores: Table. II shows the *MGRIT-AdapNR* runtime with 4, 8, 16, and 24 cores. Compared to the *Seq*, *MGRIT-AdapNR* with 24 cores achieves 2× speedup and the max error is $3mV$. The sublinear trend of Speedup is mainly caused by the overhead of communication for parallel processing. Once the #Cores exceeds a threshold, the speedup is close to linear [3]. Besides, the difference in max and avg errors of different #Cores is less than 1%. *MGRIT-AdapNR* is robust with various #Cores.

Table II: Experimental results of different #Cores using ibmpg1t-nl with 3ns simulation time and 900 time steps.

	#Cores	e_{max} (mV)	e_{avg} (mV)	Runtime (s)	Speedup (X)
Seq	1	-	-	4790.61	1
MGRIT-AdapNR	4	3.00	3.82E-3	6882.63	0.70
	8	3.00	3.83E-3	4250.47	1.13
	16	3.00	3.84E-3	3092.15	1.55
	24	3.00	3.84E-3	2493.53	1.92

Coarsening Factor (CF) and Maximum Level (ML): The CF defines the fine grid and coarse grid ratio at each level and ML defines the maximum level in multigrid. We use "genckt30" with 410K time steps to explore the optimum CF and ML to fully leverage the parallel-in-time advantage. Table III shows the results of *Seq* and *MGRIT-AdapNR* with various combinations of CF and ML. We select CF=2, 6, and 10. Then, we increase the ML from 2 to 10 with increment 2 until the time grids cannot be coarsened any more. From the results, CF=10 and ML=4 achieves the best performance. The max error is less than 1mV.

Table III: Experimental results of *Seq* and *MGRIT-AdapNR* (24 cores), with multiple combinations of CF and ML using genckt30 test case. Simulation time=6ns. #time steps=410K. Time Grid Ratio=(#Finest Time Grids)/(#Coarsest Time Grids).

	CF	ML	Time Grid Ratio	e_{max} (mV)	e_{avg} (mV)	Runtime (s)	Speedup (X)
Seq	-	-	1	-	-	1289.07	1
MGRIT-AdapNR	2	2	2	0.01	4.11E-4	1967.61	0.66
		4	8	0.06	5.86E-3	1011.81	1.27
		6	32	0.16	1.02E-2	793.43	1.62
		8	128	0.22	1.02E-2	730.05	1.77
		10	512	0.22	1.04E-2	710.14	1.82
	6	2	6	0.04	5.75E-3	938.4	1.37
		4	216	0.36	5.84E-3	445.83	2.89
		6	1296	1.10	2.91E-2	426.2	3.02
	10	2	10	0.07	7.70E-3	730.18	1.77
		4	1000	0.14	1.10E-2	390.74	3.30
		6	100000	5.60	2.86E-1	387.27	3.33

C. Main Results

Table IV shows our main results on PDNs in Table I. The simulation time of ibmpg1t-nl, ibmpg2t-nl, and ibmpg3t-nl are 3ns, 3ns, and 2ns with 900, 960, and 630 time steps, respectively. The *MGRIT-AdapNR* multigrid cycles of all three cases are 3. The *MGRIT-AdapNR* multigrid cycles and Adap. NR reduce the #NewtonIters up to 30%. Compared to *Seq*, *MGRIT-AdapNR* achieves more than 2× speedup with less than 5mV max error. The *MGRIT-AdapNR* successfully captures the transient waveform of nonlinear PDNs, as shown in Fig. 3(b).

Table IV: Experimental results of *Seq* and *MGRIT-AdapNR* (#Core=24, CF=10 and ML=4). d_{max} is the absolute error of the max voltage fluctuation of the probing nodes.

	d_{max} (mV)	e_{max} (mV)	e_{avg} (mV)	#NewtonIters		Runtime (s)		Speedup (X)
				Seq	Proposed	Seq	Proposed	
ibmpg1t-nl	5.00E-2	3.00	3.84E-3	1982	1521	4790.61	2493.53	1.92
ibmpg2t-nl	8.00E-2	3.40	8.24E-2	2304	1662	17882.07	7947.37	2.25
ibmpg3t-nl	1.00E-2	2.54	3.12E-2	1824	1256	102683.35	43430.18	2.36

VI. Conclusion

We develop the *MGRIT-AdapNR* for the transient analysis of PDNs with nonlinear load models, where the time integration is parallelized. Compared to the *Seq*, *MGRIT-AdapNR* achieves 3× speedup on long simulation time (410K time steps) and 2× speedup on the PDNs from 40K to 1M size. Without the limitation of maximum #Cores on our server, we expect that *MGRIT-AdapNR* can achieve more speedups. The future research directions include (i) exploring the performance improvement of *MGRIT-AdapNR* with more cores and (ii) improving the convergence rate using advance integrators such as Matrix Exponential [10].

VII. Acknowledgments

We acknowledge the support from NSF CCF-1564302 and the advice from Prof. Albert Chern at UC San Diego and Dr. Robert D. Falgout at Lawrence Livermore National Laboratory.

References

[1] Xbraid: Parallel multigrid in time. http://llnl.gov/casc/xbraid.

[2] V. Dobrev et al. Two-level convergence theory for multigrid reduction in time (mgrit). *SIAM Journal on Scientific Computing*, 39(5):S501–S527, 2017.

[3] R. D. Falgout et al. Parallel time integration with multigrid. *SIAM Journal on Scientific Computing*, 36(6):C635–C661, 2014.

[4] M. J. Gander et al. Analysis of the parareal time-parallel time-integration method. *SIAM Journal on Scientific Computing*, 29(2):556–578, 2007.

[5] M. J. Gander et al. Nonlinear convergence analysis for the parareal algorithm. In *Domain decomposition methods in science and engineering XVII*, pages 45–56. Springer, 2008.

[6] H. Harizi et al. Efficient modeling techniques for dynamic voltage drop analysis. In *Proc. DAC*, pages 706–711. IEEE, 2007.

[7] P.-Y. Hsu et al. Adaptive sensitivity analysis with nonlinear power load modeling. In *SLIP*, pages 1–6. IEEE, 2018.

[8] J.-L. Lions et al. A parareal in time discretization of pde's. In *C.R. Acad. Sci. Paris Ser. I Math*, pages 661–668, 2001.

[9] S. R. Nassif. Power grid analysis benchmarks. In *Proc. DAC, ASPDAC*, pages 376–381. IEEE, 2008.

[10] X. Wang et al. Stability and convergency exploration of matrix exponential integration on power delivery network transient simulation. *TCAD*, 2019.

[11] X. Zhang et al. Power distribution network design optimization with on-die voltage-dependent leakage path. In *EPEPS*, pages 87–90. IEEE, 2013.

[12] H. Zhuang et al. From circuit theory, simulation to $spice^{Diego}$: A matrix exponential approach for time-domain analysis of large-scale circuits. *IEEE Circuits and Systems Magazine*, 16(2):16–34, 2016.