Received 23 January 2018; accepted 27 January 2018. Date of publication 5 February 2018; date of current version 23 February 2018.
The review of this paper was arranged by Editor C. C. McAndrew.

Digital Object Identifier 10.1109/JEDS.2018.2801301

On the Formulation of Self-Heating Models for Circuit Simulation

LINING ZHANG [1,2] (Member, IEEE), DEBIN SONG [3], YING XIAO [2], XINNAN LIN [3], AND MANSUN CHAN [2] (Fellow, IEEE)

1 College of Electronic Science and Technology, Shenzhen University, Shenzhen 518061, China
2 Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong
3 Shenzhen Key Laboratory of Advanced Electron Device and Integration, ECE, Peking University Shenzhen Graduate School, Shenzhen 518061, China

CORRESPONDING AUTHOR: L. ZHANG (e-mail: lnzhang@ieee.org)

This work was supported in part by the Innovation and Technology Fund under Grant GHP/055/14SZ, in part by the Area of Excellence from UGC Hong Kong under Grant AOE/P-04/08, and in part by the NSFC under Grant 61704144.

ABSTRACT Different approaches to implement self-heating effects in a compact model are evaluated. The traditional approach using a subcircuit with the addition of an internal node can lead to significant increase in the simulation time. In contrast, by directly solving self-heating equations, the internal node is eliminated in the circuit Jacobian matrix. The resulting simulation time can be shortened in principle up to 60% or more without sacrificing the accuracy. The accuracy and time for self-heating simulations formulated using different approaches are compared in this paper to study their tradeoff. In addition, a generic approach to eliminate the need for internal nodes is proposed and demonstrated using the non-quasi-static effect model.

INDEX TERMS Self-heating, internal-node free, FinFETs circuits, circuit self-heating.

I. INTRODUCTION

Self-heating is getting more significant with the increased Joule heat density and reduced device volume. It is particularly important for transistors with a small channel volume surrounding by a large volume of insulator with slower thermal conduction such as silicon-on-insulator (SOI) MOSFETs [1] and multi-gate devices [2]. It is anticipated that all device models for the current technology nodes and beyond have to include self-heating effects.

The conventional approach to implement self-heating in many standard models [3]–[6] is to use a sub-circuit to solve the heat equation and track the temperature with an internal node. However, the use of internal node can significantly increase the simulation time in solving the circuit equations. For example, in a N stage ring-oscillator, the number of nodes is only $N+2$ without any internal node. When self-heating is implemented with a sub-circuit, one internal node is added to both the NMOSFET and PMOSFET of the inverter chain leading to $3N+2$ nodes in the same ring oscillator circuit. Considering solving the sparse Jacobian matrix is a $O(n^{1.4})$ [7] process, the required simulation time is almost tripled.

A method to speed up self-heating simulations is desirable. One possible way is to use the temperature solution of an earlier time point for calculating current at the present time point. However, it sacrifices the accuracy especially for signals with abrupt changes. By studying the transient simulation algorithm in SPICE, it is possible to incorporate the self-heating equations into the integration algorithm in solving the linear differential equations and eliminate the need for a sub-circuit without sacrificing accuracy. In this work, the accuracy and speed of these approaches to implement self-heating and potentially other circuits that require internal nodes are studied.

II. SOLUTION TO THE TEMPERATURE EQUATION

In compact models with the device temperature rise ΔT, the thermal capacitance C_{th}, the thermal resistance R_{th} and the Joule power P_{th} are related by the differential heat equation [3]–[5]:

$$\frac{d\,(C_{th}\cdot\Delta T)}{dt}+\frac{\Delta T}{R_{th}}=P_{th} \tag{1}$$

It resembles the form of a simple RC circuit equation:

$$\frac{d(C\cdot V)}{dt}+\frac{V}{R}=I_{cons} \tag{2}$$

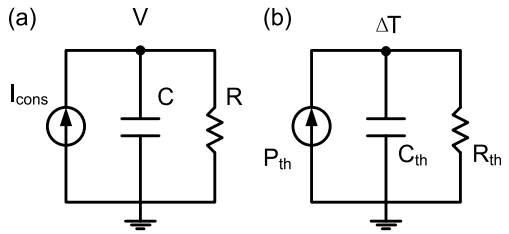

FIGURE 1. Schematic diagrams of (a) a RC circuit with a constant current source I_{cons}, and (b) the sub-circuit with a thermal node for self-heating simulations.

Eq. (2) can be represented by the circuit diagram as shown in Fig. 1 (a) with one node and three components.

By comparing the similarity between Eq. (1) and Eq. (2), a technique was developed to find the solution of equation (1) with an equivalent circuit [8], [9] as shown in Fig. 1 (b). The increase in temperatures is incorporated in the calculation of the device current to achieve a self-consistent iterative solution. Nowadays the sub-circuit (SC) method is almost a default self-heating simulation method due to its simplicity but its only drawback is an increase in simulation time.

In order to eliminate the internal node, current and voltage solutions from a previous time (PT) point at t_{i-1} can be used to calculate the temperature at the present time t_i assuming the temperature difference over a short period of time between two time points is very small. That is, the current in the present time point is calculated using $I(V(t_i), T(t_{i-1}))$ instead of $I(V(t_i), T(t_i))$. Due to the assumption of small temperature change, this approach may sacrifice accuracy especially for fast varying signal with noticeable difference between solutions at t_{i-1} and t_i. However, it serves as a reference for generic comparisons.

By more carefully examining the sub-circuit, the underlying equation is given by:

$$\left(\frac{C_{th}}{h}+\frac{1}{R_{th}}\right)\Delta T_i = I_i\left[\Delta T_i\right]V_i + \frac{C_{th}}{h}\Delta T_{i-1} \tag{3}$$

where $h = t_i - t_{i-1}$ is the time step to solve the differential equation numerically, the sub-script *i or i-1,* represents the time point. It is an implicit equation with ΔT_i appears on both side of the equation. And this equation is solved by the Newton-Raphson (NR) method in SPICE with the iterative equation below:

$$\left(\frac{C_{th}}{h}+\frac{1}{R_{th}}\right)\Delta T_i^{n+1} = I_i\left[\Delta T_i^{n}\right]V_i^{n} + \frac{C_{th}}{h}\Delta T_{i-1} \tag{4}$$

where the superscript n or $n+1$ means the number of successive iterations. The process can be illustrated in the flowchart as shown in Fig. 2 (a). It is a straightforward method and the only drawback is the increase in simulation time due to the incorporation of equation (4) into the Jacobian matrix resulting in a larger matrix size to solve. When an alternative integration method instead of the backward Euler one is used,

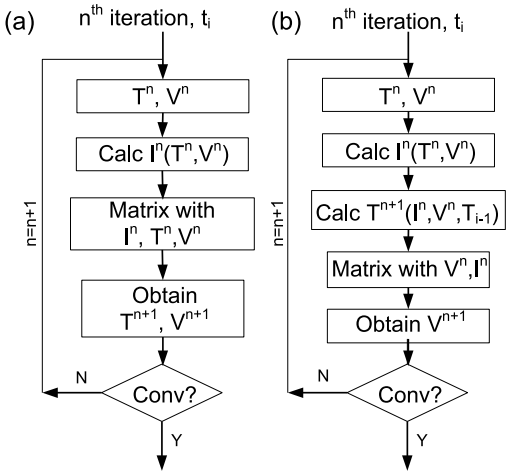

FIGURE 2. The flow chart for implementations of (a) the sub-circuit method and (b) the proposed direct evaluation method. Self-consistency between temperature and current is achieved via the NR iterations of SPICE.

e.g., the trapezoidal method, the equation for self-heating still resemble to Eq. (3):

$$\left(\frac{2C_{th}}{h}+\frac{1}{R_{th}}\right)\Delta T_i = I_i\left[\Delta T_i\right]V_i + 2C_{th}\left[\frac{\Delta T_{i-1}}{h}+T'_{i-1}\right] \tag{5}$$

where T' is the time derivative of the temperature rise. The iterative equation solved in SPICE is very similar to Eq. (4) as given below.

$$\left(\frac{2C_{th}}{h}+\frac{1}{R_{th}}\right)\Delta T_i^{n+1} = I_i\left[\Delta T_i^{n}\right]V_i^{n} + 2C_{th}\left[\frac{\Delta T_{i-1}}{h}+T'_{i-1}\right] \tag{6}$$

The differences are the time used in the integration (h or $h/2$) and the number of derivatives from previous time points. The coefficient of ΔT_i^{n+1}, current I_i and the constant terms are entered into the circuit matrix for a solution. The process is the same as that present in Fig. 2(a). If the R_{th} is temperature dependent [10], an equation similar to Eq. (4) or (6) is obtained with Taylor expansion around ΔT_{i-1}.

More careful examinations of Eq. (4) or (6) reveal that all the variables are known and except ΔT_i^{n+1}. Therefore, ΔT_i^{n+1} can be solved directly without going through the solution of the Jacobian matrix. In this approach, the flowchart is shown in Fig. 2 (b). We shall call the method described in Fig. 2 (b) the Direct Evaluation (DE) method. By comparing Fig. 2 (a) and (b), the direct evaluation method is shown to be identical to the sub-circuit method but an entry to the Jacobian matrix is eliminated, resulting in a smaller matrix to be solved. Therefore, we expect the DE method can significantly increase the simulation speed without sacrificing the accuracy.

Self-heating simulations under DC analysis are performed under the same framework of Fig. 2 (b), with the value at t_i

and T_{i-1} removed. For simulations in the frequency domain, NR iterations are not required as the Jacobian matrix is linearized. The corresponding equation in the circuit matrix for the temperature node with the sub-circuit method including both real an imaginary part is given by

$$\left(jwC_{th} + {}^{1}\!/_{R_{th}}\right)\Delta T = a \cdot v \tag{7}$$

where w is the angular frequency and v represents the small signal voltage. The coefficients of nodal voltages, a, are not expanded to avoid lengthy expressions. Similar to equation (4) and (6), the temperature ΔT can be expressed with the nodal voltages thus the algorithm in Fig. 2(b) is also applied to evaluate equation (7) without relying on an internal node. So we apply the DE method as an alternative to implement the sub-circuit in the frequency domain simulations.

III. IMPLEMENTATIONS AND COMPARISONS

To verify and compare the effectiveness of the different approaches to implement the self-heating equations, the DE method and the PT method are implemented into an open source simulator, Ngspice. We use some existing models with self-heating implemented using the SC approach as a reference. The difference between the PT, DE or SC is only the entry to the Jacobian matrix. As indicated by Eq. (4) and (6), different integration methods only differ by the way that they calculate Jacobian matrix entries, such as using different integration time step (such as *h or h*/2) and number of derivatives at pervious time points. The coefficient calculation is handled by a function call in SPICE. For example, in Ngspice, it is the NIintegrate function. Therefore, the same process can be used to evaluate the temperature with the DE method, regardless of integration methods. For the power dissipation, V^{n+1} may be used for the current I_i based on Taylor expansions, creating non-diagonal matrix entries in the temperature part of Jacobian matrix. To implement the DE method, Eq. (4) or (6) are substituted into the nodal voltage equations for the matrix filling. The temperature is updated at the beginning of $n^{th}+1$ iteration for model evaluations as shown in Fig. 2 (b). For AC simulations, Eq. (7) is used to calculate the temperature change either with an additional Jacobian matrix entry using the SC approach or without the entry in the DE approach.

Fig. 3 plots the DC simulations of a single NMOSFET with and without including the self-heating models. The negative differential-resistance of the output characteristics in Fig. 3 (a) is well reproduced by all the SC, PT and DE methods with exactly the same output. This result is expected as the three methods are just different representations of the same equation in the iterative solution process. The conclusion is more obvious when we look at the steady-state temperature change as a function of bias which is shown in Fig. 3 (b). The exactly same temperature rise is predicted using all three methods.

To study the speed of different methods for self-heating, it is necessary to simulate a complete circuit rather than

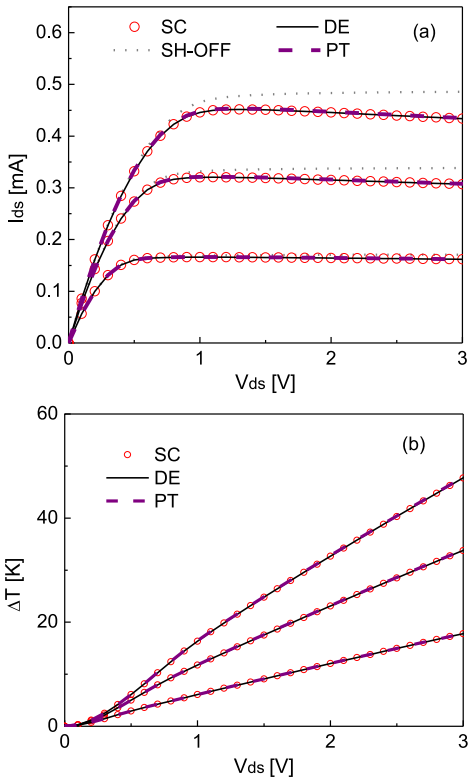

FIGURE 3. DC simulations of self-heating in a single NMOSFET with different methods. (a) The negative differential resistance is reproduced accurately with the DE and PT methods, and (b) the same temperature rise is obtained.

individual devices. A typical ring oscillator as shown in Fig. 4 (a) is simulated. Fig. 4 (b) shows that both PT and DE can significant reduce the simulation time by more than 60%. As the number of iterations (scaled by 10,000) for all three methods are similar, the reduction of simulation time is achieved from the reduction of the size of the Jacobian matrix from $3N+2$ nodes to $N+2$ nodes where N is the number of stage of the ring oscillator. It also agrees with the predicted computation time that grows with increase of the number of nodes with $O(n^{1.4})$ under modern sparse matrix solving algorithm. The DE method can be used to implement self-heating without any penalty in simulation convergence. As the number of circuit elements in larger scale circuits is around 2 to 4 times of the number of nodes, the DE method can reduce the size of the Jacobian to one third and results in significant reduction in simulation time.

To study the accuracy of the different methods to model the self-heating, we study circuits which are more susceptible to temperature changes, such as analog circuits [9]. A 2-stage amplifier as shown in Fig. 5 (a) is simulated under normal operation and large signal switching at both low and high input pulse rate to examine the time dependent self-heating behavior. Fig. 5 (b) shows the time evolution of the temperature change. The different data point at the same time point illustrated the iterative values and 2-3 iterations are required to achieve convergence. It shows that the sub-circuit

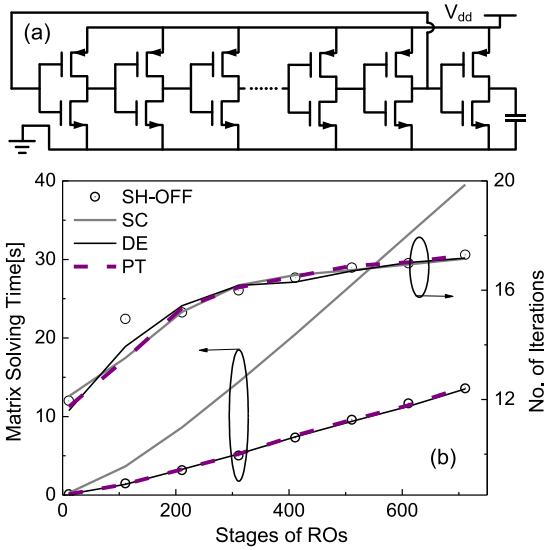

FIGURE 4. **(a) A schematic diagram of the ring-oscillator (ROs) and (b) the circuit solving time and total number of NR iterations with different approaches in simulating the ROs are compared, with a significant speed up from the direct evaluation method.**

and the DE approaches give exactly the same temperature change at all iteration point as they represent the different implementation of the same equation. The PT method shows a slightly slower temperature rise as the temperature value is not updated in the iteration process when calculating the transistor currents.

Fig. 5 (c) provides a more detail breakdown on the simulation time spent on numerical iteration and solving the Jacobian Matrix. It shows that all methods have comparable number of iterations in response to a switching signal, independent from the amount of heating. However, the matrix solving time is the longest for the sub-circuit method due to the addition of the internal node. Fig. 5 (d) shows that when the time step increases over the thermal time constant (6μs) in the same transient simulations the convergence of the DE method is reserved.

Transient characteristics of the amplifier are shown in Fig. 6 (a) and (b). When the input frequency is low, i.e., lower than $1/R_{th}C_{th}$, all three methods give similar prediction on the output under the influence of self-heating in Fig. 6 (a). The PT method shows a small deviation from the other two methods indicating some compromise in the accuracy due to the temperature difference as shown in Fig. 5 (b). When the input frequency is higher, the self-heating effect on the amplifier performance is less significant as shown in Fig. 6 (b). Again the DE method reproduces the SC method accurately, but the PT method shows its deficiency. Similarly the DE method reproduces the SC method under different integrations like the trapezoidal one.

The simulated frequency response of the amplifier is plotted in Fig. 7 (a) and (b). Consistent with Fig. 6, self-heating mainly affects the amplifier characteristics in the

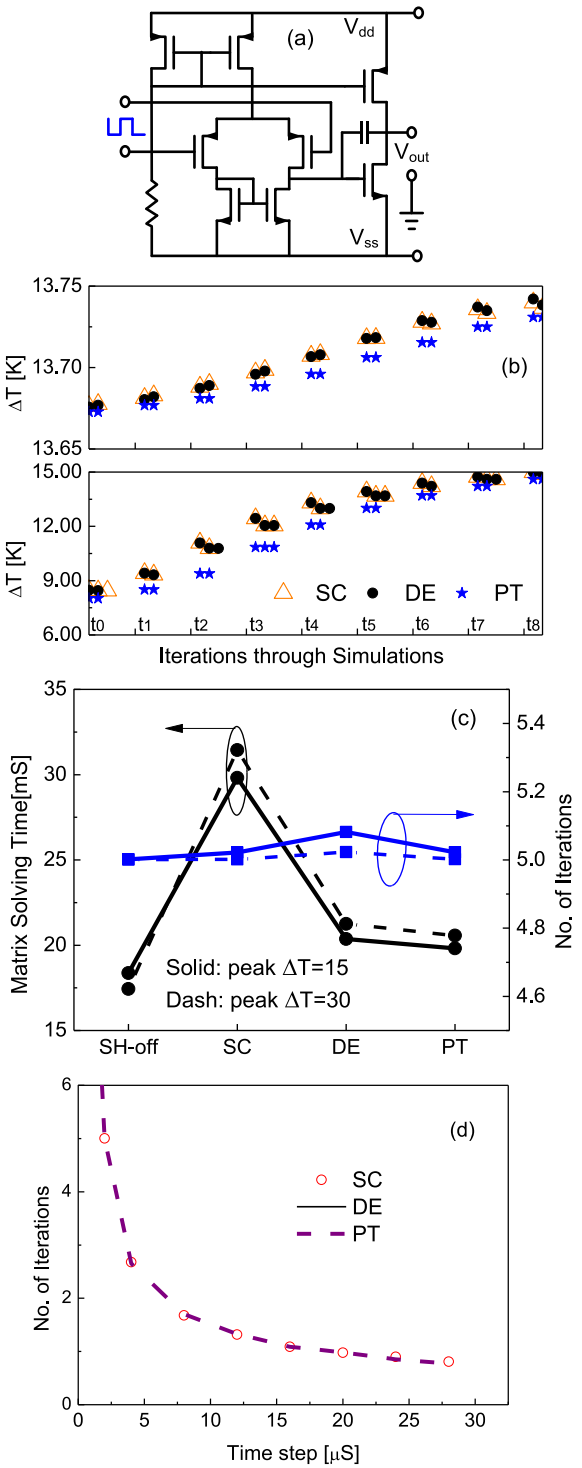

FIGURE 5. **(a) Schematic of an operational amplifier (OP) subject to self-heating and (b) typical iteration process of a device temperature in the transient simulations with slow and fast temperature changing, and (c) a comparison of the circuit solving time and number of iterations with different amount of heating, and (d) number of iterations with different time steps.**

low frequency range. The DE method reproduces the SC method to give both the magnitude and phase shift in a nice agreement.

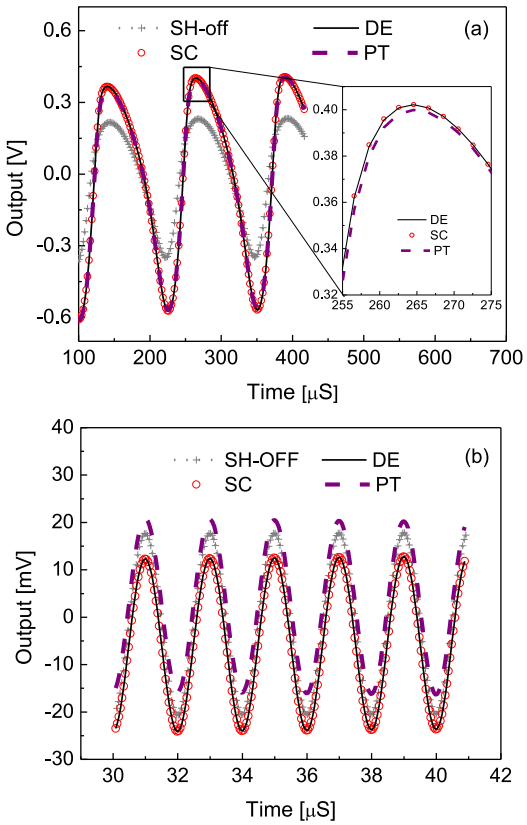

FIGURE 6. Output waveforms of the OP circuit in transient simulations without and with self-heating by different approaches. (a) The input signal frequency is lower than $1/R_{th}C_{th}$ and (b) the frequency is higher. The agreement is observed with the direct evaluation method and the sub-circuit method.

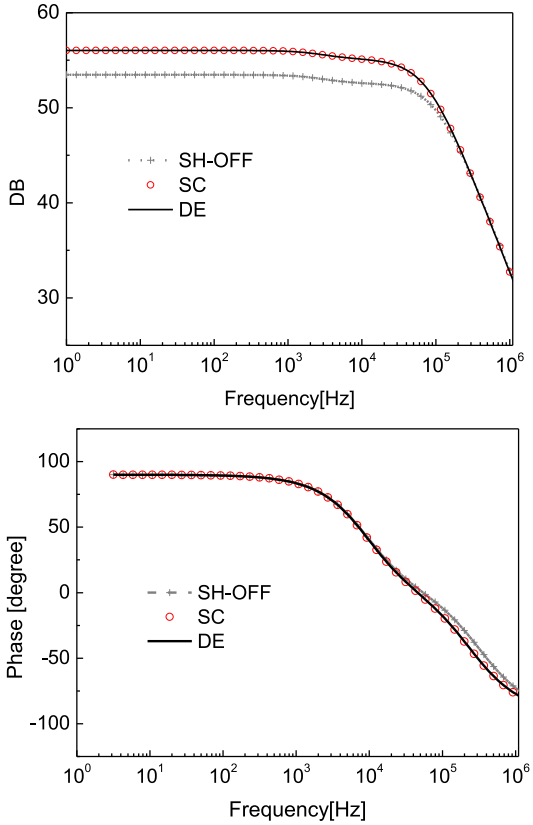

FIGURE 7. Frequency responses of the simulated amplifier without and with self-heating. (a) The magnitude characteristics and (b) the phase shift. The DE method reproduces the SC method in the small signal ac simulations.

The above simulation results verifies that the DE approach can be used to implement self-heating models to replace the sub-circuit method in the same accuracy but with significantly shorter simulation time. Application of the DE approach with Verilog-A [11] requires some changes of the compiler for the circuit matrix operations, e.g., temperature calculations and the right hand side filling. The DE method is more suitable for isolated self-heating, however, has limitations in modeling the thermal coupling [12] between devices.

IV. EXTENSIONS TO DE APPROACH

The DE approach essentially takes out the nodal equations of isolated sub-circuits from the Jacobian matrix and evaluate the node voltage directly with an explicit equation inside the model. This approach is not limited to self-heating model, but other loosely-coupled sub-circuit models as well. For example, DE approach can be used to implement the non-quasi-static (NQS) MOSFET models used in the BSIM models [13].

The BSIM NQS model is constructed with a state variable Q_{def} to keep track of the deficit (or surplus) charge given by

$$Q'_{def} = I_{cheq} - Q_{def}/\tau_{nqs} \tag{8}$$

The time dependent expression of Q_{def} at time t_i is given by:

$$\left(\frac{1}{h} + \frac{1}{\tau_{nqs}}\right) Q_{def}(t_i) = I_{cheq} + \frac{Q_{def}(t_{i-1})}{h} \tag{9}$$

As Q_{def} enters the branch current which then affects solution of the nodal voltages V_i thus also Q_{def} itself, Eq. (9) is also implicit. The iterative solution can be implemented using the following iterative form:

$$\left(\frac{1}{h} + \frac{1}{\tau_{nqs}}\right) Q^n_{def}(t_i) = I_{cheq}\left[V^n_i\right] + \frac{Q_{def}(t_{i-1})}{h} \tag{10a}$$

$$\left(I_{cheq}, Q^n_{def}\right) \Rightarrow V^{n+1}_i \tag{10b}$$

Different from self-heating, Q_{def} after its direct evaluation will enter the matrix calculation. In this way the DE method embeds the charge iteration solution in the voltage NR process. NQS simulations with DE method slow down the convergences compared with the QS simulations, but maintain similar convergent speed compared with SC method.

For verification, we again use the DE and PT approach to compare with the results from the sub-circuit model on the benchmark circuit as shown in Fig. 8 (a) [13]. By turning on the NQS effect using a sub-circuit, the circuit matrix size increases from n=14 to n=22. Fig. 8 (b) shows that

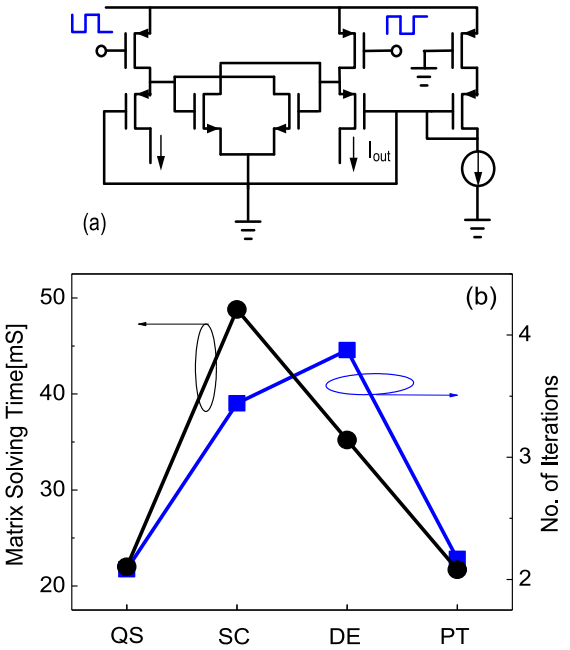

FIGURE 8. **(a) Schematic of an example circuit in which I_{out} is subject to NQS effects at edges of the input signal and (b) circuit solving time and total number of NR iterations are compared. The direct evaluation method with Eq. (7) causes an increase of iterations due to the straightforward inclusion of Q_{def} in current, however, still reduces the matrix solving time significantly.**

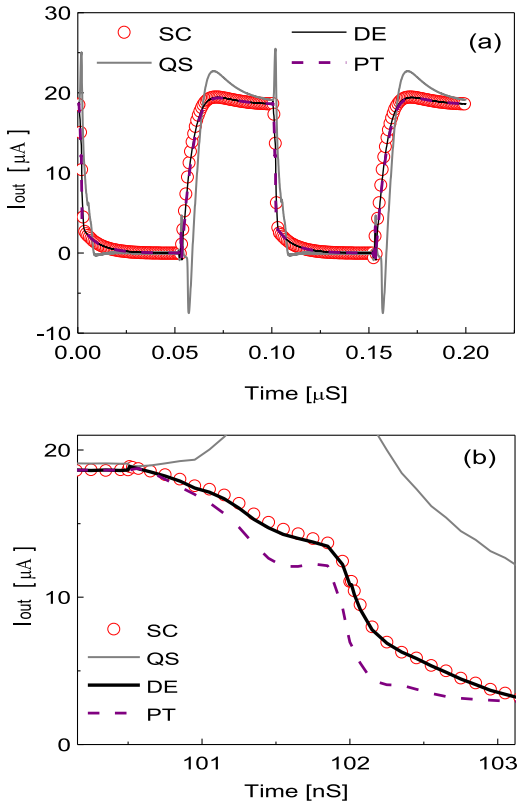

FIGURE 9. **(a) A comparison of circuit outputs with quasi-static simulations and with NQS simulations by different approaches where a generally high accuracy is observed. (b) The direct evaluation of Eq. (7) with $Q_{def}(t_i)$ matches with the sub-circuit method better than the PT method.**

the circuit matrix solving time for a transient simulation increases following due to more iterations and longer solving time of each iteration. It shows that both PT and DE can effectively reduce the simulation time. However, due to the very non-linear nature of the circuit, the sub-circuit and DE methods that perform iterative update of the state variable require more iteration to converge and requires longer time than the quasi-static simulation or the PT method. Comparing the DE and sub-circuit approach, the DE on average requires one more iteration to converge, but it is still within the uncertainty margin. Overall, the DE method still provide substantial time saving compared to the sub-circuit method even though it is slower than the PT method due to more iteration needed.

Fig. 9 (a) shows the simulation results of the current output. It is observed that all the 3 approaches to implement the NQS effect can capture the NQS behavior compared with the QS simulation. However, more detail comparison with enlarged scale shows that the PT have a small deviation from the sub-circuit approach while the DE approach show an exact match with the sub-circuit approach. It is expected that because they represented the same mathematical equations.

V. CONCLUSION

We have studied different approaches to implement the self-heating model in a MOSFET including the conventional sub-circuit approach, the previous-time point approach and the direction evaluation approach. It shows that there is a tradeoff in speed and accuracy between the PT and sub-circuit approach. However, by taking out the loosely-coupled state variable represented by the sub-circuit voltage node from the Jacobian matrix and explicitly evaluating it inside a model, the same results can be achieved but with a much significantly reduced simulation time. The approach can be extended to any loosely-coupled sub-circuit model to speed up simulation time without sacrificing the accuracy.

REFERENCES

[1] L. T. Su, J. E. Chung, D. A. Antoniadis, K. E. Goodson, and M. I. Flik, "Measurement and modeling of self-heating in SOI NMOSFET's," *IEEE Trans. Electron Devices*, vol. 41, no. 1, pp. 69–75, Jan. 1994.

[2] A. J. Scholten *et al.*, "Experimental assessment of self-heating in SOI FinFETs," in *Proc. IEDM*, Baltimore, MD, USA, 2009, pp. 1–4.

[3] W. Yang *et al.*, "BSIMSOIv4.4 MOSFET model users' manual," EECS Dept., Univ. California, at Berkeley, Berkeley, CA, USA, Dec. 2010.

[4] V. Sriramkumar *et al.*, "BSIM-CMG 106.0.0 multi-gate MOSFET compact model technical manual," Dept. Elect. Eng. Comput. Sci., Univ. California, at Berkeley, Berkeley, CA, USA, Rep., Dec. 2012.

[5] Y. Oritsuki *et al.*, "HiSIM-HV: A compact model for simulation of high-voltage MOSFET circuits," *IEEE Trans. Electron Devices*, vol. 57, no. 10, pp. 2671–2678, Oct. 2010.

[6] G. Niu, R. van der Toorn, J. C. J. Paasschens, and W. J. Kloosterman, "The mextram bipolar transistor model version 505.00," Jun. 2017.

[7] T. L. Quarles, "SPICE3 implementation guide," EECS Dept., Univ. California, at Berkeley, Berkeley, CA, USA, Rep. UCB/ERL M89/42, Apr. 1989.

[8] R. S. Vogelsong and C. Brzezinski, "Extending SPICE for electro-thermal simulation," in *Proc. CICC*, San Diego, CA, USA, 1989, pp. 21.4.1–21.4.4.
[9] M. Fulde, D. Schmitt-Landsiedel, and G. Knoblinger, "Transient variations in emerging SOI technologies: Modeling and impact on analog/mixed-signal circuits," in *Proc. ISCAS*, New Orleans, LA, USA, 2007, pp. 1249–1252.
[10] J. C. J. Paasschens, S. Harmsma, and R. van der Toorn, "Dependence of thermal resistance on ambient and actual temperature," in *Proc. IEEE BCTM*, Montreal, QC, Canada, 2004, pp. 96–99.
[11] C. C. McAndrew *et al.*, "Best practices for compact modeling in verilog-a," *IEEE J. Electron Devices Soc.*, vol. 3, no. 5, pp. 383–396, Sep. 2015.
[12] R. van der Toorn, J. C. Paasschens, W. J. Kkoosterman, and H. C. de Graaff, "Mextram," in *Compact Modeling—Principles, Techniques and Applications*, G. Gildenblat, Ed. Amsterdam, The Netherlands: Springer, 2010 ch. 7.
[13] M. Chan, K. Y. Hui, C. Hu, and P. K. Ko, "A robust and physical BSIM3 non-quasi-static transient and AC small-signal model for circuit simulation," *IEEE Trans. Electron Devices*, vol. 45, no. 4, pp. 834–841, Apr. 1998.

LINING ZHANG (M'13) received the Ph.D. degree in electronic and computer engineering from the Hong Kong University of Science and Technology. He is currently an Associate Professor with the College of Electronic Science and Technology, Shenzhen University. His current research interests include semiconductor device physics and modeling and emerging semiconductor devices.

DEBIN SONG received the B.S. degree in microelectronics from the Changchun University of Science and Technology, Changchun, China, in 2015. He is currently pursuing the M.S. degree in microelectronics and solid state electronics from Peking University Shenzhen Graduate School, Shenzhen, China.

YING XIAO is currently pursuing the Ph.D. degree with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong. He has interests in nanoscale devices fabrications, characterizations, and modeling.

XINNAN LIN, photograph and biography not available at the time of publication.

MANSUN CHAN (M'95–SM'01–F'13) received the Ph.D. degree from the University of California at Berkeley in 1995. Since 1996, he has been with the ECE Faculty, Hong Kong University of Science and Technology, Hong Kong.