An Efficient and Accurate Time Step Control Method for Power Device Transient Simulation Utilizing Dominant Time Constant Approximation

Shigetaka Kumashiro, *Senior Member, IEEE*, Tatsuya Kamei, Akira Hiroki, *Member, IEEE* and Kazutoshi Kobayashi, *Member, IEEE*

***Abstract*—An accurate metric for the time step control in the power device transient simulation is proposed. This metric contains an exponential term of the dominant time constant of the whole device structure derived from the matrix exponential term of the linearized device state equation. The proposed metric allows larger time step widths than the conventional metric of 2nd order approximation of the local truncation error. It focuses on the dominant part of the transient response and its truncation error approximation is more accurate. In the transient device simulation, box integration method and Backward Euler method are used for spatial and temporal discretization, respectively. The discretized nonlinear device equations are solved by using Newton iteration whose initial guess is given by the approximated solution of the linearized device state equation by using the dominant time constant. Total calculation time of the transient simulation of a silicon power DMOSFET by using the proposed method decreases down to 27% of that by the conventional method with keeping the current accuracy of the dominant transient response.**

***Index Terms*—Arnoldi method, local truncation error, linearized device state equation, power device, time step control, transient device simulation**

I. Introduction

DEVICE simulation is intensively used for power MOSFET design today because the uncertainties in impurity profile, geometry size and carrier transport physics of power MOSFET are all in acceptable level. It can clearly elucidate the mechanism of the phenomenon resulting from the complex coupling of plural physical effects. On the other hand, its calculation time is problematic especially in transient analysis because a large number of meshes must be used to express its large device structure and a large number of time steps must be consumed to trace its relatively slow response. Optimum time step control is important for the calculation time reduction with assuring the accuracy of transient device simulation. In this paper, we improve the speed of the transient simulation of a silicon power DMOSFET by about 4 times by adopting an accurate metric for the time step control and an improved initial value prediction for Newton iteration at each time step by using the dominant time constant information of the whole device structure derived from the matrix exponential term of the linearized device state equation.

This work was supported by the Super Cluster Program from MEXT and JST, and also by JSPS KAKENHI Grant-in-Aid for Scientific Research(C) Number 17K05142.

II. Mathematical Formulation for the Transient Device Simulation

In the transient device simulation, the following devices equations (Poisson equation and electron and hole current continuity equations) are solved.

$$\frac{\partial}{\partial t}\begin{bmatrix}0\\ n(r,t)\\ p(r,t)\end{bmatrix} = -\begin{bmatrix}\frac{1}{q}\nabla_{\mathrm{r}}\cdot\left(\varepsilon\nabla_{\mathrm{r}}\psi(r,t)\right)-n(r,t)+p(r,t)\\ -\frac{1}{q}\nabla_{\mathrm{r}}\cdot\left(-q\mu_{\mathrm{n}}n(r,t)\nabla_{\mathrm{r}}\psi(r,t)+\mu_{\mathrm{n}}k_{\mathrm{B}}T\nabla_{\mathrm{r}}n(r,t)\right)\\ \frac{1}{q}\nabla_{\mathrm{r}}\cdot\left(-q\mu_{\mathrm{p}}p(r,t)\nabla_{\mathrm{r}}\psi(r,t)-\mu_{\mathrm{p}}k_{\mathrm{B}}T\nabla_{\mathrm{r}}p(r,t)\right)\end{bmatrix} + \begin{bmatrix}-N_{\mathrm{D}}(r)+N_{\mathrm{A}}(r)\\ G_{\mathrm{n}}(r,t)-R_{\mathrm{n}}(r,t)\\ G_{\mathrm{p}}(r,t)-R_{\mathrm{p}}(r,t)\end{bmatrix} \quad (1)$$

Here, ψ is electrostatic potential, n is electron density, and p is hole density. r and t stand for spatial coordinates and time, respectively. ε, k_{B}, q, and T are permittivity of the material, Boltzmann constant, unit charge and temperature, respectively. μ_{n} and μ_{p} are electron and hole mobilities, respectively. N_{D} and N_{A} are donor and acceptor densities, respectively. G_{n} and R_{n} are generation and recombination terms for electron, respectively and G_{p} and R_{p} are generation and recombination terms for hole, respectively. Assuming 2-dimensional mesh structure as shown in Fig. 1, spatial discretization of (1) by using box integration method [1][2] and also by using Scharfetter-Gummel discretization scheme [1][3] for electron and hole current continuity equations produces (2). Here, r_i is the spatial coordinates of mesh node i, L_{ij} is the coupling coefficient between mesh nodes i and j, and A_i is the control volume of mesh node i. $B_e(x)$ is Bernoulli function expressed as (3) resulting from Scharfetter-Gummel discretization scheme.

$$
\begin{bmatrix} \vdots \\ 0 \\ \vdots \\ A_i \dfrac{\partial}{\partial t} n(r_i,t) \\ \vdots \\ A_i \dfrac{\partial}{\partial t} p(r_i,t) \\ \vdots \end{bmatrix}
= -\begin{bmatrix} \vdots \\ \dfrac{1}{q}\sum_j \varepsilon_{ij} L_{ij}\left(\psi(r_j,t)-\psi(r_i,t)\right) + A_i\left(-n(r_i,t)+p(r_i,t)\right) \\ \vdots \\ \dfrac{k_B T}{q}\sum_j \mu_{nij} L_{ij}\left(-B_e\left(\dfrac{q}{k_B T}\left(\psi(r_j,t)-\psi(r_i,t)\right)\right)n(r_j,t) + B_e\left(-\dfrac{q}{k_B T}\left(\psi(r_j,t)-\psi(r_i,t)\right)\right)n(r_i,t)\right) \\ \vdots \\ \dfrac{k_B T}{q}\sum_j \mu_{pij} L_{ij}\left(-B_e\left(-\dfrac{q}{k_B T}\left(\psi(r_j,t)-\psi(r_i,t)\right)\right)p(r_j,t) + B_e\left(\dfrac{q}{k_B T}\left(\psi(r_j,t)-\psi(r_i,t)\right)\right)p(r_i,t)\right) \\ \vdots \end{bmatrix}
+ \begin{bmatrix} \vdots \\ -A_i\left(N_D(r_i)-N_A(r_i)\right) \\ \vdots \\ A_i\left(G_n(r_i,t)-R_n(r_i,t)\right) \\ \vdots \\ A_i\left(G_p(r_i,t)-R_p(r_i,t)\right) \\ \vdots \end{bmatrix} \tag{2}
$$

$$
B_e(x) = \frac{x}{exp(x)-1} \tag{3}
$$

$$
\begin{bmatrix} \vdots \\ 0 \\ \vdots \\ A_i \dfrac{n(r_i,t_k)-n(r_i,t_{k-1})}{\Delta t_k} \\ \vdots \\ A_i \dfrac{p(r_i,t_k)-p(r_i,t_{k-1})}{\Delta t_k} \\ \vdots \end{bmatrix}
= -\begin{bmatrix} \vdots \\ \dfrac{1}{q}\sum_j \varepsilon_{ij} L_{ij}\left(\psi(r_j,t_k)-\psi(r_i,t_k)\right) + A_i\left(-n(r_i,t_k)+p(r_i,t_k)\right) \\ \vdots \\ \dfrac{k_B T}{q}\sum_j \mu_{nij} L_{ij}\left(-B_e\left(\dfrac{q}{k_B T}\left(\psi(r_j,t_k)-\psi(r_i,t_k)\right)\right)n(r_j,t_k) + B_e\left(-\dfrac{q}{k_B T}\left(\psi(r_j,t_k)-\psi(r_i,t_k)\right)\right)n(r_i,t_k)\right) \\ \vdots \\ \dfrac{k_B T}{q}\sum_j \mu_{pij} L_{ij}\left(-B_e\left(-\dfrac{q}{k_B T}\left(\psi(r_j,t_k)-\psi(r_i,t_k)\right)\right)p(r_j,t_k) + B_e\left(\dfrac{q}{k_B T}\left(\psi(r_j,t_k)-\psi(r_i,t_k)\right)\right)p(r_i,t_k)\right) \\ \vdots \end{bmatrix}
+ \begin{bmatrix} \vdots \\ -A_i\left(N_D(r_i)-N_A(r_i)\right) \\ \vdots \\ A_i\left(G_n(r_i,t_k)-R_n(r_i,t_k)\right) \\ \vdots \\ A_i\left(G_p(r_i,t_k)-R_p(r_i,t_k)\right) \\ \vdots \end{bmatrix} \tag{4}
$$

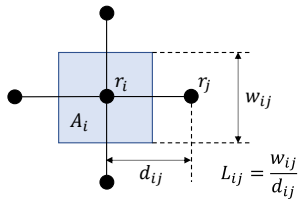

Fig. 1. Spatial discretization with 2-dimensioal mesh structure. $L_{ij} = w_{ij}/d_{ij}$ is the coupling coefficient between mesh nodes i and j, where w_{ij} and d_{ij} are the cross-section for the flux and the mesh node distance, respectively. A_i is the control volume of mesh node i.

Index j is summed up for the mesh nodes connected to the mesh node i. Temporal discretization of (2) by using Backward Euler method [4][5] produces (4). Here, t_k is the k-th analysis time in the transient simulation and $\Delta t_k \stackrel{\text{def}}{=} t_k - t_{k-1}$ is the k-th time step width. By setting proper boundary conditions and also by using matrix and vector symbols, (4) can be expressed as (5).

$$\frac{1}{\Delta t_k}\boldsymbol{C}\cdot(\boldsymbol{x}_k - \boldsymbol{x}_{k-1}) = -\boldsymbol{F}(\boldsymbol{x}_k)\cdot\boldsymbol{x}_k + \boldsymbol{F}_{\boldsymbol{C}}(\boldsymbol{x}_k) + \boldsymbol{B}\cdot\boldsymbol{u}(t_k) \quad (5)$$

Here, $\boldsymbol{x}_k$ is the variable vector at analysis time t_k and equals to ${}^t(\cdots\psi(r_i,t_k)\cdots,\cdots n(r_i,t_k)\cdots,\cdots p(r_i,t_k)\cdots)$. $\boldsymbol{C}$ is effective capacitance diagonal matrix whose elements are the control volumes of each mesh node. $\boldsymbol{F}(\boldsymbol{x}_k)$ and $\boldsymbol{F}_{\boldsymbol{C}}(\boldsymbol{x}_k)$ are nonlinear coefficient matrix and nonlinear vector term at t_k, respectively. $\boldsymbol{B}$ is input connection matrix, $\boldsymbol{u}(t_k)$ is input vector at t_k. In the Backward Euler method, all the input signals are approximated with piecewise linear waveforms. In this paper, only single device and voltage sources connected to the device terminals are considered and no other external circuit element is considered for simplicity. Since (5) is nonlinear with respect to $\boldsymbol{x}_k$, Newton method is used to solve (5), i.e. the following (6)-(9) are iteratively calculated with the initial value $\boldsymbol{x}_{k,1}$ until the norm of $\boldsymbol{\delta x}_{k,n+1}$ becomes sufficiently small.

$$\boldsymbol{G}(\boldsymbol{x}_{k,n}) = \left.\frac{\partial(\boldsymbol{F}(\boldsymbol{x})\cdot\boldsymbol{x})}{\partial\boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}_{k,n}} - \left.\frac{\partial\boldsymbol{F}_{\boldsymbol{C}}(\boldsymbol{x})}{\partial\boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}_{k,n}} \quad (6)$$

$$\left(\frac{1}{\Delta t_k}\boldsymbol{C} + \boldsymbol{G}(\boldsymbol{x}_{k,n})\right)\cdot\boldsymbol{\delta x}_{k,n+1} = -\left(\frac{1}{\Delta t_k}\boldsymbol{C}\cdot(\boldsymbol{x}_{k,n} - \boldsymbol{x}_{k-1}) + \boldsymbol{F}(\boldsymbol{x}_{k,n})\cdot\boldsymbol{x}_{k,n} - \boldsymbol{F}_{\boldsymbol{C}}(\boldsymbol{x}_{k,n}) - \boldsymbol{B}\cdot\boldsymbol{u}(t_k)\right) \quad (7)$$

$$\boldsymbol{x}_{k,n+1} = \boldsymbol{x}_{k,n} + \alpha\boldsymbol{\delta x}_{k,n+1} \quad (8)$$

$$n = n + 1 \quad (9)$$

Here, n is the number of Newton iteration. $\boldsymbol{x}_{k,n+1}$ is the approximate solution of $\boldsymbol{x}_k$ after n-th Newton iteration, and $\boldsymbol{\delta x}_{k,n+1}$ is the correction vector at n-th Newton iteration. α is damping factor to prevent divergence when electrostatic potential variation is too large. $\boldsymbol{G}(\boldsymbol{x}_{k,n})$ is also equivalent to the Jacobian used for obtaining DC solution by Newton method with initial value $\boldsymbol{x}_{k,n}$ after the transient has been decayed with very large time step width i.e. $\Delta t_k \to \infty$. To solve (7), LU-factorization of $\boldsymbol{C}/\Delta t_k + \boldsymbol{G}(\boldsymbol{x}_{k,n})$ is used here. For LU-factorization, Crout method without pivoting is used here, and the variables are reordered as ${}^t(\psi(r_1,t_k), n(r_1,t_k), p(r_1,t_k), \psi(r_2,t_k), n(r_2,t_k), \cdots)$ so that a narrow width band matrix is formed to reduce the calculation time. After the Newton iteration converges, the analysis time is advanced by setting:

$$\boldsymbol{x}_k = \boldsymbol{x}_{k,n+1} \quad (10)$$

$$t_{k+1} = t_k + \Delta t_{k+1}. \quad (11)$$

If the Newton iteration does not converge or the Local Truncation Error (LTE) is greater than the specified criteria, the time step is cancelled and re-calculated with smaller time step width. The overall algorithm is shown in Fig. 2. As for Step 6-8 and Step 25-27 of the algorithm, detailed explanation is given in later sections. The mathematical method explained in this section is used for the transient device simulation throughout this paper.

Transient device simulation algorithm: Solve the discretized non-linear equations by Newton method.

1. $k = 0, t_0 = 0$.
2. Set initial time step width Δt_1, $t_1 = t_0 + \Delta t_1$.
3. Do {
4. $k = k + 1, n = 0$.
5. Set boundary condition $\boldsymbol{B}\cdot\boldsymbol{u}(t_k)$.
6. If 'Exp_LTE_metric' is used then
7. Calculate dominant time constant τ_c by Arnoldi algorithm (Fig. 8).
8. Set initial guess for $\boldsymbol{x}_{k,1}$ by (32).
9. else
10. $\boldsymbol{x}_{k,1} = \boldsymbol{x}_{k-1}$.
11. endif
12. Do {
13. $n = n + 1$.
14. Calculate $\boldsymbol{F}(\boldsymbol{x}_{k,n})$, $\boldsymbol{F}_{\boldsymbol{C}}(\boldsymbol{x}_{k,n})$, $\boldsymbol{G}(\boldsymbol{x}_{k,n}) = \left.\frac{\partial(\boldsymbol{F}(\boldsymbol{x})\cdot\boldsymbol{x})}{\partial\boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}_{k,n}} - \left.\frac{\partial\boldsymbol{F}_{\boldsymbol{C}}(\boldsymbol{x})}{\partial\boldsymbol{x}}\right|_{\boldsymbol{x}=\boldsymbol{x}_{k,n}}$.
15. LU-factorization: $\boldsymbol{L}\cdot\boldsymbol{U} = \frac{1}{\Delta t_k}\boldsymbol{C} + \boldsymbol{G}(\boldsymbol{x}_{k,n})$.
16. $\boldsymbol{\delta x}_{k,n+1} = -\boldsymbol{U}^{-1}\left(\boldsymbol{L}^{-1}\left(\frac{1}{\Delta t_k}\boldsymbol{C}\cdot(\boldsymbol{x}_{k,n} - \boldsymbol{x}_{k-1}) + \boldsymbol{F}(\boldsymbol{x}_{k,n})\cdot\boldsymbol{x}_{k,n} - \boldsymbol{F}_{\boldsymbol{C}}(\boldsymbol{x}_{k,n}) - \boldsymbol{B}\cdot\boldsymbol{u}(t_k)\right)\right)$.
17. If $|\boldsymbol{\delta\psi}_{k,n}| > b$ then
18. $\alpha = \frac{a\cdot log(|\boldsymbol{\delta\psi}_{k,n}| - b + a) + b - a\cdot log(a)}{|\boldsymbol{\delta\psi}_{k,n}|}$
19. else
20. $\alpha = 1$
21. endif
22. $\boldsymbol{x}_{k,n+1} = \boldsymbol{x}_{k,n} + \alpha\boldsymbol{\delta x}_{k,n+1}$.
23. } while $|\boldsymbol{\delta x}_{k,n+1}| > E_{rn}\cdot|\boldsymbol{x}_{k,n}| + E_{an}$ and $n \le n_{max}$.
24. $\boldsymbol{x}_k = \boldsymbol{x}_{k,n+1}$.
25. Evaluate Local Truncation Error (LTE).
26. if $LTE > E_r\cdot|\boldsymbol{x}_k| + E_a$ or $n > n_{max}$ then $k = k - 1$.
27. Calculate Δt_{k+1} based on LTE and n.
28. $t_{k+1} = t_k + \Delta t_{k+1}$
29. } while $t_{k+1} < t_{end}$ and $k < k_{max}$.

Fig. 2. Flow chart of transient device simulation algorithm used in this paper.

III. An Exponential Based Local Truncation Error Metric

In the conventional transient device simulation by using Backward Euler method, 2nd_order_LTE (Local Truncation Error) which is a product of second time derivative of carrier density and squared time step width is usually used for the time step width control metric [5] [6]. 2nd_order_LTE is derived as follows. Consider that the following ordinary differential equation

$$\frac{dx}{dt} = f(x). \tag{12}$$

Time discretization of (12) by using Backward Euler method yields

$$\frac{x_{\mathrm{BE}}(t_0 + \Delta t) - x_{\mathrm{BE}}(t_0)}{\Delta t} = f\big(x_{\mathrm{BE}}(t_0 + \Delta t)\big), \tag{13}$$

where x_{BE} is an approximated solution by Backward Euler method, t_0 is the previous time, and Δt is a time step width. Taylor expansion of the exact solution $x(t)$ at $t = t_0 + \Delta t$ becomes

$$x(t_0) = x(t_0 + \Delta t) - \dot{x}(t_0 + \Delta t) \cdot \Delta t + \frac{1}{2!}\ddot{x}(t_0 + \Delta t) \cdot (\Delta t)^2 + O((\Delta t)^3). \tag{14}$$

Subtracting (13) from (14) assuming $x_{BE}(t_0 + \Delta t) = x(t_0 + \Delta t)$ produces LTE definition:

$$LTE \stackrel{\text{def}}{=} |x_{\mathrm{BE}}(t_0) - x(t_0)| = \left|\frac{1}{2}\ddot{x}(t_0) \cdot (\Delta t)^2 + O((\Delta t)^3)\right|. \tag{15}$$

Finally, 2nd_order_LTE-metric is obtained as follows by omitting unknown higher order terms.

$$2^{nd}\_order\_LTE\_metric \stackrel{\text{def}}{=} \left|\frac{1}{2}\ddot{x}(t_0) \cdot (\Delta t)^2\right|. \tag{16}$$

In the transient analysis, Δt is controlled so that (16) should be less than a predetermined criterion such as:

$$\Delta t \le \sqrt{\frac{2 \cdot (E_{\mathrm{r}} \cdot |x(t_0)| + E_{\mathrm{a}})}{|\ddot{x}(t_0)| + \delta}}, \tag{17}$$

where E_{r} and E_{a} are relative and absolute LTE tolerances, respectively and δ is a ceiling value to protect the denominator from becoming zero. However, the accuracy of this metric is dubious because it only adopts the most dominant term in the Taylor expansion of the LTE and the contributions from the other higher order terms are ignored. In the case of exponentially decaying transient response, cancellation effect between the Taylor expansion terms is expected since all the odd-order derivatives become negative and all the even-order derivatives become positive in the Taylor expansion of $\exp(-x)$. Therefore, (16) may overestimate the actual LTE and may raise false alarm to decrease Δt.

On the other hand, suppose that the response of the whole device structure can be well approximated by a dominant time constant τ_c. This assumption is not valid for general circuits, especially for complex RLC circuits which have a lot of time constants uniformly scattered in a wide range. However, as will be shown in the later sections, it holds for power devices which have a large and well-separated time constant for the main response. If the dominant time constant τ_c is known a priori, the LTE can be estimated as Exp_LTE_metric as follows. As defined in (15), the LTE is the difference between an approximated solution by Backward Euler method and the exact solution as

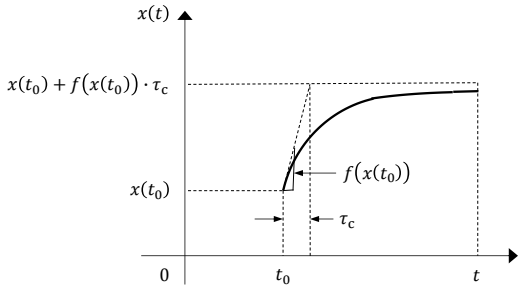

Fig. 3. Approximation of the exact solution by using a dominant time constant τ_c.

$$LTE \stackrel{\text{def}}{=} |x_{\mathrm{BE}}(t_0) - x(t_0)| = \left|-\frac{1}{2!}\ddot{x}(t_0 + \Delta t) \cdot (\Delta t)^2 + \frac{1}{3!}\dddot{x}(t_0 + \Delta t)(\Delta t)^3 - \cdots\right|. \tag{18}$$

Assume the exact solution $x(t)$ can be well-approximated by using a dominant time constant τ_c of the whole device structure as:

$$x(t) = x(t_0) + f\big(x(t_0)\big) \cdot \tau_{\mathrm{c}} \cdot \left(1 - \exp\left(-\frac{t - t_0}{\tau_{\mathrm{c}}}\right)\right). \tag{19}$$

The shape of (19) is shown in Fig. 3. By substituting (19) into (18) and putting the Taylor series into an exponential term, the Exp_LTE_metric can be derived as:

$$Exp\_LTE\_metric \stackrel{\text{def}}{=} \left|\dot{x}(t_0) \cdot \tau_{\mathrm{c}} \cdot \left[1 - \left(1 + \frac{\Delta t}{\tau_{\mathrm{c}}}\right)\exp\left(-\frac{\Delta t}{\tau_{\mathrm{c}}}\right)\right]\right|. \tag{20}$$

Equation (20) is expected to be more accurate than (16) since all the Taylor expansion terms are included. For Exp_LTE_metric, Δt is determined so that (20) should be less than a predetermined criterion such as:

$$1 - \left(1 + \frac{\Delta t}{\tau_{\mathrm{c}}}\right)\exp\left(-\frac{\Delta t}{\tau_{\mathrm{c}}}\right) \le \frac{E_{\mathrm{r}} \cdot |x(t_0)| + E_{\mathrm{a}}}{|\dot{x}(t_0) \cdot \tau_{\mathrm{c}}| + \delta}. \tag{21}$$

Equation (21) can be solved by using such as bi-sectional method [7]. Fig. 4 shows the time step width (Δt) dependence of the main Exp_LTE_metric term in (20) (i.e. LHS of (21)) with respect to $|\tau_{\mathrm{c}}|$. This figure tells us two important features. The first one is that the main Exp_LTE_metric term exponentially decreases as Δt becomes smaller than $|\tau_{\mathrm{c}}|$. This feature is advantageous for adopting large time step width by filtering out the unimportant responses whose time constants are much smaller than the dominant one. On the contrary, as the conventional 2nd_order_LTE_metric watches local variations only, it cannot overlook any quick response even if it is unimportant. The second feature is that although the τ_{c} of the active device can be either positive or negative, the main

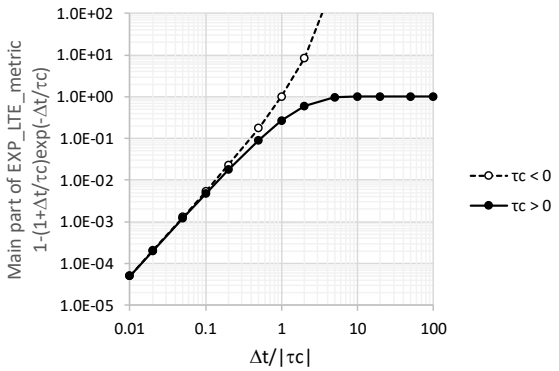

Fig. 4. Time step width dependence of the main part of Exp_LTE_metric with respect to the dominant time constant τ_c. The filled circle with solid line is for positive τ_c and the open circle with broken line is for negative τ_c. As long as $\tau_c > 0$, the main part of Exp_LTE_metric never exceeds 1.

Exp_LTE_metric term never exceeds 1 as long as the τ_c is positive. This means that if $|\dot{x}(t_0) \cdot \tau_c|$ in (20) is smaller than the required LTE criterion, it is possible to take Δt as infinite. On the other hand, Δt in the 2nd_order_LTE_metric of (16) never fails to be restricted to some limited value unless $|\ddot{x}(t_0)|$ is 0. If the false alarm is suppressed by replacing 2nd_order_LTE_metric with Exp_LTE_metric, it is possible to enjoy the efficiency of large Δt.

IV. Dominant Time Constant of a Device

Spatially discretized device equations by using mesh can be expressed as an equivalent circuit shown in Fig. 5 (a) [8]. Here, electrostatic potential (ψ) and carrier densities (n, p) are taken as state variables and nonlinear resistance and linear capacitance resulting from the shape of the control volume [2] and physical nature of each mesh point are used as circuit elements. The capacitances between carrier density nodes and potential nodes in Fig. 5 (a) correspond to A_i in LHS of (2) and are also the elements of the effective capacitance diagonal matrix $\boldsymbol{C}$ in (5). The capacitances between potential nodes correspond to the $\varepsilon_{ij} L_{ij}/q$ in RHS of (2). The 4 capacitances around each potential node must satisfy charge conservation law required by Poisson equation.

In the case of 1-dimensional PN-diode at thermal equilibrium, the resistance value becomes high where the relevant carrier density is low. If such nonconductive high resistances are removed, the equivalent circuit becomes the one shown in Fig. 5 (b). The shortest response time constants of this equivalent circuit result from the resistances and capacitances in the same control volume expressed by the thick arrows in Fig. 5 (b). These time constants are equal to the dielectric relaxation time of doped silicon (τ_d) expressed as:

$$\tau_d = \frac{\varepsilon_{si}}{q(\mu_n n + \mu_p p)} \qquad (22)$$

Here, ε_{Si} is dielectric constant of silicon. Dielectric relaxation time easily becomes less than 1E-12 s where the doping density is high. Using an explicit time discretization scheme such as Forward Euler method is not practical in transient device simulation because the maximum time step width is restricted

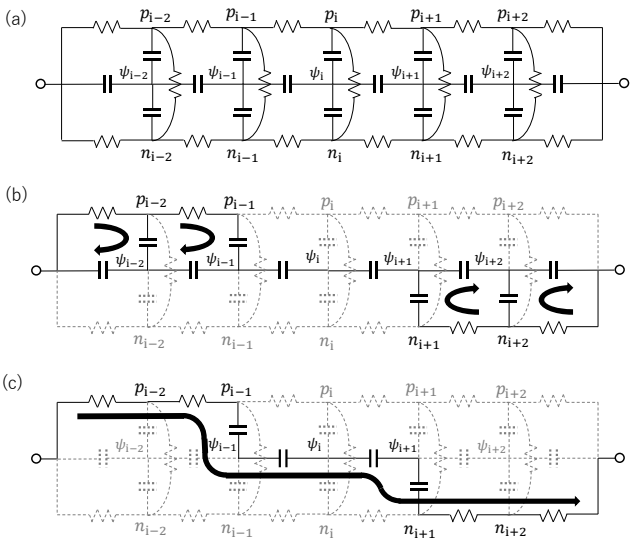

Fig. 5. (a) Equivalent circuit for device state variables (electrostatic potential ψ and carrier densities n, p) in terms of non-linear resistances and linear capacitances derived by spatial mesh discretization. (b) Equivalent circuit for PN-diode where the low conductive resistances corresponding to the low carrier densities are ignored. The thick arrows express the shortest time constants in the circuit which are equal to the dielectric relaxation time of the highly doped control volumes. (c) Equivalent circuit where the capacitances in the control volumes whose dielectric relaxation time is less than 1E-12 s are ignored. The thick arrow expresses the longest time constant which restrict the response of the whole device structure.

by the dielectric relaxation time.

Generally, in the time scale longer than the dielectric relaxation time, it is possible to treat semiconductor as simple resistance with ignoring the relevant capacitance. Since the cutoff frequency of the PN-diode at thermal equilibrium is a few hundreds of GHz, the capacitances in the control volumes whose dielectric relaxation times are less than 1E-12 s can be ignored. In this case, the equivalent circuit becomes the one shown in Fig. 5 (c). The thick arrow in Fig. 5 (c) shows the series resistances and capacitances corresponding to the longest response time constant which restricts the response of the whole device structure. Compact transistor model further simplifies this equivalent circuit with small number of lumped elements. The longest response time constant expresses the most principal and important transient response from a view point of practical device operation. It is called dominant time constant in this paper.

V. Dominant Time Constant Extraction

A. Mathematical Background

Consider device equations (1). As explained in Section II, in terms of spatial discretization by box integration method and Scharfetter-Gummel discretization scheme, with proper boundary conditions, the following symbolic device state equation is derived [9]-[11]:

$$\boldsymbol{C} \cdot \frac{\partial \boldsymbol{x}}{\partial t} = -\boldsymbol{F}(\boldsymbol{x}) \cdot \boldsymbol{x} + \boldsymbol{F_C}(\boldsymbol{x}) + \boldsymbol{B} \cdot (\boldsymbol{u_0} + \boldsymbol{u_1} \cdot t). \qquad (23)$$

Here, $\boldsymbol{x}$ is state variable vector and equals to

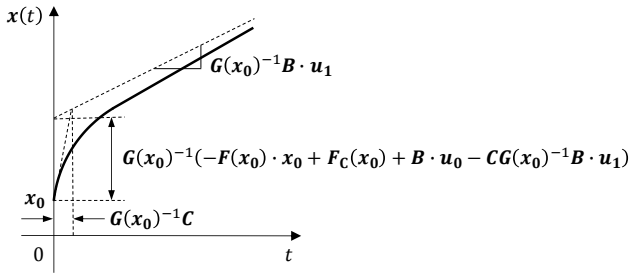

Fig. 6. Time dependent behavior of the formal solution of the linearized device state equation.

$^t(\cdots \psi(r_i,t)\cdots, \cdots n(r_i,t)\cdots, \cdots p(r_i,t)\cdots)$. Although (23) is similar to (5), temporal discretization is not applied here. Moreover, only step input vector $\boldsymbol{u_0}$ and ramp input vector $\boldsymbol{u_1}$ are considered here because all the input signals are approximated with piecewise linear waveforms as explained in Section II. Linearization of (23) by using Taylor expansion with respect to $\boldsymbol{x}$ at $\boldsymbol{x} = \boldsymbol{x_0}$ results in:

$$\boldsymbol{C}\cdot\frac{\partial \boldsymbol{x}}{\partial t} = -\boldsymbol{F}(\boldsymbol{x_0})\cdot\boldsymbol{x_0} + \boldsymbol{F_C}(\boldsymbol{x_0}) + \boldsymbol{B}\cdot(\boldsymbol{u_0}+\boldsymbol{u_1}\cdot t) + \left(-\frac{\partial(\boldsymbol{F}(\boldsymbol{x})\cdot\boldsymbol{x})}{\partial \boldsymbol{x}}\bigg|_{x=x_0} + \frac{\partial \boldsymbol{F_C}(\boldsymbol{x})}{\partial \boldsymbol{x}}\bigg|_{x=x_0}\right)\cdot(\boldsymbol{x}-\boldsymbol{x_0}) \quad (24)$$

or by using (6) [9]-[11],

$$\boldsymbol{C}\cdot\frac{\partial \boldsymbol{x}}{\partial t} = -\boldsymbol{G}(\boldsymbol{x_0})\cdot\boldsymbol{x} + \boldsymbol{G}(\boldsymbol{x_0})\cdot\boldsymbol{x_0} - \boldsymbol{F}(\boldsymbol{x_0})\cdot\boldsymbol{x_0} + \boldsymbol{F_C}(\boldsymbol{x_0}) + \boldsymbol{B}\cdot(\boldsymbol{u_0}+\boldsymbol{u_1}\cdot t). \quad (25)$$

Formal solution of (25) can be obtained as:

$$\boldsymbol{x}(t) = \boldsymbol{x_0} + \left(\boldsymbol{I} - \exp(-\boldsymbol{C}^{-1}\boldsymbol{G}(\boldsymbol{x_0})\cdot t)\right)\cdot\boldsymbol{G}(\boldsymbol{x_0})^{-1}(-\boldsymbol{F}(\boldsymbol{x_0})\cdot\boldsymbol{x_0} + \boldsymbol{F_C}(\boldsymbol{x_0}) + \boldsymbol{B}\cdot\boldsymbol{u_0} - \boldsymbol{C}\boldsymbol{G}(\boldsymbol{x_0})^{-1}\boldsymbol{B}\cdot\boldsymbol{u_1}) + \boldsymbol{G}(\boldsymbol{x_0})^{-1}\boldsymbol{B}\cdot\boldsymbol{u_1}\cdot t, \quad (26)$$

where $\boldsymbol{I}$ is identity matrix. One important thing to be noted is that diagonal matrix $\boldsymbol{C}$ is singular since Poisson equation, the first row in (1), does not have time derivative term. Therefore, in reality, (26) does not hold. However, as will be explained in the latter part of this section, there is no need to calculate $\boldsymbol{C}^{-1}$ in the actual approximated calculation of the matrix exponential term $\exp(-\boldsymbol{C}^{-1}\boldsymbol{G}(\boldsymbol{x_0})\cdot t)$. Here, (26) is temporary used for explaining the structure of the formal solution. Time development behavior of (26) is shown in Fig. 6. The term $\boldsymbol{x_0}$ is the initial value of the device state equation. The term:

$$\boldsymbol{I} - \exp(-\boldsymbol{C}^{-1}\boldsymbol{G}(\boldsymbol{x_0})\cdot t) \quad (27)$$

in (26) is the transient decay factor. The term:

$$\boldsymbol{G}(\boldsymbol{x_0})^{-1}(-\boldsymbol{F}(\boldsymbol{x_0})\cdot\boldsymbol{x_0} + \boldsymbol{F_C}(\boldsymbol{x_0}) + \boldsymbol{B}\cdot\boldsymbol{u_0} - \boldsymbol{C}\boldsymbol{G}(\boldsymbol{x_0})^{-1}\boldsymbol{B}\cdot\boldsymbol{u_1}) \quad (28)$$

in (26) expresses the transient response amplitude. The term:

$$\boldsymbol{G}(\boldsymbol{x_0})^{-1}\boldsymbol{B}\cdot\boldsymbol{u_1}\cdot t \quad (29)$$

in (26) corresponds to the time development by ramp input. The term:

$$-\boldsymbol{F}(\boldsymbol{x_0})\cdot\boldsymbol{x_0} + \boldsymbol{F_C}(\boldsymbol{x_0}) \quad (30)$$

in (28) stands for the continuity unbalance at initial state. The term:

$$\boldsymbol{B}\cdot\boldsymbol{u_0} - \boldsymbol{C}\boldsymbol{G}(\boldsymbol{x_0})^{-1}\boldsymbol{B}\cdot\boldsymbol{u_1} \quad (31)$$

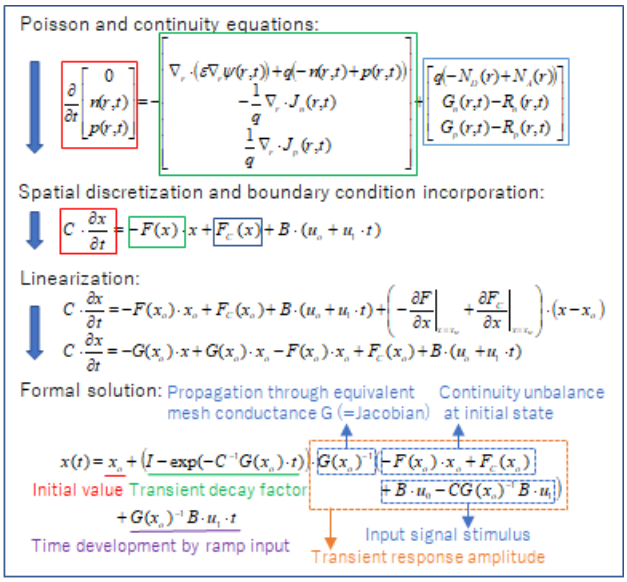

Fig. 7. Derivation procedure of the formal solution of the linearized device state.

in (28) is the input signal stimulus. Equation (28) tells us that the transient response amplitude is the result of the propagation of the source signals (30)-(31) through the equivalent mesh conductance $\boldsymbol{G}(\boldsymbol{x_0})$. The derivation procedure of the formal solution of the linearized device state equation explained above is summarized in Fig. 7. Since the original device state equation is highly nonlinear system, especially due to the Bernoulli function (3), the formal solution of the linearized system (26) is valid only within limited time duration. Therefore, transient analysis with temporal discretization such as Backward Euler method explained in Section II is necessary to get long range solution [10]-[11].

Since the matrix exponential term in (27) is infinite series of matrix product which is derived by applying Taylor expansion of exponential function to matrix, it is difficult to carry out actual calculation. To overcome this difficulty, various approximation methods have been proposed [12]-[23]. In this paper, matrix exponential is replaced by scalar exponential function in terms of the largest eigenvalue of $\boldsymbol{G}^{-1}\boldsymbol{C}$ [9]. The largest eigenvalue of $\boldsymbol{G}^{-1}\boldsymbol{C}$ corresponds to the smallest eigenvalue of $\boldsymbol{C}^{-1}\boldsymbol{G}$ in (27). It also corresponds to the dominant time constant of the whole device structure τ_c (= the longest time constant). Then, the approximation of (26) becomes as follows:

$$\boldsymbol{x}(t) = \boldsymbol{x_0} + \left(1 - \exp\left(-\frac{t}{\tau_c}\right)\right)\cdot\boldsymbol{G}(\boldsymbol{x_0})^{-1}(-\boldsymbol{F}(\boldsymbol{x_0})\cdot\boldsymbol{x_0} + \boldsymbol{F_C}(\boldsymbol{x_0}) + \boldsymbol{B}\cdot\boldsymbol{u_0} - \boldsymbol{C}\boldsymbol{G}(\boldsymbol{x_0})^{-1}\boldsymbol{B}\cdot\boldsymbol{u_1}) + \boldsymbol{G}(\boldsymbol{x_0})^{-1}\boldsymbol{B}\cdot\boldsymbol{u_1}\cdot t. \quad (32)$$

Note that (32) can be used to give a good initial guess of Newton iteration performed at each time step in the transient analysis defined as Step 8 in Fig. 2.

To obtain the largest eigenvalue of $\boldsymbol{G}^{-1}\boldsymbol{C}$, Arnoldi algorithm [9] [16] shown in Fig. 8 is adopted here. As mentioned before,

Arnoldi algorithm: Embed the eigen values of $\boldsymbol{G}(\boldsymbol{x_0})^{-1}\boldsymbol{C}$ into Hessenberg matrix H from larger side and set the largest eigenvalue of H to the dominant time constant τ_c.

1. Set m as the number of eigenvalues to be obtained.
2. LU-factorization of $\boldsymbol{G}(\boldsymbol{x_0})$: $\boldsymbol{L}\cdot\boldsymbol{U}=\boldsymbol{G}(\boldsymbol{x_0})$
3. Calculate transient response amplitude (28) as $\boldsymbol{v}$:
$$\boldsymbol{v}=\boldsymbol{U}^{-1}\Big(\boldsymbol{L}^{-1}\big(\boldsymbol{F}(\boldsymbol{x_0})\cdot\boldsymbol{x_0}-\boldsymbol{F_C}(\boldsymbol{x_0}) -\boldsymbol{B}\cdot\boldsymbol{u_0}+\boldsymbol{C}\cdot\boldsymbol{U}^{-1}\big(\boldsymbol{L}^{-1}(\boldsymbol{B}\cdot\boldsymbol{u_1})\big)\big)\Big)$$
4. $\boldsymbol{v_1}=\boldsymbol{v}/\|\boldsymbol{v}\|$
5. For $j=1$ To m
6. $\quad\boldsymbol{w}=\boldsymbol{U}^{-1}\left(\boldsymbol{L}^{-1}\left(\boldsymbol{C}\cdot\boldsymbol{v}_{\mathrm{j}}\right)\right)$
7. $\quad$For $i=1$ To j
8. $\quad\quad h_{\mathrm{i,j}}=\boldsymbol{w}^{T}\cdot\boldsymbol{v}_{\mathbf{i}}$
9. $\quad\quad\boldsymbol{w}=\boldsymbol{w}-h_{\mathrm{i,j}}\cdot\boldsymbol{v}_{\mathbf{i}}$
10. $\quad i++$
11. $\quad h_{\mathrm{j+1,j}}=\|\boldsymbol{w}\|$
12. $\quad\boldsymbol{v}_{\mathbf{j+1}}=\boldsymbol{w}/h_{\mathrm{j+1,j}}$
13. $j++$
14. $H=\begin{bmatrix}h_{1,1} & h_{1,2} & \cdots & h_{1,\mathrm{m}}\\ h_{2,1} & h_{2,2} & \cdots & h_{2,\mathrm{m}}\\ 0 & \ddots & \ddots & \vdots\\ 0 & 0 & h_{\mathrm{m,\,m-1}} & h_{\mathrm{m,m}}\end{bmatrix}$
15. Calculate the largest eigenvalue (λ_{max}) of H.
16. Set $\tau_c=\lambda_{max}$.

Fig. 8. Flow chart of Arnoldi algorithm [9] [16]. There is no need to calculate $\boldsymbol{C}^{-1}$ in this algorithm. After Arnoldi algorithm has been completed, the largest eigenvalue of Hessenberg matrix is calculated by such as QR method [24] or directly solving the characteristics equation.

no $\boldsymbol{C}^{-1}$ calculation appears in Fig. 8. This is because not the smallest eigenvalue of $\boldsymbol{C}^{-1}\boldsymbol{G}$ but the largest eigenvalue of $\boldsymbol{G}^{-1}\boldsymbol{C}$ is calculated here. Arnoldi algorithm embeds the eigenvalues of $\boldsymbol{G}^{-1}\boldsymbol{C}$ into Hessenberg matrix from larger side. Then, the largest eigenvalue of the Hessenberg matrix is calculated by such as QR method [24] or directly solving the characteristics equation. Since the device equations may result in complex conjugate eigenvalue pairs, the Hessenberg matrix size m in Fig. 8 should be at least 2. Moreover, according to our experiment, $m=2$ is sufficient to get the largest eigenvalue as is shown as Fig. 13 in Section V *B*.

Arnoldi algorithm shown in Fig. 8 is executed as Step 7 in the transient device simulation flow in Fig. 2. As will be shown in Section VI, LU-factorization in the transient device simulation algorithm and in Arnoldi algorithm is the most CPU-time consuming part. Since LU-factorization is executed at Step 7 and Step 15 in Fig. 2, the total number of LU-factorization per time step is the number of Newton iterations plus one. Therefore, the calculation cost of Arnoldi algorithm is nearly equal to one more extra Newton iteration.

B. Example of Dominant Time Constant Extraction

Fig. 9 (a) shows a 1-dimensional N+P-diode structure and

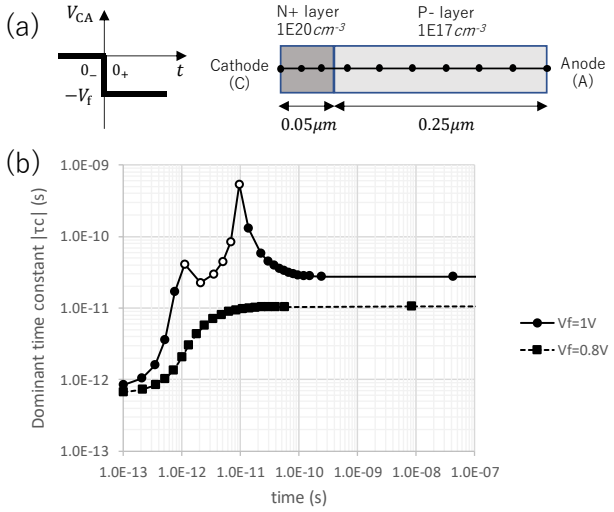

Fig. 9. (a) 1-D N+P-diode structure and forward step bias applied to it. (b) Time variation of the dominant time constants with the progress of the transient analysis after 1V (circles with solid line) and 0.8V (squares with broken line) forward step bias is applied to the N+P-diode. The open circles mean that the dominant time constant is negative.

Fig. 9 (b) shows an example of the variation of the dominant time constant with the progress of the transient analysis after forward step bias is applied to the diode. The filled squares with broken line are the case of moderate (0.8 V) forward bias and the filled and open circles with solid line are the case of strong (1.0 V) forward bias. At the time less than 2E-13 s where the carrier injection has not taken place yet, short dominant time constants (~1E-12 s) are extracted. They result from the product of the initial depletion layer capacitance and the neutral region resistance and are almost independent of the applied biases. On the other hand, after reaching steady state (>1E-10 s), the dominant time constants become larger reflecting the large equivalent inductances due to the conduction current by the injected carriers. In this region, the dominant time constant becomes larger if the applied forward bias is larger (1E-11 s at 0.8 V and 3E-11 s at 1.0 V). The open circles mean that the extracted dominant time constants are negative.

In order to verify the validity of the extracted dominant time constants especially for the negative case, AC analysis [25] is conducted for the same N+P-diode. The small signal admittance of the N+P-diode draws hemi-circle in complex plane according to the variation of angular frequency ω of the small signal as shown in Fig. 10. Whether the hemi-circle stays in the upper part (Fig. 10 (a)) or lower part (Fig. 10 (b)) of the complex plane depends on the bias condition of the diode. If reverse or small forward bias is applied, the depletion layer behaves as capacitance and the small signal equivalent circuit becomes the one shown in Fig. 10 (a). On the other hand, if large forward bias is applied, the depletion layer disappears and the forward conduction current behaves as inductance due to the inertia of carrier traveling to the terminals so that the small signal equivalent circuit becomes the one shown in Fig. 10 (b). AC analysis results of the N+P-diode conducted at each time step of the transient simulation after 1 V forward step bias

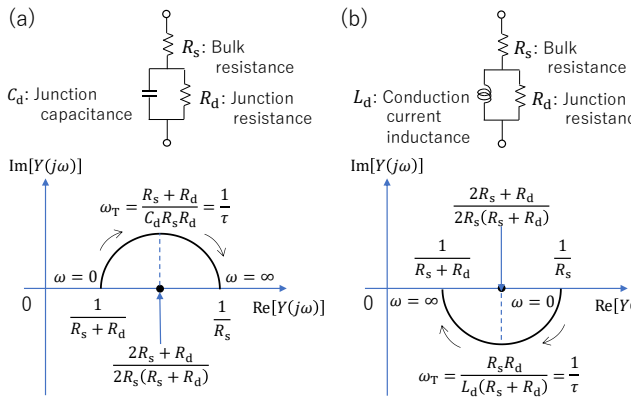

Fig. 10. Small signal equivalent circuit of the N+P-diode and its admittance vector trajectory in complex plane with respect to the angular frequency of the applied small AC signal. Time constant of the equivalent circuit is given by the inverse of the angular frequency at the top or bottom of the hemi-circle. (a) In the case of reverse or small forward bias, the depletion layer behaves as capacitance. (b) In the case of large forward bias, the forward current behaves as inductance.

application are shown in Fig. 11. The small signal frequency is scanned from 1 MHz to 100 GHz at each time step. Before 3.45E-13 s, the diode admittance stays capacitive. Then, it becomes inductive as shown in Fig. 11 (a). After 1.12E-12 s, the hemi-circle flips to the left-half part of the complex plane as shown in Fig. 11 (b), which means the small signal (= differential) bulk resistance R_s becomes negative. This is due to the onset of conductivity modulation in the p-type region such that the higher the carrier injection level becomes, the smaller R_s becomes. Since the time constant of the equivalent circuit is given by the inverse of the small signal angular frequency at the bottom of the hemi-circle as

$$\tau = \frac{1}{\omega_T} = \frac{L_d(R_s + R_d)}{R_s R_d} \tag{33}$$

as shown in Fig. 10 (b), the time constant becomes also negative if $R_s < 0$ and $|R_s| \ll R_d$. In the case of negative time constant, the time step width is usually restricted by the convergence of Newton iteration rather than LTE. This is due to the high sensitivity of the positive feedback system so that even small $\boldsymbol{\delta x}_{k,n+1}$ in (8) causes large variation of $\boldsymbol{F}(\boldsymbol{x}_{k,n+1})$ in (7) at the next Newton iteration. After 1.38E-11 s, the hemi-circle flips back to the right-half part of the complex plane and the time constant also goes back to positive because the conductivity modulation begins to saturate due to the increase of SRH (Shockley-Read-Hall) recombination. Dominant time constants obtained by AC analysis are compared with those by Arnoldi method in Fig. 12. The open symbols correspond to the negative values. These two results are in good agreement with the averaged relative error of 2.8 %, and therefore, it is proved that Arnoldi method can give valid dominant time constants and the negative time constants reflect the actual physical phenomenon. The calculation speed of AC analysis is very slow because it uses complex double precision variables. For example, in the case of our implementation by using Microsoft Visual Basic [26], complex double matrix inversion takes about 10 times longer CPU-time than that of real double matrix inversion. More importantly, it is not easy for AC analysis to extract a

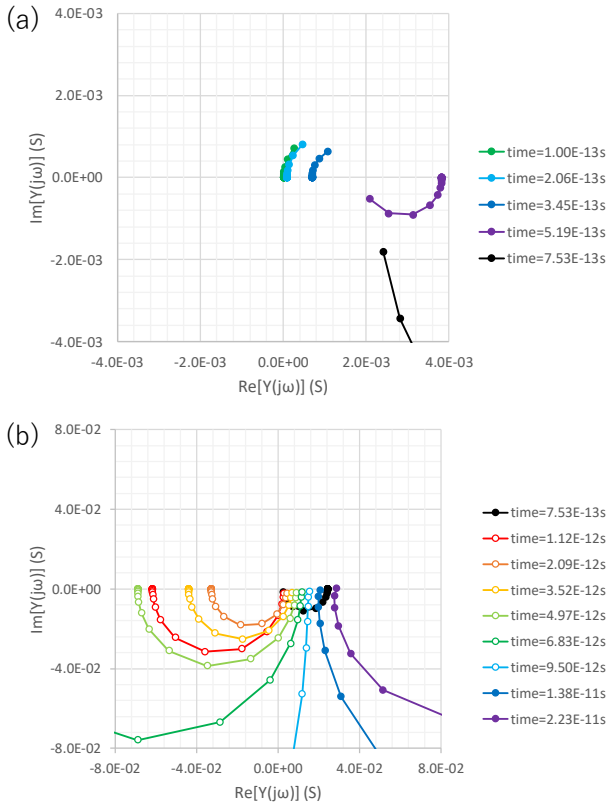

Fig. 11. Small signal admittance vector trajectory of the N+P-diode at each time step of a transient analysis. The small signal frequency is scanned from 1 MHz to 100 GHz. (a) Analysis time = 1E-13 s - 7.53E-13 s. Before 5.19E-13 s, the admittance hemi-circle stays in the upper half part of the complex plane (= capacitive). After that, the hemi-circle moves to the lower half part (= inductive). (b) Analysis time = 7.53E-13 s - 2.23E-11 s. After 1.12E-12 s, the admittance hemi-circle flips from right half part to the left half part of the complex plane (= negative time constant). After 1.38E-11 s, the hemi-circle flips back to the left half part (= positive time constant).

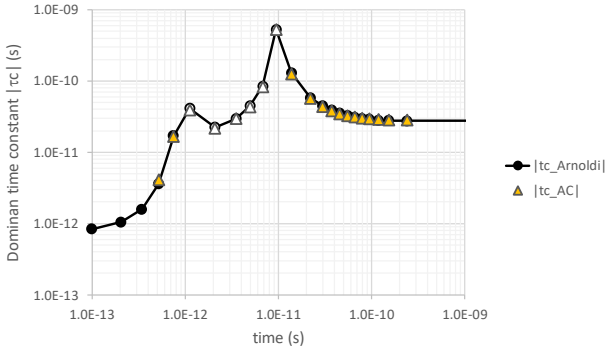

Fig. 12. Comparison of the dominant time constants for 1 V forward bias between Arnoldi method (circles with solid line) and AC analysis (triangles). Open symbols stand for negative time constants. For the analysis time less than 5.19E-13 s, the dominant time constants cannot be extracted by AC analysis because the admittance hemi-circles are not formed with the frequency range of 1 MHz ~ 100 GHz.

dominant time constant if the equivalent circuit has complex topology. On the other hand, Arnoldi method is fast and sure to

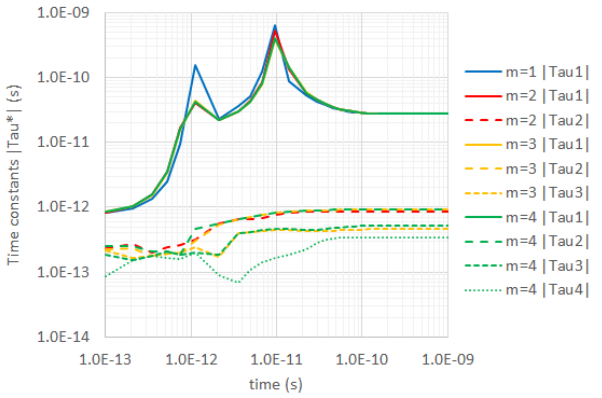

Fig. 13. Behavior of the absolute value of the time constants extracted from larger side by Arnoldi method for $m = 1{\sim}4$. The largest time constant is about one order of magnitude larger than the second one. The largest time constant hardly changes even if m is increased from 2 to 3 or 4 except for the peak portion which has high sensitivity.

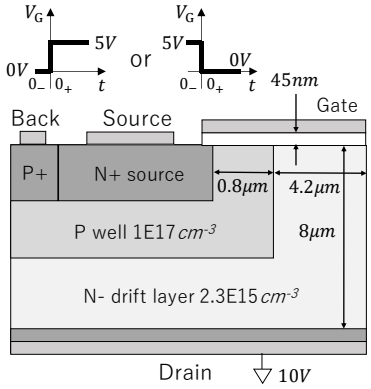

Fig. 14. Power DMOSFET structure [27] used for the performance verification of the proposed method. Step V_G of $\pm$5V is applied for the transient analysis of turn-on and turn-off.

extract a dominant time constant for any arbitrary device.

Fig. 13 shows the behavior of the absolute value of the time constants extracted from larger side by Arnoldi method for $m = 1{\sim}4$. The largest time constant is about one order of magnitude larger than the second one, which indicates the dominant time constant is separated well from the other smaller ones. The dominant time constant hardly changes even if m is increased from 2 to 3 or 4 except for the peak portion which has high sensitivity.

VI. Performance Verification by Transient Simulation of 2-D Power DMOFET

Transient simulation of a 2-dimensional power DMOSFET [27] after step V_G application as shown in Fig. 14 is used for the performance verification of the proposed method. The used computer is equipped with Intel Core i3-4130 CPU of 3.40 GHz clock frequency and 4 GB DRAM. The simulation program is written with Microsoft Visual Basic. Dominant time constant variation and time step width distribution with respect to the analysis time are plotted in Fig. 15. Fig. 15 (a) shows the turn-on case and Fig. 15 (b) shows the turn-off case. The dominant time constants shown by circles with solid lines stay around 1E-10 s which is almost equal to the electron traveling time in the n- drift layer. The open circles in Fig. 15 (a) are negative time constants which appear at the onset of charging the channel (~ 1E-12 s) and at the onset of charging the n- drift layer (~ 2E-11 s). At the onset of charging, the resistance of the channel and the n- drift layer begin to decrease due to electron injection and this causes negative differential resistance that leads to negative dominant time constant. In the case of turn-off, resistance continues to increase due to electron depletion and this keeps the dominant time constant positive (all filled circles) as shown in Fig. 15 (b). Time step width distribution when Exp_LTE_metric is used with relative LTE tolerance (i.e. E_r in (21)) = 1 % for carrier densities and absolute LTE tolerance (i.e. E_a in (21)) = 1 % of the net doping densities is also shown by open triangles with dashed lines in Fig. 15. Multiple time step

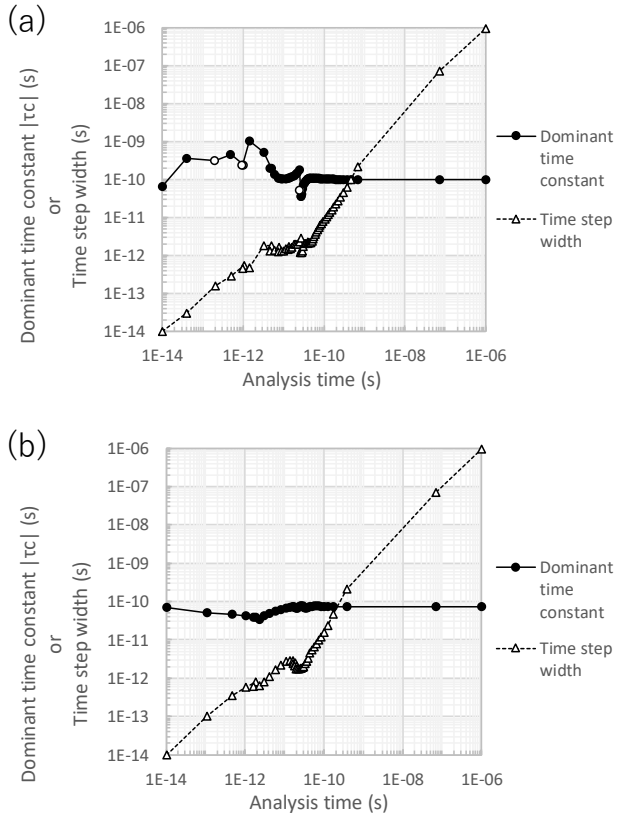

Fig. 15. Dominant time constants variation (circles with solid line) and time step widths distribution (triangles with broken line) with respect to the analysis time. Dominant time constants stay around 1E-10 s which is almost equal to the electron traveling time in the n- drift layer. Open circles correspond to the negative time constants. (a) Turn-on case. (b) Turn-off case.

widths at the same analysis time mean that the time step width was rejected there as the Exp_LTE_metric exceeded the LTE tolerance. In the turn-on case, there are dense time steps between the channel formation (~ 3E-12 s) and transient decay (~ 3E-10 s). Similarly, in the turn-off case, time steps are dense between the onsets of channel depletion (~ 1E-12 s) and n- layer

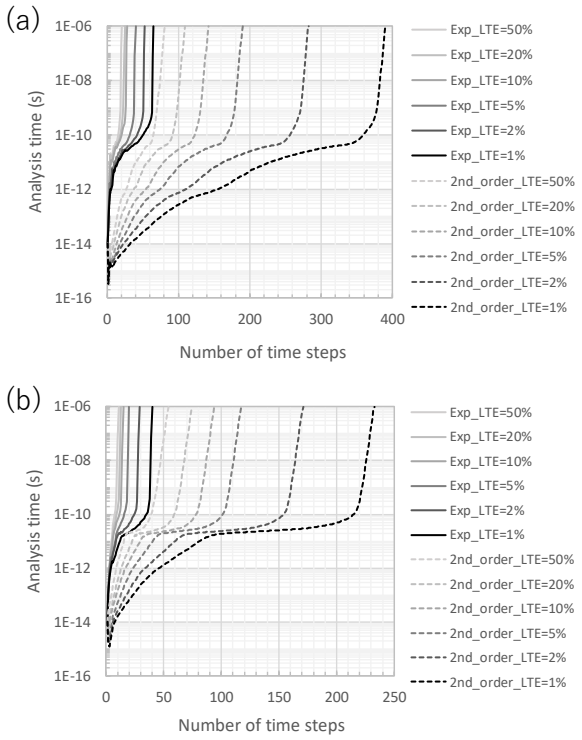

Fig. 16. Analysis time progress with respect to the number of the time steps for relative LTE tolerance (E_r) of 1 ~ 50 % for carrier densities and absolute LTE tolerance (E_a) of 1-50 % of the net doping densities. (a) Turn-on case. (b) Turn-off case.

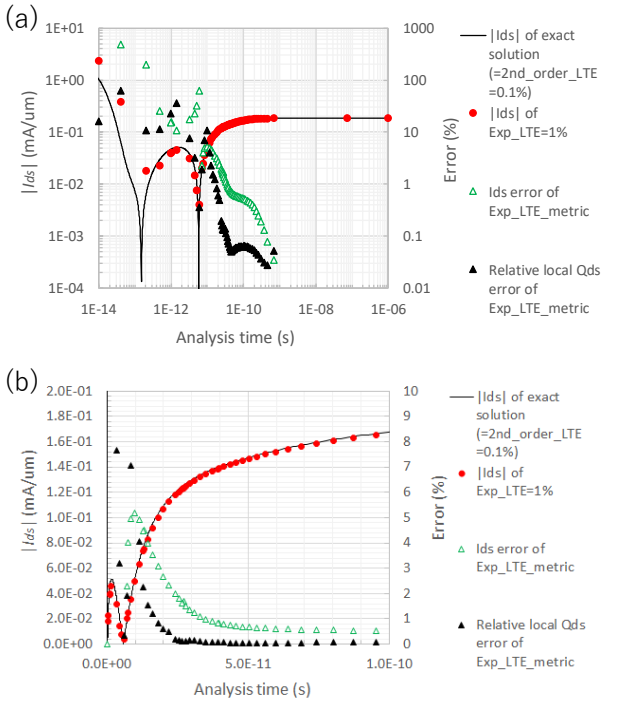

Fig. 17. Drain current (I_{ds}) comparison for turn-on case between the solution with Exp_LTE_tolerance=1% (filled circles) and the exact solution obtained by setting 2nd_order_LTE_tolerance to 0.1% (solid line). Open triangles stand for the drain current errors. Filled triangles stand for the relative local drain charge error defined by (34). (a) Whole response with log-scale. (b) Main response with linear-scale.

depletion (~ 1E-10 s). In overall, the time step width is almost monotonically increasing in Fig. 15. This is due to the fact that $|\dot{x}(t_0)|$ in (20) is almost monotonically decreasing.

Analysis time progresses with respect to the number of time steps are shown in Fig. 16 for E_r of 1-50 % for carrier densities and E_a of 1-50 % of the net doping densities. Fig. 16 (a) shows the turn-on case and Fig. 16 (b) shows the turn-off case. Exp_LTE_metric (solid lines) shows about 5 times more rapid time progress at 1E-6 s than 2nd_order_LTE_metric (broken lines). This comes from the following two reasons. First, since 2nd_order_LTE_metric ignores higher order Taylor expansion terms, cancellation effect between Taylor expansion terms in the case of positive dominant time constant is not included and therefore it overestimates the actual LTE. Second, since Exp_LTE_metric focuses on the dominant time constant, shorter time responses which are less important for the actual device operation are not highly considered.

Fig. 17 and 18 compares the drain current (I_{ds}) between the solution with Exp_LTE_tolerance = 1 % (filled circles) and the exact solution obtained by setting 2nd_order_LTE_tolerance to 0.1 % (solid line). Open triangles stand for the drain current error of Exp_LTE_metric case. Filled triangles stand for the relative local drain charge error in a time step with respect to the cumulative drain charge defined as:

$$Relative\_local\_Q_{ds}\_error \stackrel{\text{def}}{=} \left| \frac{\int_{t_0}^{t_0+\Delta t} (I_{ds\_Exp\_LTE} - I_{ds\_exact})dt}{\int_0^{t_0+\Delta t} I_{ds\_exact}dt} \right| \quad (34)$$

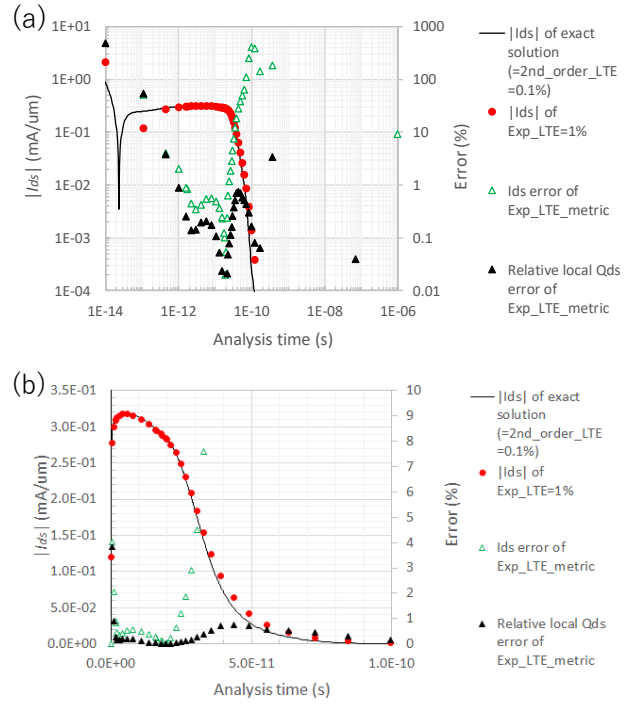

Fig. 18. Drain current (I_{ds}) comparison for turn-off case between the solution with Exp_LTE_tolerance=1% (filled circles) and the exact solution obtained by setting 2nd_order_LTE_tolerance to 0.1% (solid line). Open triangles stand for the drain current errors. Filled triangles stand for the relative local drain charge error defined by (34). (a) Whole response with log-scale. (b) Main response with linear-scale.

Fig. 17 shows the turn-on case and Fig. 18 shows the turn-on case. Since step V_G biases are applied here, displacement current coupling between the terminals occurs through the series capacitances in the middle row of Fig. 5 (a), and this coupling produces femto-second order current responses. From a practical view point, ideal step V_G application is unrealistic and the obtained results in femto-second order are physically dubious considering the classical physics models used in the simulation. Total amount of the transported or stored charges by such an ultra-fast current response is negligible compared with that of the later dominant current response. Therefore, less current accuracy in the ultra-fast response part is not a problem from a view point of total charge conservation. Exp_LTE_metric just ignores such an ultra-fast time response and this is the reason why the drain current error (open triangles) and the relative local drain charge error (filled triangles) become large in this duration. In the case of turn-on (Fig. 17), both the drain current error and the relative local drain charge error become less than 3 % after the analysis time reaches about 1/5 of the dominant time constant. In the case of turn-off (Fig. 18), both the drain current error and the relative local drain charge error also stay less than 3 % during the electron discharge (1E-12 s - 3E-11 s). Steep increase of the drain current error in the decay part of the turn-off is observed. However, the relative local drain charge error in this part is still kept less than 1% because the drain current level is low.

CPU-time between 2nd_order_LTE_metric and Exp_LTE_metric is compared for 1-50 % LTE tolerance in Fig. 19. Fig. 19 (a) shows the turn-on case and Fig. 19 (b) shows the turn-off case. CPU-time per time step of Exp_LTE_metric when initial guess is applied before Newton iteration by using (32) (filled triangles with solid lines) is about 40 % longer than that of 2nd_order_LTE_metric (filled triangles with dashed lines). This is mainly due to the calculation cost of the dominant time constant by Arnoldi method and partly due to the increase in the number of Newton iterations induced by the larger time step width. However, the total CPU-time of Exp_LTE_metric with the initial guess (filled circles with solid line) is as small as about 27 % of 2nd_order_LTE_metric (filled circles with dashed line). By using Exp_LTE_metric, the total number of time steps reduces down to about 20 % of 2nd_order_LTE_metric as shown in Fig. 16 and this is the main cause of CPU-time reduction.

Smaller number of time steps means larger time step widths per time step, which may induce convergence problem of Newton iteration. Therefore, it is important to use a good initial guess calculated by (32) to assure stable simulation. Fig. 20 compares the number of Newton iterations per time step for Exp_LTE_metric = 1 % between the initial guess by (32) and simple use of the solution at the preceding time step. By adopting the initial guess by (32), the number of Newton iterations and therefore also their CPU-time can be reduced by 25%. On the other hand, the smaller the number of Newton iteration becomes, the more significant the relative calculation overhead of the LU-factorization in Arnoldi algorithm becomes. Consequently, the total calculation cost per time step with the Exp_LTE_metric is larger than that with

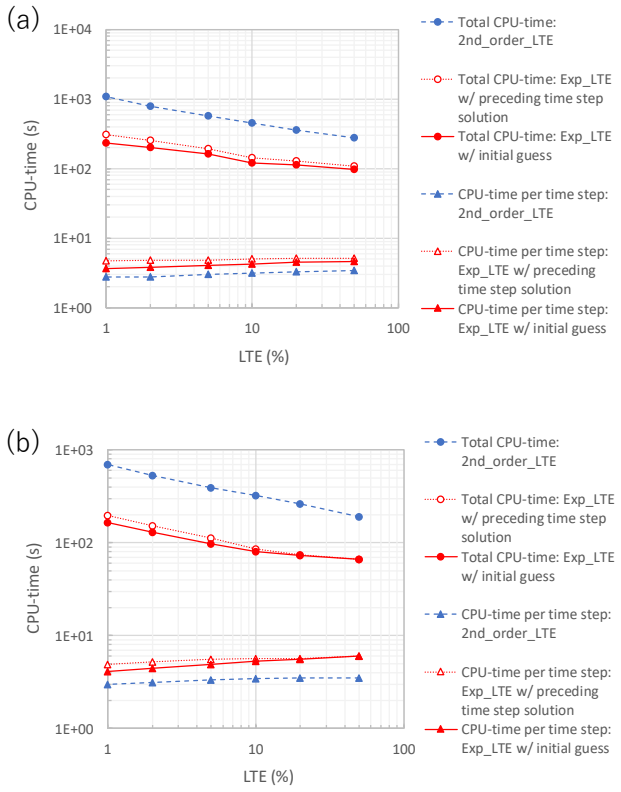

Fig. 19. CPU-time comparison between 2nd_order_LTE_metric and Exp_LTE_metric for 1-50% LTE tolerance. Solid circles and solid triangles with solid line correspond to the case that initial guess is applied before Newton iteration by using (32). (a) Turn-on case. (b) Turn-off case.

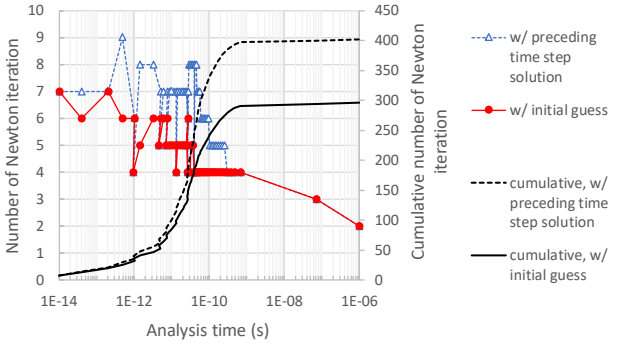

Fig. 20. Comparison of the number of Newton iteration between with initial guess by (32) and simple use of the solution at the preceding time step for turn-on case. Even for the wide time step widths produced by Exp_LTE_metric, good initial guess helps the convergence.

2nd_order_LTE_metric as shown in Fig. 19.

Mesh size dependence of the CPU-time for Exp_LTE_metric and 2nd_order_LTE_metric with 1% LTE tolerance is shown in Fig. 21. Here, $n_x \times n_y$ rectangular mesh is used and therefore, the total mesh size (n_m) is $n_m = n_x n_y$ and the resultant matrix size is $3n_m \times 3n_m$. If the variables are reordered as ${}^t(\psi(r_1,t_k), n(r_1,t_k), p(r_1,t_k), \psi(r_2,t_k), n(r_2,t_k), \cdots)$ as

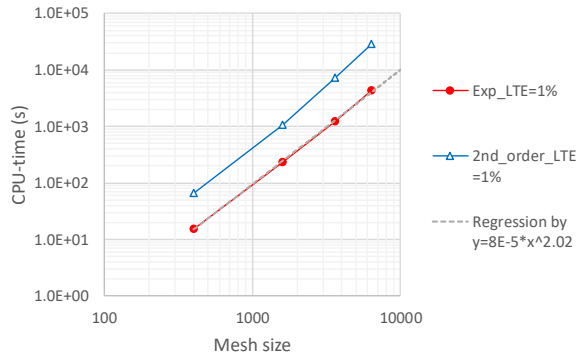

Fig. 21. Mesh size dependence of the CPU-time for Exp_LTE_metric and 2nd_order_LTE_metric with 1% LTE tolerance. The CPU-time is almost proportional to the square of the mesh size, which means LU-factorization of the band matrix is the dominant calculation cost.

explained in Section II, the band width of the resultant matrix is almost equal to $3\sqrt{n_m}$. Although this matrix is sparse as it has only 15 diagonal and off-diagonal element stripes, all the elements within the band width are filled in after LU-factorization. Therefore, the time complexity of LU-factorization is $O\left(n_m \cdot \sqrt{n_m} \cdot \sqrt{n_m}\right) = O(n_m^2)$. The time complexity of the other matrix-vector operations such as forward-backward substitution or matrix-vector product is $O\left(n_m \cdot \sqrt{n_m}\right) = O(n_m^{1.5})$. Since the CPU-time is almost proportional to $n_m^{2.02}$ as shown in Fig. 21, LU-factorization of the band matrix is dominant in the calculation cost.

VII. Discussion

It is also possible to adopt the proposed Exp_LTE_metric to other devices than power DMOSFET. However, in advanced logic CMOS, the response time of the whole device structure is equal to the inversion layer forming time which is less than 1E-12 s. Therefore, CPU-time reduction will not be significant if Exp_LTE_metric is adopted. For advanced logic CMOS devices, the following two cases may be effective: 1) Soft-error or latch up caused by charged particle irradiation, which is restricted by the response time of the whole well structure. 2) The case that the whole circuit response is restricted by the response time of external load. Still, the most effective application is power devices which need to charge and discharge long n- drift layer, and whose external load has long response time.

In Exp_LTE_metric, only one dominant time constant is considered. This is enough for power devices whose dominant time constants are well-separated from the other smaller time constants. However, there may be problems such that multiple time constants are uniformly scattered in an interested time range and/or fast responses also need to be calculated accurately. To ensure the accuracy throughout the entire simulation time range for general problems, a mechanism of choosing appropriate time constants for Exp_LTE_metric depending on the analysis time may be necessary.

VIII. Conclusions

A new accurate Exp_LTE_metric for time step control for power device transient simulation is derived by utilizing dominant time constant information of the whole device structure. The dominant time constant is extracted as the negative inverse of the smallest eigen value of the matrix which appears in the matrix exponential term in the formal solution of the linearized device state equation. By using the proposed Exp_LTE_metric, CPU-time of the 2-dimensional power DMOSFET transient simulation successfully decreases down to 27 % of the conventional 2nd_order_LTE_metric with assuring the current accuracy of the dominant transient response. Initial guess by using the formal solution of the linearized device state equation improves the convergence of Newton iteration by 25% for the large time step widths derived by Exp_LTE_metric.

Acknowledgment

The authors would like thank Prof. Daisuke Ueda of Kyoto Institute of Technology for his kind lecture on the power DMOSFET device structure and operation.

References

[1] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Wien New York: Springer-Verlag, 1984, ch. 2

[2] E. Buturla, P. Cottrell, B. Grossman, and K. Salsburg, "Finite-element analysis of semiconductor devices: The FIELDAY program," *IBM J. Res. Develop.*, vol. 25, no. 4, pp. 218-231, 1981.

[3] D. Scharfetter and H. Gummel, "Large-signal analysis of a silicon Read diode oscillator," *IEEE Trans. Electron Devices*, vol. ED-16, no. 1, pp. 64-77, 1969.

[4] H. Togawa, *Numerical Calculation of Differential Equations* (in Japanese), 1st ed. Tokyo, Japan: Ohmsha, 1981, ch. 5, sec. 1, p.38, ch. 12, sec. 4, p.103.

[5] R. Bank, W. Coughran, Jr., W. Fichtner, E. Grosse, D. Rose, and R. Smith, "Transient Simulation of Silicon Devices and Circuits," *IEEE Trans. Electron Devices*, vol. ED-32, no. 10, pp. 1992-2007, 1985.

[6] H. Togawa, *Numerical Calculation of Differential Equations* (in Japanese), 1st ed. Tokyo, Japan: Ohmsha, 1981, ch. 4, sec. 1-2, pp.28-29, ch. 6, sec. 2, p.53.

[7] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *NUMERICAL RECIPES in C* (Japanese Ed.), 1st ed. Cambridge University Press, Tokyo, Japan: Gijutsu Hyoron-sha, 1993, ch. 9, sec. 1, pp.257-258.

[8] T. Ohtsuki, and K. Kani, "A Unified Modeling Scheme for Semiconductor Devices With Application of State-Variable Analysis," *IEEE Trans. Circuit Theory*, vol. Ct-17, no. 1, pp. 26-32, 1970.

[9] S. Kumashiro, T. Kamei, A. Hiroki, and K. Kobayashi, "An Accurate Metric to Control Time Step of Transient Device Simulation by Matrix Exponential Method," in *Proc. SISPAD*, Kamakura, Japan, 2017, pp 37-40.

[10] S. Kumashiro, "Transient Simulation of Passive and Active VLSI Devices Using Asymptotic Waveform Evaluation," Ph.D. dissertation, ECE, CMU, Pittsburgh, PA, USA, 1992, pp. 84-90.

[11] H. Read, S. Kumashiro, and A. Strojwas, "Efficient transient device simulation with AWE macromodels and domain decomposition," *IEICE Trans. Electron.*, vol. E77-C, no. 2, pp. 236-247, 1994.

[12] C. Moler, and C. Van Loan, "Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later," *SIAM Review*, vol. 5, no. 1, pp. 3-49, 2003.

[13] L. Pillage, and R. Rohrer, "Asymptotic Waveform Evaluation for Timing Analysis," *IEEE Trans. Computer-Aided Design*, vol. CAD-9, no. 4, pp. 352-366, 1990.
[14] Y. Saad, "Analysis of some Krylov subspace approximations to the matrix exponential operator," *SIAM J. Numer. Anal.*, vol.29, no. 1, pp. 209-228, 1992.
[15] P. Feldmann, and R. Freud, "Efficient Linear Circuit Analysis by Pade Approximation via Lanczos Process," *IEEE Trans. Computer-Aided Design*, vol. CAD-14, no. 5, pp. 639-649, 1995.
[16] H. Zhuang, X. Wang, Q. Chen, P. Chen, and C. Cheng, "From circuit theory, simulation to SPICE^{Diego}: A matrix exponential approach or time domain analysis of large-scale circuits," *IEEE Circuits and Systems Magazine*, pp. 16-34, 2016.
[17] S. Weng, Q. Chen, and C. Cheng, "Time-Domain Analysis of Large-Scale Circuits by Matrix Exponential Method With Adaptive Control," *IEEE Trans. Computer-Aided Design*, vol. CAD-31, no. 8, pp. 1180-1193, 2012.
[18] S. Weng, Q. Chen, N. Wong, and C. Cheng, "Circuit Simulation via Matrix Exponential Method for Stiffness Handling and Parallel Processing," in *Proc. ICCAD*, San Francisco, CA, USA, 2012, pp. 407-414.
[19] H. Zhuang, S. Weng, J. Lin, and C. Cheng, "MATEX: A Distributed Framework for Transient Simulation of Power Distribution Networks," in *Proc. DAC*, San Francisco, CA, USA, 2014, pp. 1-6.
[20] H. Zhuang, W. Yu, I. Kang, X. Wang, and C. Cheng, "An Algorithmic Framework for Efficient Large-Scale Circuit Simulation Using Exponential Integrators," in *Proc. DAC*, San Francisco, CA, USA, 2015, pp. 1-6.
[21] Q. Mei, W. Schoenmaker, S. Weng, H. Zhuang, and C. Cheng, "An Efficient Transient Electro-Thermal Simulation Framework for Power Integrated Circuits," *IEEE Trans. Computer-Aided Design*, vol. CAD-35, no. 5, pp. 832-843, 2016.
[22] H. Zhuang, W. Yu, S. Weng, I. Kang, J. Lin, X. Zhang, R. Coutts, and C. Cheng, "Simulation Algorithm With Exponential Integration for Time-Domain Analysis of Large-Scale Power Delivery Networks," *IEEE Trans. Computer-Aided Design*, vol. CAD-35, no. 10, pp. 1681-1694, 2016.
[23] X. Wang, H. Zhuang, and C. Cheng, "Exploring the Exponential Integrators with Krylov Subspace Algorithms for Nonlinear Circuit Simulation," *Proc. ICCAD*, Irvine, CA, USA, 2017, pp.163-168.
[24] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *NUMERICAL RECIPES in C* (Japanese Ed.), 1st ed. Cambridge University Press, Tokyo, Japan: Gijutsu Hyoron-sha, 1993, ch. 11, sec. 6, pp.369-375.
[25] S. Laux, "Techniques for Small-Signal Analysis of Semiconductor Devices," *IEEE Trans. Electron Devices*, vol. ED-32, no. 10, pp.2028-2037, 1985.
[26] Microsoft Developer Network (in Japanese) [Online]. Available: https://msdn.microsoft.com/ja-jp/library/system.numerics.complex(v=vs.110).aspx
[27] D. Fuoss, "Vertical DMOS power field-effect transistors optimized for high-speed operation," in *IEDM Tech. Dig*., San Francisco, CA, USA, 1982, pp. 250-253.

Shigetaka Kumashiro (M'90–SM'03) received his B. E. and M. E. degrees from the University of Tokyo in 1981 and 1983, respectively, and his Ph. D. degree from Carnegie Mellon University in 1992. He worked for NEC Corporation from 1983 to 2002, NEC Electronics Corporation from 2002 to 2010 and Renesas Electronics Corporation from 2010 to present, respectively. He has also worked with STARC and MIRAI-Selete. He has been working in the field of the modeling, simulation and reliability of ULSI processes and devices. Now he is a principal specialist of Renesas Electronics Corporation. In 2016, he was also appointed as a research professor of Green Innovation Center of Kyoto Institute of Technology. He received ASP-DAC best paper award in 2001. Dr. Kumashiro is a member of IEICE and JSAP.

Tatsuya Kamei received the B. E. degree from Kyoto Institute of Technology, Japan in 2017, where he is currently pursuing the M. E. degree with the Department of Electronic Engineering.
His current research interests include device simulation for Power MOSFETs.

Akira Hiroki (M'99) received the B.S and M.S. degrees in physics from Chiba University in 1983 and 1985, respectively, and the Ph.D. degree in electrical engineering from Osaka University in 2000.
In 1985, he joined the Semiconductor Research Center, Matsushita Electric Industrial Co., Ltd., Osaka, Japan, where he engaged in device modeling and simulation. Since 2002, he has been an associate professor with Kyoto Institute of Technology, where he has been involved in device modeling and quantum transport.

Kazutoshi Kobayashi (M'96) received his B. E., M. E. and Ph.D. in Electronic Engineering from Kyoto University, Japan in 1991, 1993, 1999, respectively.
Starting as an Assistant Professor in 1993, he was promoted to associate professor in the Graduate School of Informatics, Kyoto University, and stayed in that position until 2009. For two years during this time, he acted as associate professor of VLSI Design and Education Center (VDEC) at the University of Tokyo. Since 2009, he has been a professor at Kyoto Institute of Technology. While in the past he focused on reconfigurable architectures utilizing device variations, his current research interest is in improving the reliability (Soft Errors, Random Telegraph Noise, Bias Temperature Instability and Plasma Induced Damage) of current and future VLSIs. He started a research related to gate drivers for power transistors since 2013. He was the recipient of the IEICE best paper award in 2009 and the IRPS best poster award in 2013.