國立臺灣大學電機資訊學院電子工程學研究所

博士論文

Graduate Institute of Electronics Engineering
College of Electrical Engineering & Computer Science
National Taiwan University

Doctoral dissertation

三維電晶體之自發熱效應及

三維積體電路穿矽連通柱之應變/應力

Self-heating Effects in 3D Transistors and Strain/Stress of 3D-IC TSV

顏智洋
Jhih-Yang Yan

指導教授：劉致為 博士

Advisor: Chee Wee Liu, Ph.D.

中華民國 107 年 7 月

July, 2018

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

國立臺灣大學博士學位論文
口試委員會審定書

三維電晶體之自發熱效應及
三維積體電路穿矽連通柱之應變/應力
Self-heating Effects in 3D Transistors and Strain/Stress of 3D-IC TSV

本論文係顏智洋君（F01943063）在國立臺灣大學電子工程學研究所完成之博士學位論文，於民國 107 年 7 月 21 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

（指導教授）

系主任、所長

doi:10.6342/NTU201801788

Self-heating Effects in 3D Transistors and Strain/Stress of 3D-IC TSV

By

Jhih-Yang Yan

Dissertation

Submitted in partial fulfillment of the requirement
for the degree of Doctor of Philosophy
in Electronics Engineering
at National Taiwan University
Taipei, Taiwan, R.O.C.

July 2018

Approved by :

Chu-Hsuan Lin

Shu-Tong Chang Ming-Han Liao

Min-Hung Lee Min-Jang Chen Kuo-Hsing Kao

Advised by :

Approved by Director :

doi:10.6342/NTU201801788

Related Publication (相關論文發表)

A: Journal Paper（學術期刊論文）

1. **Jhih-Yang Yan**, Sun-Rong Jan, Yi-Chung Huang, Huang-Siang Lan, Y.-H. Huang, Bigchoug Hung, K.-T. Chan, Michael Huang, M.-T. Yang and C. W. Liu, "Asymmetric Keep-out Zone of Through-Silicon Via using 28nm Technology Node", *IEEE Electron Device Letter*, Vol. 36, No. 9, pp. 938-940, 2015.
2. S.-T. Fan, **J.-Y. Yan**, D.-C. Lai, C. W. Liu, "The hysteresis-free negative capacitance field effect transistors using non-linear poly capacitance," *Solid-State Electronics*, Volume 122, Pages 13-17, 2016.
3. I-Hsieh Wong, Yen-Ting Chen, **Jhih-Yang Yan**, Huang-Jhih Ciou, Yu-Sheng Chen and C. W. Liu, "Fabrication and Low Temperature Characterization of Ge (110) and (100) p-MOSFETs" IEEE Transactions on Electron Devices, Vol. 61, No. 6, pp. 2215, 2014.

B: Conference Paper (學術會議論文)

1. **Jhih-Yang Yan**, Chia-Che Chung, Sun-Rong Jan, H. H. Lin, W. K. Wan, M.-T. Yang, and C. W. Liu, "Comprehensive Thermal SPICE Modeling of FinFETs and BEOL with Layout Flexibility Considering Frequency Dependent Thermal Time Constant, 3D Heat Flows, Boundary/Alloy Scattering, and Interfacial Thermal Resistance with Circuit Level Reliability Evaluation," accepted by ***Symposium on VLSI Technology (VLSI-Technology), Honolulu, Hawaii, June 18-22, 2018.***
2. **Jhih-Yang Yan**, Sun-Rong Jan, Yu-Jiun Peng, H. H. Lin, W. K. Wan, Y.-H. Huang, Bigchoug Hung, K.-T. Chan, Michael Huang, M.-T. Yang, and C. W. Liu, "Thermal Resistance Modeling of Back-end Interconnect and Intrinsic FinFETs, and Transient Simulation of Inverters with Capacitive Loading Effects," *p.898-*

doi:10.6342/NTU201801788

*901, **International Electron Devices Meeting (IEDM), 2016**.*

3. I-Hsieh Wong, Fang-Liang Lu, Shih-Hsien Huang, Hung-Yu Ye, Chun-Ti Lu, **Jhih-Yang Yan**, Yu-Cheng Shen, Yu-Jiun Peng, Huang-Siang Lan, and C. W. Liu, "High Performance Ge Junctionless Gate-all-around NFETs with Simultaneous I_{on} =1235 μA/μm at $V_{OV}=V_{DS}$=1V, SS=95 mV/dec, high I_{on}/I_{off}=2E6, and Reduced Noise Power Density using S/D Dopant Recovery by Selective Laser Annealing," *p.842-845, **International Electron Devices Meeting (IEDM), 2016**.*
4. C. W. Liu, **Jhih-Yang Yan**, and Sun-Rong Jan, "Modeling and Simulation of TSV Induced Keep-out Zone Using Silicon Data," 13th International Conference on Solid-State Integrated Circuit & Technology (ICSICT 2016), Hangzhou, China, Oct. 25-28, 2016.
5. **Jhih-Yang Yan**, Sun-Rong Jan, Yi-Chung Huang, Huang-Siang Lan, C. W. Liu, Y.-H. Huang, Bigchoug Hung, K.-T. Chan, Michael Huang, and M.-T. Yang, "Compact Modeling and Simulation of TSV with Experimental Verification," *International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA)*, Hsinchu, Taiwan, 2016.
6. **Jhih-Yang Yan**, Pin-Shiang Chen, Jiun-Ian Pai, Wen-Wei Hsu and C. W. Liu, "The Incorporation of Electromagnetic Effects on Through Silicon Vias in TCAD Simulation" International Semiconductor Device Research Symposium, 2013.
7. I-Hsieh Wong, Yen-Ting Chen, Huang-Jhih Ciou, Yu-Sheng Chen, **Jhih-Yang Yan** and C.W. Liu, "Mobility Strain Response and Low Temperature Characterization of Ge p-MOSFETs," 71st Annual Device Research Conference.

doi:10.6342/NTU201801788

C: Patents (專利)

1. **Jhih-Yang Yan**, C. W. Liu, and Der-Chuan Lai, "Semiconductor device and transistor," (US 9,679,893 B2).
2. **Jhih-Yang Yan**, Samuel C. Pan, C. W. Liu, Hung-Yu Yeh, and Da-zhi Zhang, "Three-Dimensional Transistor and Methods of Manufacturing Thereof," (US 9,627,411 B2).
3. **Jhih-Yang Yan**, Fang-Liang Lu, and Chee Wee Liu, "Highly thermally conductive via structure in MEOL and BEOL for integrated circuits, and intradevice highly thermally conductive trench structure for multi-channel devices," (application in progress).
4. **Jhih-Yang Yan**, Yu-Jiun Peng, H. H. Lin, and Chee Wee Liu, "Thermal via0s to cool down multi-channel devices," (application in progress).

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

誌謝

於台大電子所求學的六年轉瞬已過，從碩士班一年級帶著衝勁卻不知前路有何物，到如今取得博士學位在相關領域嘗試貢獻己力，是得眾位師長、同儕、朋友、家人及合作單位的協助與鼓勵。感謝指導教授 劉致為老師在研究所期間的指導、包容及身教，舉凡：引導研究方向、指正思考邏輯、精確論文著述及訓練報告口條等皆使我裨益匪淺；您更多次於國外研討會中鞭策我在每日會後深入閱讀文獻、隔天與會發問及踴躍於報告結束後與國外學者討論交流研究成果，讓我的研究能與國外團隊之成果比肩。感謝口試委員：張廖貴術教授、林吉聰教授、李敏鴻教授、陳敏璋教授、張書通教授、廖洺漢教授、林楚軒教授以及高國興教授撥冗批閱論文初稿，並在口試時提供建議及方向。

感謝指導過我的學長們：孫宏彰、許文瑋，在我碩一時於文獻蒐集、投影片製作及與合作公司彙報進度等方面給予幫助，使我能步上軌道迅速適應研究室生活。感謝在電二 514 那兩年多的時光，陳彥廷學長及翁翊軒學長讓我視野寬廣，於最前沿半導體元件及材料有更深刻的認識，並在生活中多了許多美好的回憶。感謝研究主題合作的同學及學弟：白君彥、黃奕中、賴德全及彭裕鈞，對於研究主題相關模擬上的幫助讓研究進度能順利推進；特別感謝彭裕鈞學弟，在我想法、直覺最豐富的兩年內，於實作上分擔許多壓力，我們才能將諸多假說及發想一一驗證並發表。感謝詹孫戎學長在應力應變模擬軟體上的教導，讓我免去許多不必要的嘗試。

感謝貫串博班研究生涯合作計畫的聯發科 PTD 部門的各位長官，在每個月的月報中，您們總針對我的研究進度以業界觀點提出問題、給予討論及必要的回饋支持。謝謝科技部與台積電-台大聯合研發中心的支持，使我能在博士學位就讀期間專心致志，毋須為生活四處奔波。

最後要感謝我的父母對於我無條件的支持與愛護，我的求學期間與您們周遭所知、所遇的年輕人相比是如此的漫長，辛苦您們耐心的等待及一路陪伴。

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

摘要

本篇論文之目的為對三維電晶體之自發熱效應及三維積體電路穿矽連通柱造成之截止區提出模型並驗證。三維電晶體相比平面電晶體能提供更好的電性表現，然而單位面積功率的提高、材料的低熱導率皆使元件操作溫度升高，而元件溫度升高將導致元件表現衰退及可靠度問題，因此需要有準確的模型及驗證方法，才能監控三維電晶體中的自發熱行為；三維積體電路比起平面系統單晶片有許多優點：更大的頻寬、更短的導線長度、更高的製程選擇靈活度等，然而其所使用之穿矽連通柱於周遭矽晶圓上造成額外的應變將影響元件表現，因此需要有準確的截止區模型作為電路設計的參考。

在自然對流的狀況下，空氣與散熱表面的邊界需考慮一極大的額外熱阻，此熱阻並在晶片面朝上及晶片面朝下時造成元件散熱路徑的不同。為了能準確描述元件的本質自發熱行為，吾人藉由提出之「二階偽等溫面模型」精準的計算後段繞線熱阻，方得以於晶片面朝下及晶片面朝上的條件分別由總熱阻中萃取出鰭式電晶體的本質熱阻。由於鰭式電晶體具備一垂直非對稱結構，在晶片面朝上時，鰭式電晶體有比起晶片面朝下時更大的本質熱阻。藉由電熱模擬軟體，本篇論文指出因矽鍺源/汲極的低熱導係數，在一反向器中，p 型元件比起 n 型元件有更高的操作溫度。反向器中的溫度高低及高溫持續時間可藉由輸出負載電容及元件電流大小來調變。吾人並指出當以交流訊號驅動反向器後，殘存於元件通道中及第一層金屬導線上的溫度比起操作時溫度皆過低，導致量測困難，以至於利用此溫度作為可靠度評估指標可能高估生命期。

本篇論文提出具分佈式熱阻熱容網路之「以積體電路為重點的模擬程式所建立的熱模型」(thermal SPICE)，以解決現有之雙時間常數及單時間常數模型皆無法準確預估交流元件的自發熱效應的問題。在鰭式電晶體中，熱時間常數並非單一數值，而是與輸入頻率成線性相關。邊界散射、合金散射及介面熱阻皆會提高溫度，並皆被本篇論文提出之 thermal SPICE 納入考量。藉由模組化鰭、金屬

doi:10.6342/NTU201801788

導線以及絕緣層，此熱模型具有元件布局及連線的彈性。於美商英特爾(intel)最先進的鰭式電晶體技術文獻指出，其將採用以鈷製成之導線，因鈷導線具有比銅導線長達 5 倍的電致遷移生命期，然而，當使用本論文提出之 thermal SPICE 對於環型震盪器進行模擬可發現，因鈷導線的熱傳導係數較低，以所得之鈷導線溫度進行理論投射，其電致遷移生命期在考慮金屬導線溫度上升的情況下對銅導線的優勢將減少至僅剩 2.44 倍。吾人發現模擬出之溫度可藉由額外擺放於環形震盪器上的第二層連通柱而降低。藉著第二層連通柱達成較低的鈷金屬溫度，可使得鈷導線比起原情況下的銅導線有 5.56 倍的電致遷移生命期。後段繞線及多鰭鰭式電晶體的熱阻可分別藉由加入額外的熱連通柱及額外的第零層連通柱來降低。

後段繞線完成後才製作之穿矽連通柱在附近的元件中施加額外的應變。本篇論文中於十二吋晶圓上量測 28 奈米節點元件因穿矽連通柱造成的導通電流變異。在此條件下，穿矽連通柱的應力場受到後段繞線絕緣層的影響，吾人觀察到非對稱之應力場並提出其模型。有別於先前的文獻，由於徑向應力及切線應力絕對值大小並不相同，穿矽連通柱造成的截止區並不對稱。藉由量測結果及三維有限元素分析模擬，吾人提出並驗證了一非對稱截止區模型。並推測因為內建應力大小的不同，n 型元件與 p 型元件被發現有大小相近的截止區。

關鍵字： 自發熱效應、熱阻模型、鰭式電晶體、後段繞線、以積體電路為重點的模擬程式所建立的熱模型、穿矽連通柱、非對稱截止區

doi:10.6342/NTU201801788

Abstract

The purpose of this dissertation is to model the self-heating effect in 3D transistors and the keep-out zone of 3D-IC TSVs. 3D transistors provide better electrostatic as compared to planar transistors. However, the junction temperature increases in 3D transistors due to the increased power density and low thermal conductivity of materials. The high junction temperature degrades device performance and reliability. Accurate self-heating effect modeling can help to monitor the junction temperature for lifetime prediction. 3D-IC has the larger bandwidth, shorter interconnect length, and is more cost effective than 2D SoC. However, the through-silicon vias (TSVs) induce additional strain in nearby Si substrate. The strain field leads to the performance variation in the transistors neighboring to the TSV. Precise keep-out zone (KOZ) modeling is important for the circuit design.

With the accurate modeling of the thermal resistance of back-end-of-line (BEOL) by two-step pseudo isothermal plane, the intrinsic thermal resistances of FinFETs are extracted with face-down and face-up configurations. The intrinsic thermal resistance is affected by the direction of heat flow, and it is higher for the face-up configuration than the face-down configuration. The free convection of the air leads to a large thermal resistance. Due to the low thermal conductivity of SiGe S/D, the maximum temperature of pFET is found higher than nFET in an inverter. The output capacitive loading and

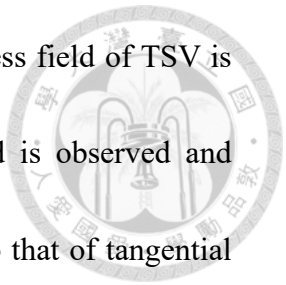
doi:10.6342/NTU201801788

the current of the inverter can control the maximum temperature and the high temperature duration. With AC input, the temperature in M1 layer and the residual temperature in the channel are found too low as compared to the real device temperature. Using the temperature on metal line as the T_j indicator may overestimate the reliability lifetime.

Two τ_c and one τ_c models failed to predict accurate AC self-heating results, and a thermal SPICE modeling with distributed R_{th}-C_{th} network is proposed in this dissertation. In FinFETs, the thermal time constant of the hotspot is linearly dependent to input frequency instead of a constant. Boundary scattering, alloy scattering, and interfacial thermal resistance raise the temperature and are included in the SPICE. The device layout and interconnect routing flexibilities are achieved by using modularized components of fins, metals, and IMDs. The reported intrinsic electromigration improvement of Co interconnect (5X) could be countervailed (5X→2.44X) by the increasing T_{metal} with the projection of Black's theory. T_j (FinFET) and T_{metal} are lowered by placing additional V2 on the power line of a ring oscillator. The predicted EM MTTF of Co interconnect with the lowered T_{metal} by V2 insertion is ~5.65X of W/Cu interconnect. The $R_{th,BEOL}$ and $R_{th0,FinFET}$ can be reduced by adding thermal via in the BEOL and increasing via0s on the multi-fin FinFETs, respectively.

The via-last TSV induced additional strain in nearby devices. The I_{on} variation is

doi:10.6342/NTU201801788



measured using 28nm node devices across 12 inch wafers. The stress field of TSV is affected by the BEOL dielectrics, and an asymmetric stress field is observed and modeled. The absolute value of radial stress ($|\sigma_r|$) does not equal to that of tangential stress ($|\sigma_\theta|$) and leads to the asymmetric KOZ, different from previously reported. With the help of experiment data and 3D finite element analysis (FEA) simulation, a modified KOZ model with the asymmetric radial and tangential stress field is proposed, fitted, and verified. Different internal stress in device channel leads to comparable KOZ size for nFETs and pFETs.

Keywords: Self-heating effect, thermal resistance modeling, FinFET, BEOL, thermal SPICE modeling, through-silicon via, asymmetric keep-out zone.

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

Table of Contents

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

List of Figures

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

List of Tables

doi:10.6342/NTU201801788

doi:10.6342/NTU201801788

Chapter 1 Introduction

1.1 Motivation

The improvements in device **P**erformance, **P**ower, **A**rea, and **C**ost (**PPAC**) were delivered by classic Dennard scaling over the past few decades. However, the Dennard scaling met the limitation at about 15 years ago, and the performance of transistors started to be boosted by transistor innovations (**Fig. 1. 1**) instead of simply shrinking. Strained Si was introduced by SiGe stressor and contact-etching-stop-layer (CESL) for pFETs and nFETs, respectively, at 90nm node. High-k metal gate was proposed to lower the gate leakage by the large physical thickness as compared to the SiO_2 with the same EOT at 45nm node. Since 2012, 3D transistors (Si FinFETs [1.1]) and 3D-ICs (Integrated Fan-out, InFO [1.2]) have become the mainstream of semiconductor industry in commercial products to overcome the short channel effects and to achieve low cost and small form factor, respectively. Nowadays, **R**eliability (**R**) issues are eye-catching in advanced technology nodes, and to meet the requirements of **PPACR** simultaneously has become a must.

Thermal management is getting critical in ultra-scaled 3D transistors due to the increasing device density, the confined device geometry, and the high thermal resistivity of some materials. High temperature in the channel induced by self-heating

doi:10.6342/NTU201801788

effect (SHE) is reported not only to degrade the device performance [1.3]-[1.5] but also to poison the reliability due to accelerated channel hot-carrier degradation [1.6]-[1.8], bias-temperature instability [1.9], and time-dependent dielectric breakdown [1.10]. Therefore, precise device temperature is essential for accurate reliability lifetime projection.

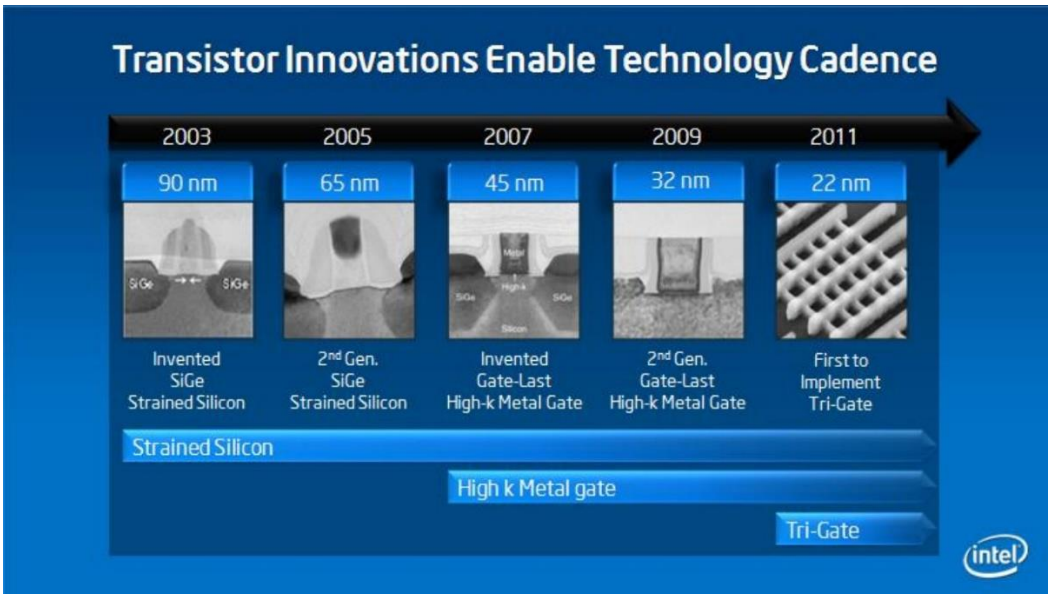


Fig. 1. 1 Transistor performance boosters since 90nm node (Source: Intel Corp.)

While InFO serves as a cost effective option of 3D-IC for the application processor in mobile devices, **Fig. 1. 2** shows the 3D super chip by using the through-silicon vias (TSVs) to connect the vertical stacked chips for heterogeneous integration such as RF, photonics, memories, and logics, etc, which is a promising choice for high-end 3D-IC applications. However, the stress field induced by TSVs due to the mismatch of materials' coefficients of thermal expansion (CTEs) add additional strain in the transistors, which changes the device current. Accurate keep-out zone (KOZ) modeling

doi:10.6342/NTU201801788

is needed for circuit design to avoid the performance variation.

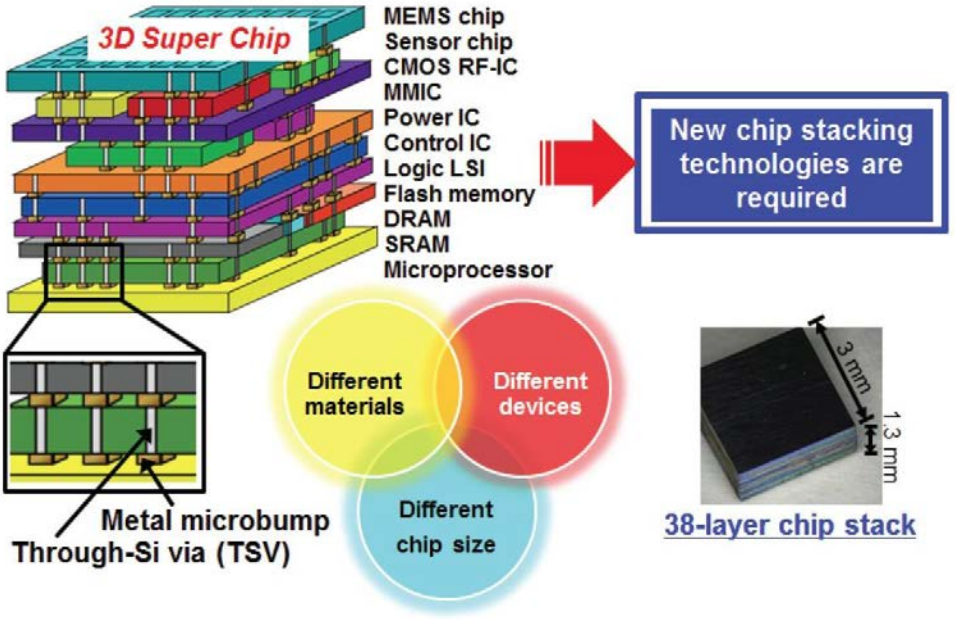


Fig. 1. 2 3D Super chip by TSV [1.11]

In this dissertation, the SHE in scaled FinFETs and the stress field induced by via-last TSVs are simulated and modeled.

1.2 Dissertation Organization

The objections of this dissertation are to provide the temperature mapping both in transistors and the BEOL with the modeling of the SHE in scaled FinFETs both for DC and AC inputs, and to predict the proximity effects of via-last TSVs on nearby devices by the proposed KOZ model with experimental verifications.

In chapter 2, the SHE of FinFETs is simulated using TCAD based on the structure of Intel 14nm FinFETs considering the reduced thermal conductivities of materials suffering from alloy scattering and boundary scattering due to confined geometry.

doi:10.6342/NTU201801788

Thermal resistance (R_{th}) circuit consist of the FinFET, the BEOL, and the substrate is schematically drawn. Non-flat isothermal plane is observed in the BEOL. A two-step pseudo isothermal plane for BEOL R_{th} modeling is proposed and verified by the simulation results. Method of enhancing the heat dissipation is demonstrated. The direction of heat flow is affected by the large thermal resistance of the free convection of air ($R_{th,convection}$). The intrinsic thermal resistances of a single channel ($R_{th0,device}$) in the multi-fin FinFETs are extracted from the overall R_{th}. The $R_{th0,device}$ is dependent on FinFET layout, geometry, and thermal boundary condition.

In chapter 3, the transient junction temperature (T_j) response in an inverter with capacitive loading is simulated by mix-mode electro-thermal TCAD simulation. The cooling time is decided by the volume of hot spot in device channel. The capacitive loading of the inverter in the real circuit is calculated by the output current experimentally. The low thermal conductivity of SiGe S/D leads to the higher T_j in pFETs than in nFETs. The temperature in back-end interconnect and the dependence of device temperature on the output loading are investigated with AC input. The residual T_j and the temperature on M1 are found too low to reflect the real device temperature, which may lead to an underestimation of device temperature with transient AC input during measurements.

In chapter 4, thermal SPICE Modeling of FinFETs and BEOL with distributed R_{th}-

doi:10.6342/NTU201801788

C_{th} network is demonstrated. The thermal time constant of the hotspot ($\tau_{hotspot}$) in scaled FinFETs is found to be frequency dependent instead of being a constant. Guideline for determining the grid size is reported using fitted slope of the frequency dependence of $\tau_{hotspot}$. The discretized fin is modeled with device geometry and material thermal conductivity. The completely modularized fin is connected to the middle-end-of-line (MEOL) and inter-metal dielectric (IMD) to form a FinFET. $R_{th0,device}$ values of the multi-fin FinFETs implemented using our modularized model are verified by the TCAD simulation results to ensure the layout flexibility of the model. Interfacial thermal resistance can be considered in the thermal SPICE model, different from the device simulation.

In chapter 5, the validity of the thermal SPICE model with AC input is ensured by the similar results with device simulation. The $C_{th,BEOL}$ is calculated by the extended two-step pseudo isothermal plane model with identical correction factors for the non-flat isothermal plane in BEOL. $\tau_{hotspot}$, $\tau_{FinFET+MEOL}$, and $\tau_{BEOL+Sub}$ are extracted from the envelope of transient T_j evolution by the time when ΔT_j increases to $1\text{-}e^{-1}$ of the final value. A ring oscillator (RO) with stacked power line and V2 bundles is assembled using our modularized SPICE model. Nodal temperatures at the V0/M1 interface are extracted for BEOL reliability evaluation. The effect of Co incorporation on BEOL EM MTTF considering its low thermal conductivity as compared to Cu and W is

doi:10.6342/NTU201801788

demonstrated. The EM MTTF can be improved by the layout design of BEOL.

In chapter 6, the performance variation caused by the stress field near a through-silicon via (TSV) is measured using 28nm node devices across 12-inch wafers. The TSV is fabricated by a via-last process. The back-end-of-line (BEOL) dielectrics on TSV cause the asymmetric stress field, i.e., the absolute value of radial stress ($|\sigma_r|$) does not equal to that of tangential stress ($|\sigma_\theta|$) on silicon and leads to the asymmetric keep-out zone (KOZ), different from previously reported. A modified KOZ model with the asymmetric radial and tangential stress field is proposed and verified by 3D finite element analysis (FEA) simulation and experiment data. The physics behind the asymmetry is also described. Comparable KOZ size for nFETs and pFETs is observed.

Finally, chapter 7 summarizes the conclusions and the contributions of this work. The further improvements of thermal SPICE model of scaled FinFET and circuit level reliability evaluation are discussed in the future work.

1.3 References

[1.1] C. Auth, C. Allen, A. Blattner, D. Bergstrom, M. Brazier, M. Bost, M. Buehler, V. Chikarmane, T. Ghani, T. Glassman, R. Grover, W. Han, D. Hanken, M. Hattendorf, P. Hentges, R. Heussner, J. Hicks, D. Ingerly, P. Jain, S. Jaloviar, R. James, D. Jones, J. Jopling, S. Joshi, C. Kenyon, H. Liu, R. McFadden, B. McIntyre, J. Neirynck, C. Parker, L. Pipes, I. Post, S. Pradhan, M. Prince, S.

doi:10.6342/NTU201801788

Ramey, T. Reynolds, J. Roesler, J. Sandford, J. Seiple, P. Smith, C. Thomas, D. Towner, T. Troeger, C. Weber, P. Yashar, K. Zawadzki, K. Mistry, "A 22nm high performance and lowpower CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors," in *Proc. Symposium on VLSI Technology*, June 2012, pp. 131-132.

[1.2] Christianto C. Liu, Shuo-Mao Chen, Feng-Wei Kuo, Huan-Neng Chen, En-Hsiang Yeh, Cheng-Chieh Hsieh, Li-Hsien Huang, Ming-Yen Chiu, John Yeh, Tsung-Shu Lin, Tzu-Jin Yeh, Shang-Yun Hou, Jui-Pin Hung, Jing-Cheng Lin, Chewn-Pu Jou, Chuei-Tang Wang, Shin-Puu Jeng, Douglas C.H. Yu, "High-performance integrated fan-out wafer level packaging (InFO-WLP): Technology and system integration," in *IEEE IEDM Tech. Dig.*, Dec. 2012, pp. 14.1.1-14.1.4.

[1.3] Liping Wang, Andrew R. Brown, Mihail Nedjalkov, Craig Alexander, Binjie Cheng, Campbell Millar, and Asen Asenov, “Impact of self-heating on the statistical variability in bulk and SOI FinFETs,” *IEEE Transactions on Electron Devices,* Vol. 62, No. 7, July 2015, pp. 2106-2112.

[1.4] E. Bury, B. Kaczer, P. Roussel, R. Ritzenthaler, K. Raleva, D. Vasileska, and G. Groeseneken, "Experimental validation of self-heating simulations and projections for transistors in deeply scaled nodes," *Reliability Physics*

doi:10.6342/NTU201801788

Symposium (IRPS), *2014 IEEE International*, pp. XT.8.1- XT.8.6.

[1.5] E. Bury, B. Kaczer, J. Mitard, N. Collaert, N.S. Khatami, Z. Aksamija, D. Vasileska, K. Raleva, L. Witters, G. Hellings, D. Linten, G. Groeseneken, and A. Thean, "Characterization of self-heating in high-mobility Ge FinFET pMOS devices," in *Proc. Symposium on VLSI Technology*, June 2015, pp. T60–T61.

[1.6] Hai Jiang, Xiaoyan Liu, Nuo Xu, Yandong He, Gang Du, and Xing Zhang, "Investigation of self-heating effect on hot carrier degradation in multiple-Fin SOI FinFETs," *Electron Device Letters, IEEE* 36.12 (2015): pp.1258-1260.

[1.7] Steven Mittl and Fernando Guarín, " Self-Heating and Its Implications on Hot Carrier Reliability Evaluations," *Reliability Physics Symposium (IRPS)*, *2015 IEEE International,* pp. 4A.4.1-4A.4.4.

[1.8] A. Laurent, X. Garros, S. Barraud, G.Mariniello, G. Reimbold, D. Roy, E.Vincent, and G.Ghibaudo, "Hot Carrier Degradation in Nanowire Transistors: Physical mechanisms, Width dependence and Impact of Self-Heating," in *Proc. Symposium on VLSI Technology*, June 2016, pp. T48–T49.

[1.9] S. E. Liu, J. S. Wang, Y. R. Lu, D. S. Huang, C. F. Huang, W. H. Hsieh, J. H. Lee, Y. S. Tsai, J. R. Shih, Y.-H. Lee, and K. Wu, "Self-heating effect in FinFETs and its impact on device reliability characterization," *Reliability Physics Symposium (IRPS)*, *2014 IEEE International,* pp. 4A.4.1-4A.4.4.

doi:10.6342/NTU201801788

[1.10] S. Mei, N. Raghavan, M. Bosman, D. Linten, G. Groeseneken, N. Horiguchi, and K.L. Pey, " New understanding of dielectric breakdown in advanced FinFET devices – physical, electrical, statistical and multiphysics study," in *IEEE IEDM Tech. Dig.*, Dec. 2016, pp. 15.5.1-15.5.4.

[1.11] Mitsumasa Koyanagi, " Heterogeneous 3D Integration - Technology Enabler toward Future Super-Chip," in *IEEE IEDM Tech. Dig.*, Dec. 2013, pp. 1.2.1-1.2.8.

doi:10.6342/NTU201801788

Chapter 2 Thermal Resistance Modeling of Intrinsic FinFETs and Back-end-of-line

2.1 Introduction

Self-heating (SH) in scaled FinFET circuits retards the performance, degrades device reliability, and accelerates the electro-migration (EM) in back-end-of-line (BEOL). Temperature mapping with high accuracy is essential for precise reliability lifetime prediction.

Several methods have been reported for SHE measuring:

1. Resistance thermometry structure [2.1],[2.2]: the electrical resistance changes of the device metal gate and the M1 metal are measured. However, the extracted temperature is the average temperature of the entire metal line, and the junction temperature (T_j) projection needs to be carefully calibrated by device simulation.

2. Heater-sensor structure [2.2]: changes in the electrical properties of the nearby device are measured. In this case, the T_j of device under test (DUT) is not measured directly. The estimate T_j of DUT could be affected by device layouts and boundary condition. Device simulation is also needed for the T_j projection.

3. Thermoreflectance image [2.3]: it images the top surface temperature of transistors. However, the different boundary condition and heat dissipation path

doi:10.6342/NTU201801788

between conventional measurement and flip-chip packaging will lead to the T_j difference between measurement and real circuit. These two configurations will be discussed in 2.2.2.

4. Scanning Thermal Microscopy [2.4]: the local temperature of the top surface of the sample is measured by a nanoscale sharp tip through the mechanical contact. This method suffers from the similar boundary condition issue as thermoreflectance image does.

Moreover, it has been reported that none of the existing SHE measurement methods can provide precise AC response of spatial and temporal temperatures, which are essential for accurate reliability lifetime projection in circuit level. Therefore, device simulation is the only way to know the exact temperature distribution and time evolution in the channel. Device simulation of the SHE in FinFET was reported [2.3],[2.5], but the difference between measurement and real circuit was not considered, and the intrinsic thermal resistance was not decoupled.

In this chapter, a two-step pseudo isothermal plane model of the thermal resistance of BEOL ($R_{th,BEOL}$) is proposed and verified by numerical simulation. With $R_{th,BEOL}$, the intrinsic thermal resistance of FinFET ($R_{th,device}$) is decoupled from the thermal resistance circuit. The dependence of intrinsic thermal resistance of a single channel in FinFETs ($R_{th0,device}$) on the device geometry and layout are discussed.

doi:10.6342/NTU201801788

2.2 Electro-thermal Simulation and Boundary Conditions

2.2.1 FinFET Structure and Simulation Settings

The simulation structure is composed of the FinFET, S/D contacts, gate metal, via0s, and part of the Si substrate with the thickness of 250nm. (**Fig. 2. 1**). Diamond shape S/D structure and the taper fin are considered according to the cross-sectional TEM images. **Table 2. 1** shows the geometry parameters which follow the reported 14nm FinFET technology by intel [2.6]. The bottom side of FinFETs is connected to the chassis temperature ($T_{chassis}$) by the thermal resistance of Si substrate ($R_{th,sub}$) and the via0s are connected to $T_{chassis}$ through $R_{th,BEOL}$ (**Fig. 2. 1 (a)**). The $T_{chassis}$ is assumed to be 40ºC. The four sidewalls of simulation structure are set adiabatic to represent the boundary condition in the hot spot of a chip (no heat flux between two turned on device).

Top views of FinFETs with 1-fin-1-finger, 2-fin-1-finger, and 1-fin-2-finger are shown in **Fig. 2. 1 (b).** The tungsten M0 of source and drain, TiN metal gate, and copper via0s are implemented in the structure. The extension of metal gate is considered following the design rule.

Note that in multi-fin FinFETs, e.g. the 2-fin-1-finger device, the number of via0s remains three as compared to the 1-fin-1-finger device. On the other hand, the number of via0s increases with the increasing finger number in multi-finger FinFETs. For example, there are five via0s for the 1-fin-2-finger device.

doi:10.6342/NTU201801788

Thermal conductivities of materials used in the simulation are shown in **Table 2. 2**. It is reported [2.7],[2.8] that the thermal conductivity of SiGe S/D and silicon fin are lower than that of bulk Si due to alloy scattering and boundary scattering, respectively.

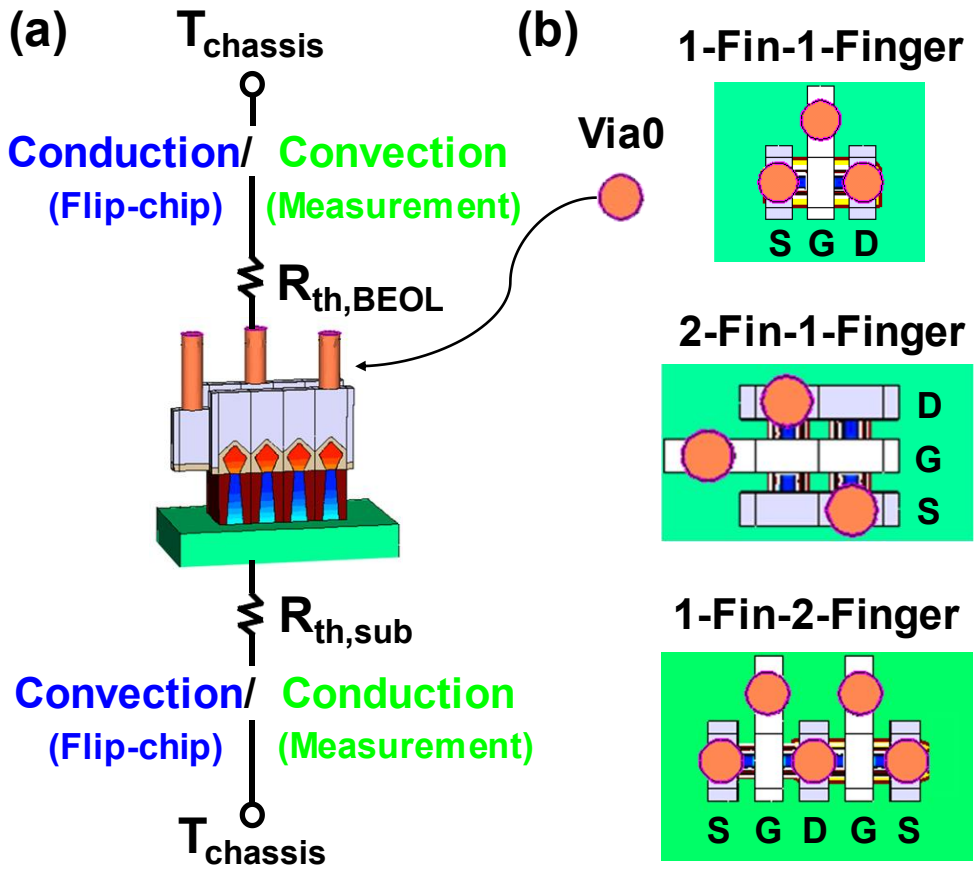


Fig. 2. 1 (a) FinFET structure and thermal resistances of BEOL and substrate, and **(b)** device layouts

doi:10.6342/NTU201801788

Table 2. 1 14nm FinFET Parameters [2.6]

Gate length (L_g)	22 nm	Contact poly pitch	70 nm
Fin height (H_{fin})	40 nm	I_D	42 (μA/fin)
Average fin width ($W_{fin, avg}$)	~10 nm	V_{DD}	0.7 V
Fin pitch (FP)	42nm		

Table 2. 2 Thermal conductivities of materials

Material	κ_{th} ($Wm^{-1}K^{-1}$)	Material	κ_{th} ($WM^{-1}K^{-1}$)
Si (bulk)	148	Tungsten	174
Si (10nm)	**13**	TiN	19.2
SiO_2	1.4	Copper	401
SiGe	**2**		

2.2.2 Conventional Measurement and Flip-chip Packaging

In the common SHE measurements, the samples are placed face-up on the stage and only the devices under test are turned on. On the other hand, the ICs such as the application processors (AP) in mobile phones are usually packaged in the face-down

doi:10.6342/NTU201801788

configuration (flip-chip), and many devices are turned on.

Dissipating the heat through the free convection of the air is way less efficient than the conduction [2.9], and leads to an additional thermal resistance ($R_{th,convection}$) at the boundary. The $R_{th,\ convection}$ is much larger than $R_{th,BEOL}$ and $R_{th,Sub}$. Most of the heat (>99.99%) dissipates by conducting through the Si substrate to the stage for face-up configuration (**Fig. 2. 2 (a)**). For the face-down configuration, most of the heat dissipates by conducting through the BEOL mounted on PCB using μ-bumps (**Fig. 2. 2 (b)**).

With $T_{chassis}$ = 40ºC, the four sidewalls of simulated FinFETs are set to be adiabatic (no heat flux between two on devices). For face-up configuration, the FinFETs are connected to the calculated $R_{th,Sub}$ for the area size of 5μm x 5μm through the substrate side to represent the single turned on device during the measurement (**Fig. 2. 3** (a)). For face-down configuration, the FinFETs are connected to the modeled $R_{th,BEOL}$ with the device layout size through the via0s, just as the hotspot in the circuit of a flip-chip packaged die (**Fig. 2. 3 (b)**).

doi:10.6342/NTU201801788

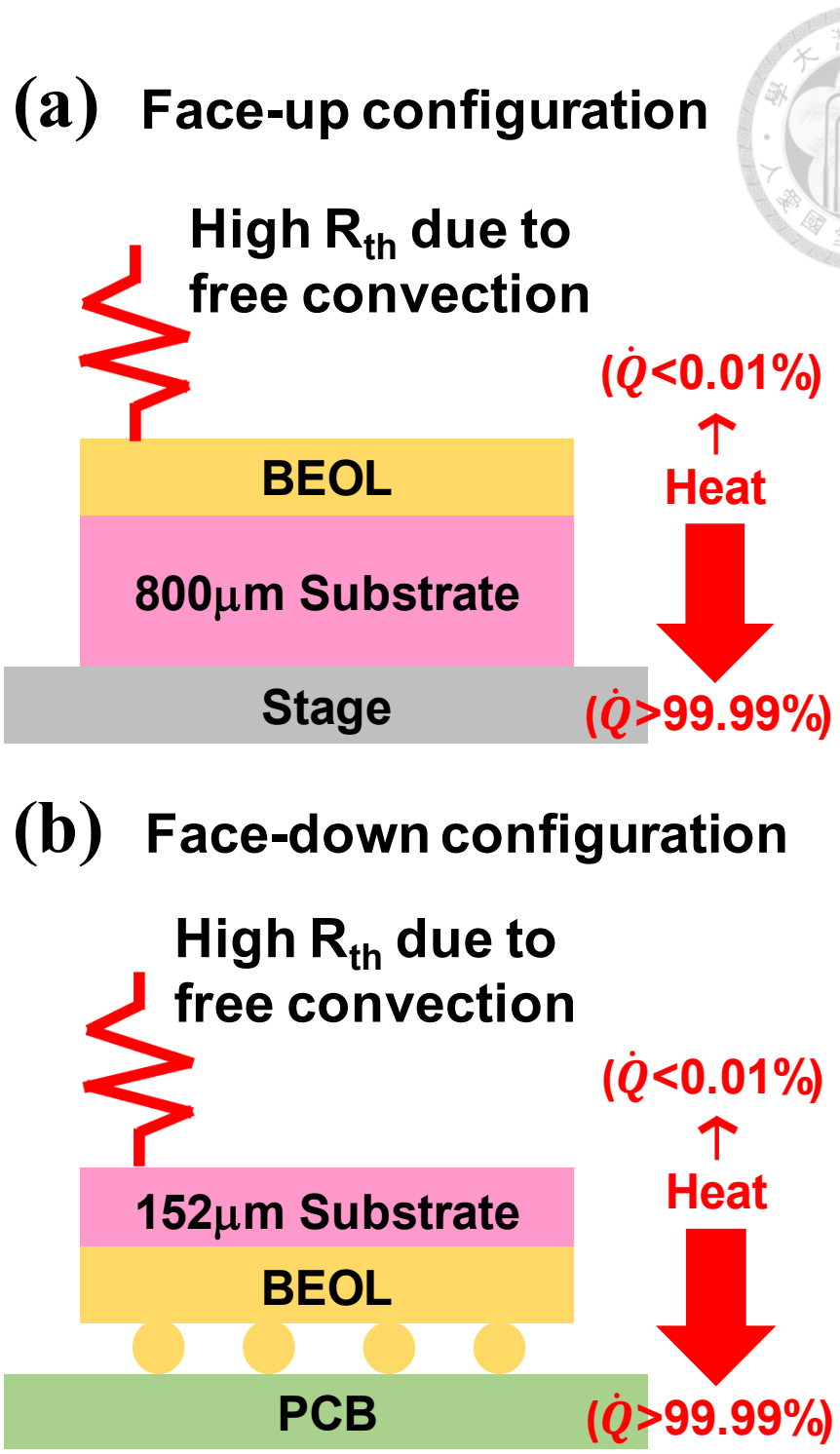


Fig. 2. 2 Heat dissipation paths in **(a)** conventional measurement (face-up) and **(b)** flip-chip packaging (face-down)

doi:10.6342/NTU201801788

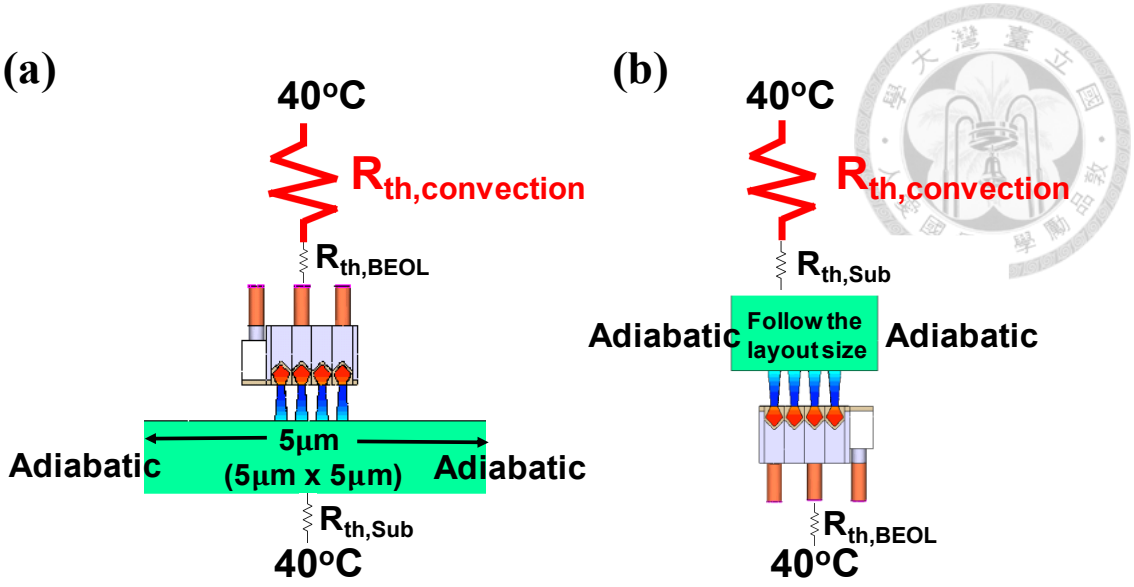


Fig. 2. 3 Simulation structure, boundary condition, and thermal resistances of **(a)** face-up and **(b)** face-down configurations

2.3 Thermal Resistance Modeling of BEOL

2.3.1 Thermal Resistance Circuit

The thermal resistance circuit is schematically drawn for both face-up and face-down configurations (**Fig. 2. 4**). The total thermal resistance ($R_{th,total}$) can be calculated by:

$$R_{th,total} = \frac{T_j - T_{chassis}}{\dot{Q}} \tag{2.1}$$

, where

$$\dot{Q} = I_{on} \times V_{DS} \tag{2.2}$$

is the Joule heat of device.

The corresponding intrinsic thermal resistance of device ($R_{th,device}$) can be

doi:10.6342/NTU201801788

decoupled from the $R_{th,total}$ by:

$$R_{th,device} = R_{th,total} - [(R_{th,BEOL}+R_{th,convection})//R_{th,Sub}] \quad (2.3)$$

, and

$$R_{th,device} = R_{th,total} - [R_{th,BEOL}//(R_{th,Sub}+R_{th,convection})] \quad (2.4)$$

for face-up and face-down configurations, respectively.

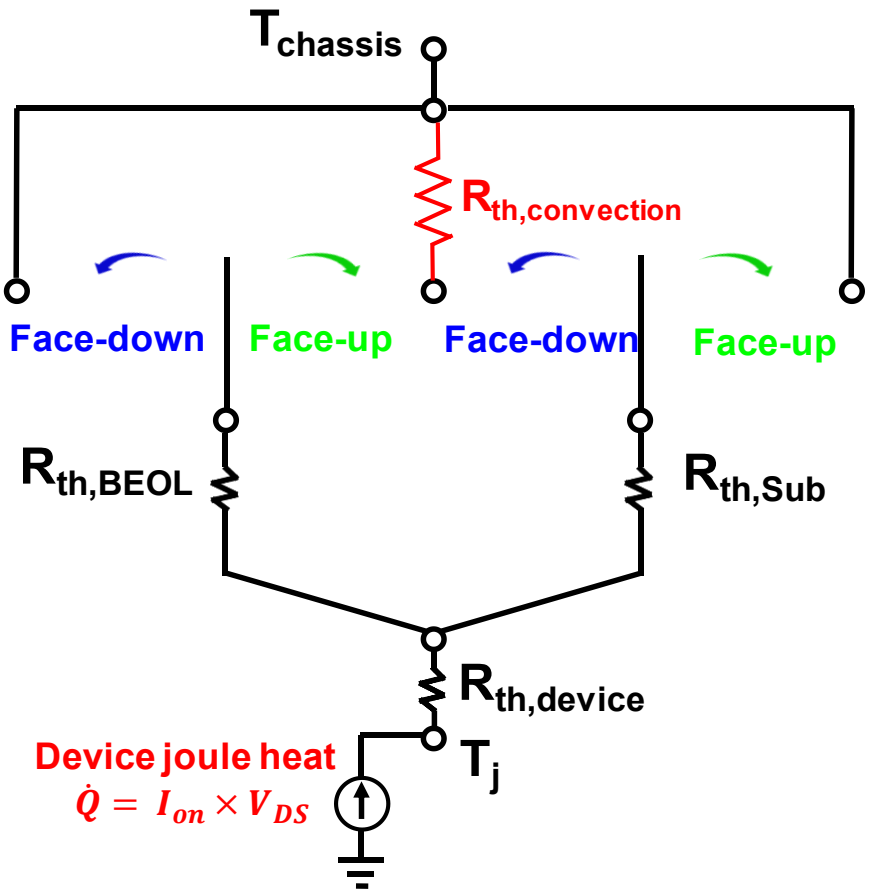


Fig. 2. 4 Thermal resistance circuit of device, BEOL, substrate, and free convection

While $R_{th,Sub}$ can be simply calculated by:

$$R_{th,Sub} = \frac{1}{k_{th,\text{Si (bulk)}}} \cdot \frac{thickness}{area} \quad (2.5)$$

doi:10.6342/NTU201801788

,precise $R_{th,BEOL}$ modeling is needed to extracted the $R_{th,device}$ from the thermal resistance circuit based on equation (2.3) and (2.4).

2.3.2 Two-step Pseudo Isothermal Plane Modeling

Fig. 2. 5 shows the isothermal plane in the simulated BEOL structure. Both of the thicknesses of via layer and metal layer are 100nm. The metal line and via are made of Cu, and the IMD is assumed pure SiO_2. Non-flat isothermal plane in BEOL is observed, different from the physical structure. The spacing between isothermal plane in SiO_2 outside the via is wider than near the via due to the thermal conductivity difference between metal and dielectric.

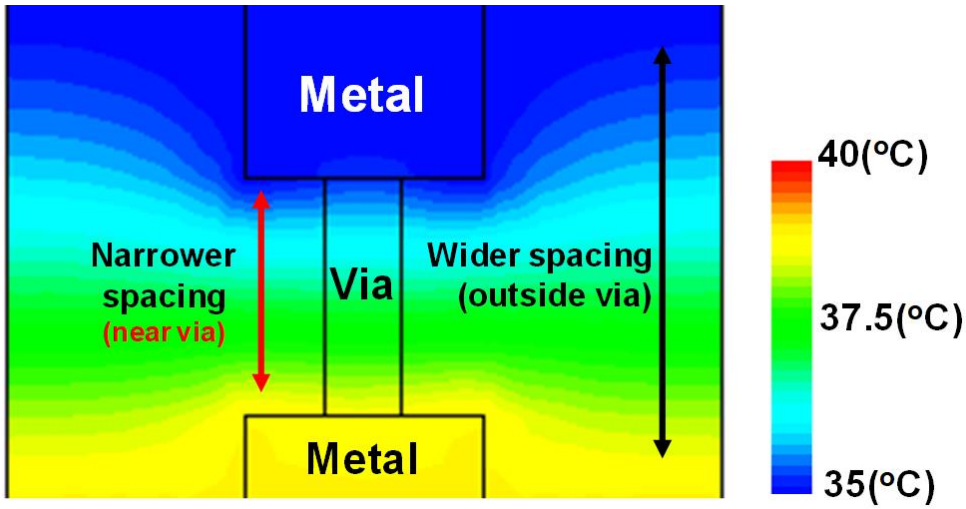


Fig. 2. 5 Simulated temperature contour in BEOL

A two-step pseudo isothermal plane model with the corrected oxide thicknesses x and y for SiO_2 outside the metal layers and SiO_2 between the metal layers, respectively,

doi:10.6342/NTU201801788

is proposed based on the simulation results (**Fig. 2. 6**). With the physical thickness of via layer (h_1) and metal layer (h_2), thermal resistances are calculated layer by layer based on the via densities and metal densities. In via layers, the thermal resistance is:

$$R_{th,via} = R_{cu} // R_{SiO2\_between\ metal\ layers} // R_{SiO2\_outside\ the\ metal\ layers} \quad (2.6)$$

In metal layers, the thermal resistance is:

$$R_{th,metal} = (R_{cu} + R_{SiO2\_between\ metal\ layers}) // R_{SiO2\_outside\ the\ metal\ layers} \quad (2.7)$$

The total $R_{th,BEOL}$ is the series of all $R_{th,via}$ and $R_{th,\ metal}$.

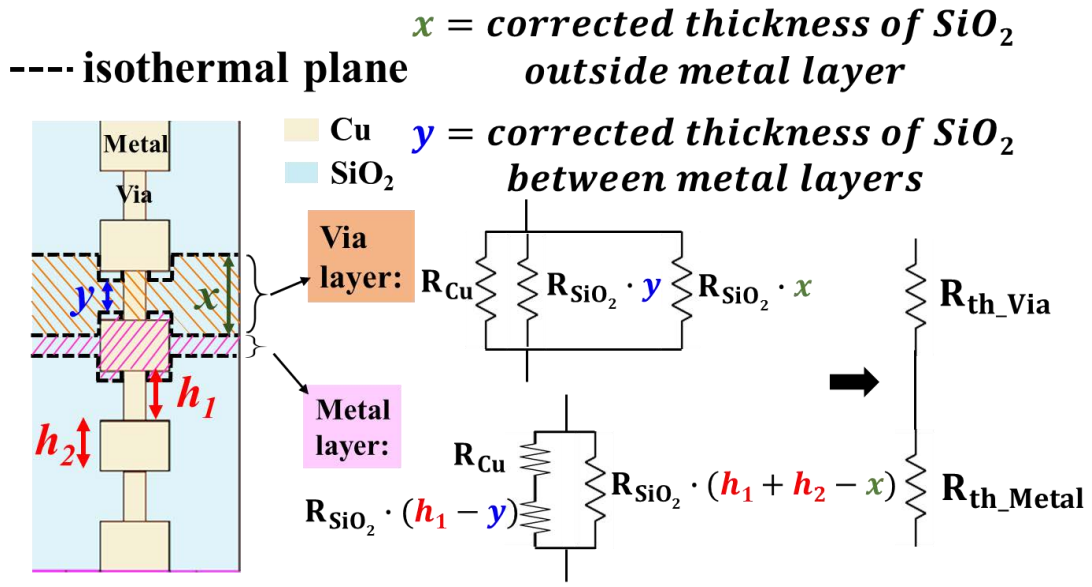


Fig. 2. 6 Two-step pseudo isothermal plane modeling of $R_{th,BEOL}$ with the corrected IMD thickness

To verify the proposed model, BEOL structures (**Fig. 2. 7**) consist of five metal layers with different via densities are simulated. The top contact and the bottom contact

doi:10.6342/NTU201801788

are set to be at 35°C and 40°C, respectively. The thermal resistance is calculated by the temperature difference (5°C) and the simulated heat flow through the contacts. There are some dummy metals in the BEOL in real products, which has also been considered using parts of none connected metals in the cases with low via density (<1%).

The two-step pseudo isothermal model with $h_1=h_2=1\mu m$, $x=0.14\mu m$, and $y=0.098\mu m$ is verified by the simulation results (**Fig. 2. 8**). The model can provide accurate $R_{th,BEOL}$ as compared to the flat isothermal plane model.

With the $R_{th,BEOL}$ model, the specific thermal resistivities (ρ, thermal resistance per unit area) are calculated for both BEOL using listed metal density and Si substrate, and compared to the free convection (**Table 2. 3**). The $\rho_{th,convection}$ is over five orders of magnitude lager than $\rho_{th,BEOL}$ and $\rho_{th,sub}$.

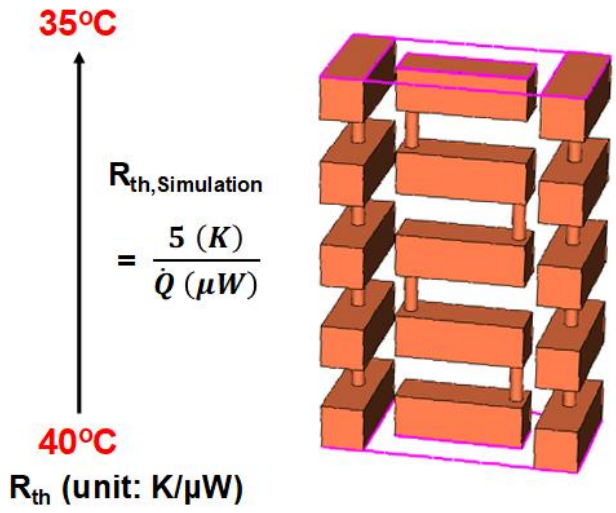


Fig. 2. 7 Simulation structure and settings for $R_{th,BEOL}$ extraction

doi:10.6342/NTU201801788

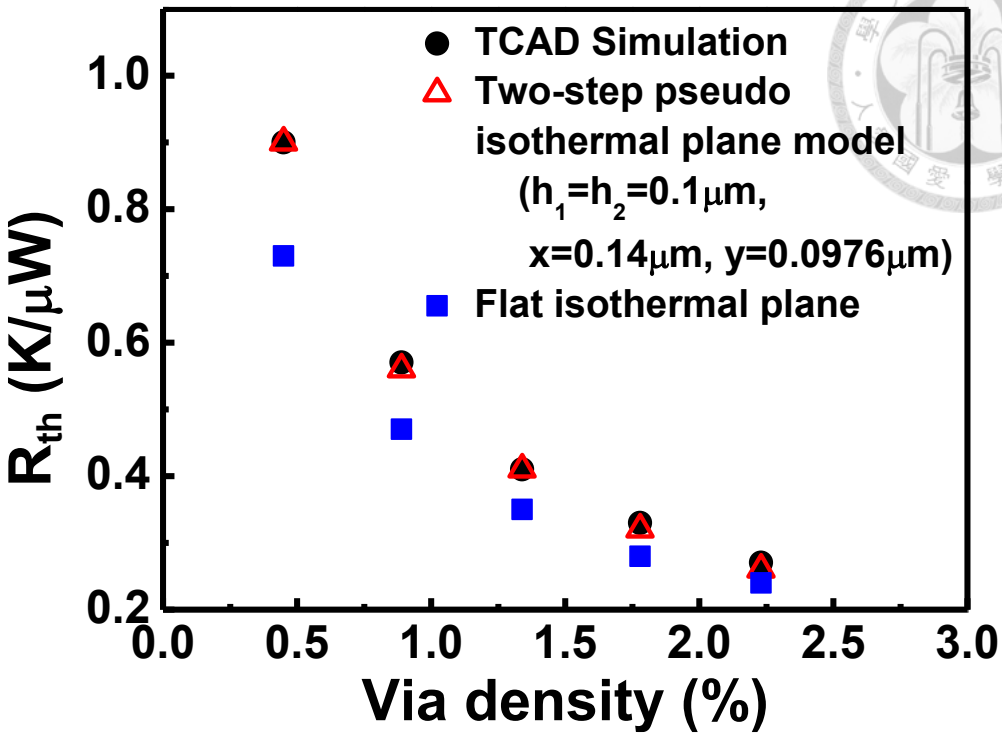


Fig. 2. 8 Comparison of $R_{th,BEOL}$ models with simulation results

Table 2. 3 BEOL metal densities and specific thermal resistivity

BEOL metal densities	
Via layers	5%
M1~M6	60%
M7 and M8	80%
Specific thermal resistivity (K•μm²/W)	
$\rho_{th,BEOL}$	4.8E4
$\rho_{th,sub}$ (152μm)	1.0E6
$\rho_{th,sub}$ (800μm)	5.4E6
$\rho_{th,convection}$	2.0E11

doi:10.6342/NTU201801788

2.4 Thermal Resistance Modeling of FinFET

2.4.1 Intrinsic Thermal Resistance of FinFET

The intrinsic thermal resistance of the entire device $R_{th,device}$ decreases with the increasing device layout due to the enlarged heat dissipation area. To evaluate the SHE in FinFETs with different layout size, the intrinsic thermal resistance of a single channel ($R_{th0,device}$) should be used for the convenience (**Fig. 2. 9**). Given the $R_{th,device}$ value of a multi-fin or multi-finger FinFET, the $R_{th0,device}$ of a single channel is obtained as:

$$R_{th0,device} = R_{th,device} \times (\text{fin \#}) \times (\text{finger \#}) \quad (2.8)$$

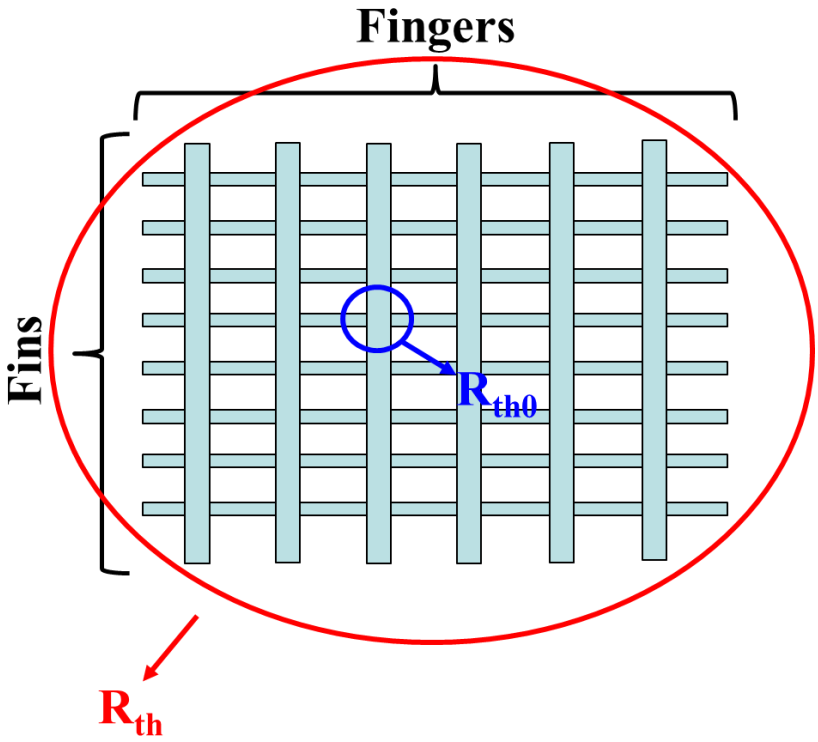


Fig. 2. 9 Thermal resistance of the entire device and a single channel

doi:10.6342/NTU201801788

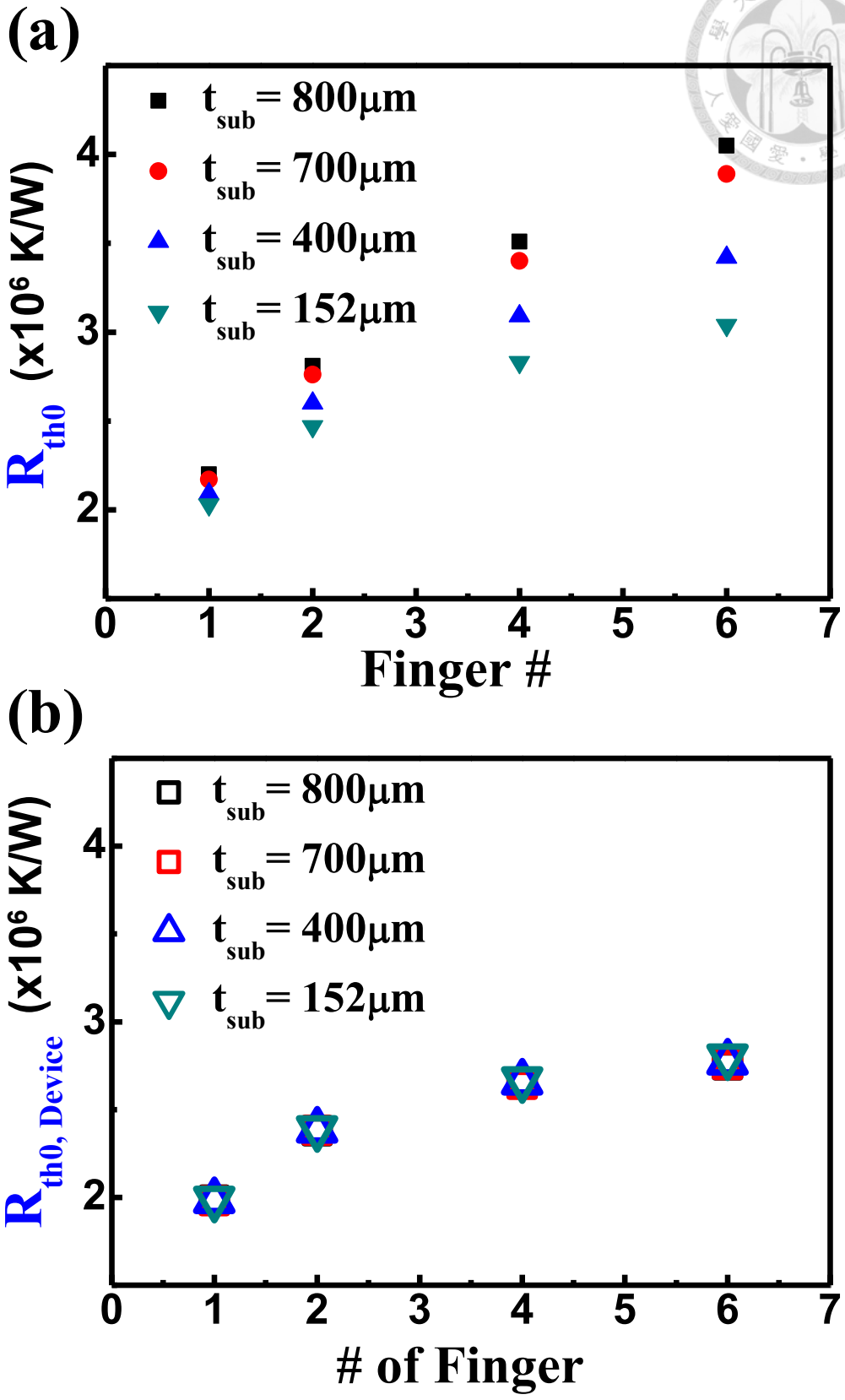


Fig. 2. 10 (a) $R_{th0,total}$ and **(b)** $R_{th0,device}$ with different substrate thickness

doi:10.6342/NTU201801788

It is important to extract the $R_{th0,device}$ from the $R_{th0,total}$. The typical substrate thickness of a Si wafer is about 800 μm. However, to achieve small form factor of the mobile phone, the AP dies are usually thinned down to 6 mils (152μm). The simulated $R_{th0,total}$ under face-up configuration is observed to be highly dependent on the substrate thickness (**Fig. 2. 10 (a)**). Using the measured or simulated $R_{th0,total}$ with the substrate thickness of 800μm could overestimate the device T_j after the thin down process.

The $R_{th0,device}$ is decoupled using equation (2.3), and shows no substrate thickness dependence (**Fig. 2. 10 (b)**). It is the intrinsic characteristic of the device and would not be affected by substrate or BEOL.

2.4.2 Layout and Geometry Dependence of Intrinsic Thermal Resistance

The device T_j can be predicted by the following equation:

$$T_j - T_{Chassis} = R_{th} \times \dot{Q}_{total} = R_{th0} \times \dot{Q}_{single\ channel} \quad (2.9)$$

, where $\dot{Q}_{total}$ is total power of the device and $\dot{Q}_{one\ channel}$ is power of a channel.

To see the layout and geometry dependence of SHE in FinFETs, $R_{th0,device}$ is extracted and fitted for both face-up and face-down configurations.

The $R_{th0,FinFET}$ is observed to be linearly dependent on the fin number (**Fig. 2. 11**) and can be fitted by:

doi:10.6342/NTU201801788

$$R_{th0,FinFET} = 1.98 \times [0.13(fin\ \#) + 0.87]\left(\frac{K}{\mu W}\right) \quad (2.10)$$

, and

$$R_{th0,FinFET} = 1.85 \times [0.13(fin\ \#) + 0.89]\left(\frac{K}{\mu W}\right) \quad (2.11)$$

for face-up and face-down configurations, respectively

On the contrary, the $R_{th0,FinFET}$ saturates as the number of finger increases (**Fig. 2. 12**) and is fitted by:

$$R_{th0,FinFET} = 1.98 \times \left[1.35\left(1 - e^{-1.35\times(finger\ \#)}\right)\right]\left(\frac{K}{\mu W}\right) \quad (2.12)$$

, and

$$R_{th0,FinFET} = 1.85 \times \left[1.08\left(1 - e^{-2.30\times(finger\ \#)}\right)\right]\left(\frac{K}{\mu W}\right) \quad (2.13)$$

for face-up and face-down configurations, respectively.

The difference trends of $R_{th0,FinFET}$ with increasing number of fin and increasing number of finger is due to the number of via0 on the FinFETs. As fin number increases, the number of via0 remains three. For fixed three via0s, the $R_{th0,FinFET}$ increases linearly with the incrasing total area. On the other hand, the number of via0 increases with the increasing finger number. The enhanced heat dissipating ability by additional via0s makes $R_{th0,FinFET}$ to saturate.

The results reveals that adding additional via0s can ease SHE, and will be further discussed in chapter 6.

doi:10.6342/NTU201801788

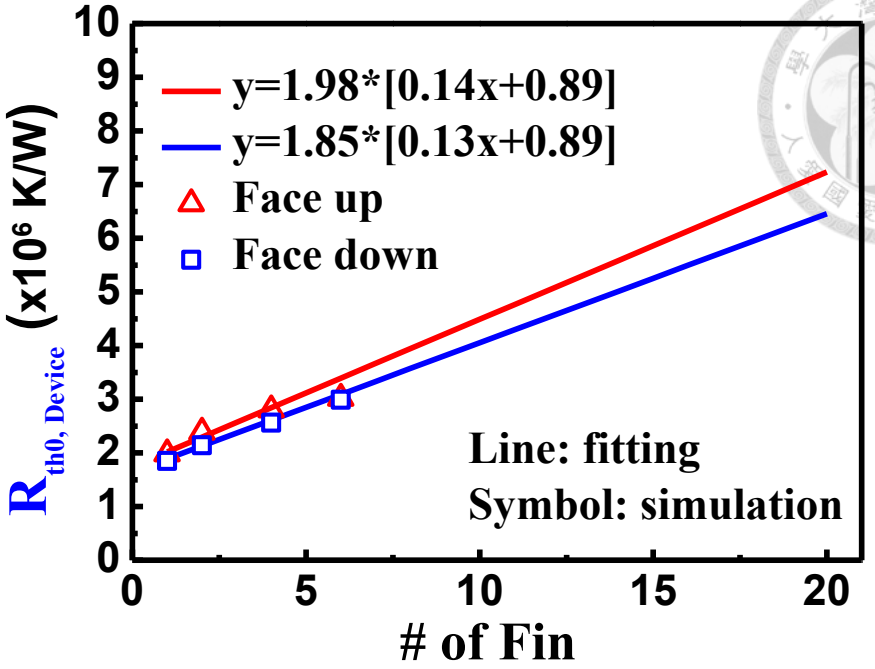


Fig. 2. 11 $R_{th0,device}$ fitting as a function of fin number for face-up and face-down configurations

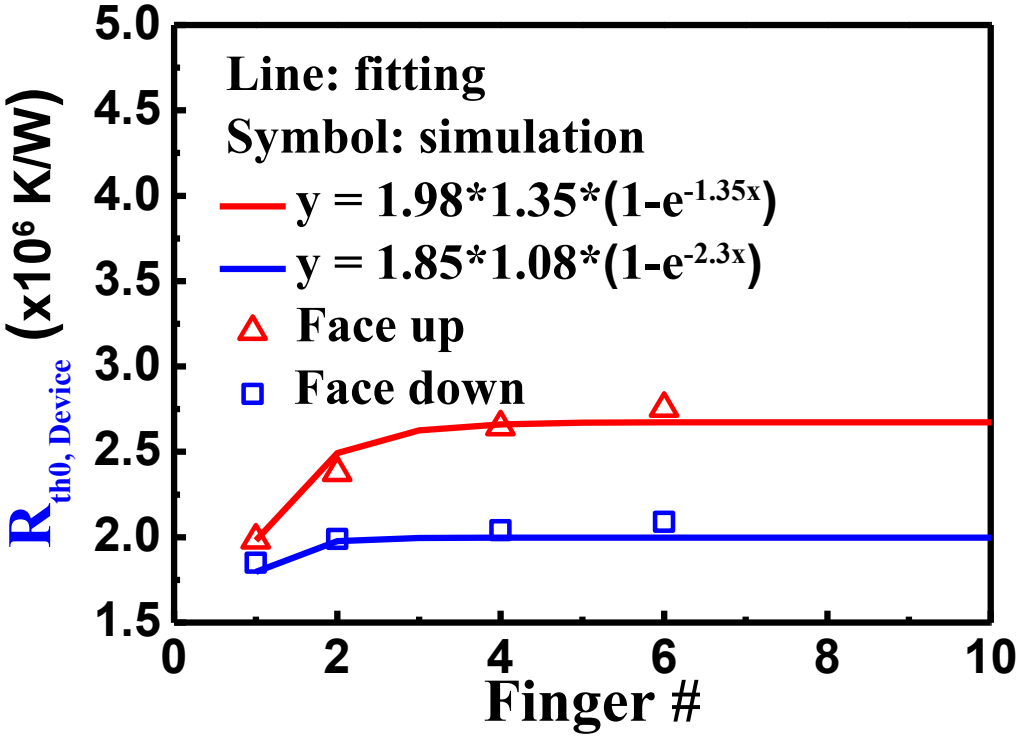


Fig. 2. 12 $R_{th0,device}$ fitting as a function of finger number for face-up and face-down configurations

doi:10.6342/NTU201801788

The contact area of metal gate increases with the increasing fin height, leading to the decreasing $R_{th0,FinFET}$ in both cases.

The fitted $R_{th0,FinFET}$ are:

$$R_{th0,FinFET} = 1.98 \times \left[-0.88\left(\frac{H_{fin}}{40nm}\right) + 2.31\right]\left(\frac{K}{\mu W}\right) \quad (2.14)$$

, and

$$R_{th0,FinFET} = 1.85 \times \left[-1.2\left(\frac{H_{fin}}{40nm}\right) + 2.59\right]\left(\frac{K}{\mu W}\right) \quad (2.15)$$

for face-up and face-down configurations, respectively.

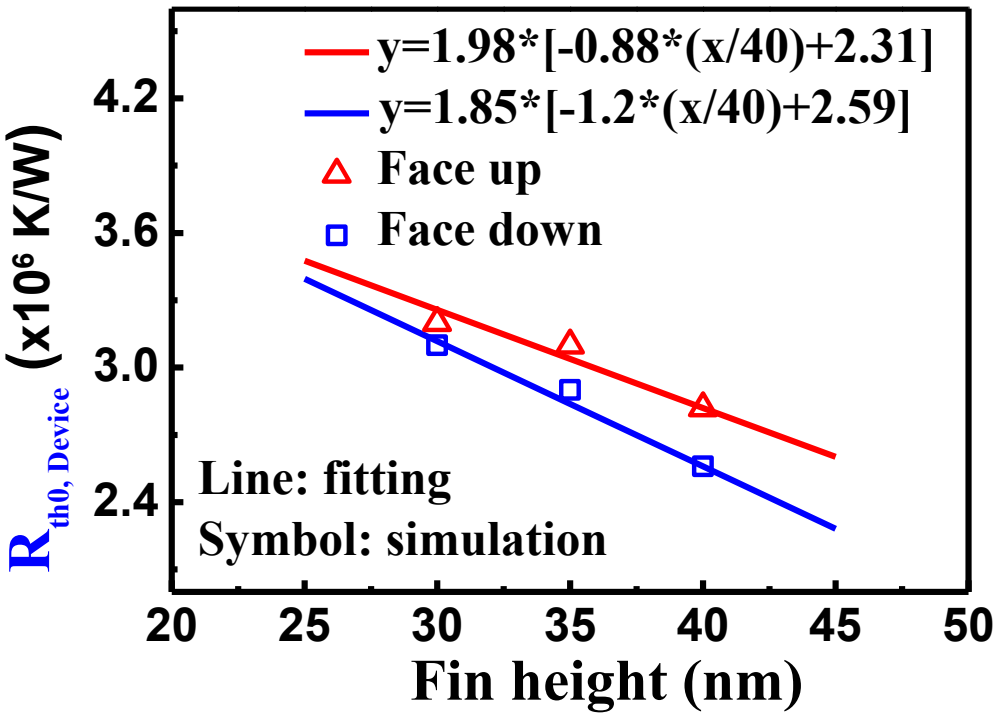


Fig. 2. 13 $R_{th0,device}$ fitting as a function of fin height for face-up and face-down configurations

doi:10.6342/NTU201801788

The summarized final equations are:

Face-up:

$$R_{th0,device} = 1.98 \times [0.13(fin\ \#) + 0.87] \times \left[1.35\left(1 - e^{-1.35\times(finger\ \#)}\right)\right] \times \left[-0.88\left(\frac{H_{fin}}{40nm}\right) + 2.31\right] \left(\frac{K}{\mu W}\right) \quad (2.16)$$

Face-down:

$$R_{th0,device} = 1.85 \times [0.13(fin\ \#) + 0.89] \times \left[1.08\left(1 - e^{-2.30\times(finger\ \#)}\right)\right] \times \left[-1.2\left(\frac{H_{fin}}{40nm}\right) + 2.59\right] \left(\frac{K}{\mu W}\right) \quad (2.17)$$

The face-down configuration has a lower $R_{th0,device}$ because the FinFET is not a vertically zygomorphic structure. With face-down, the heat flux flows through the via0s which have a smaller thermal resistivity as compared to the channel stopper and STI of the FinFET.

2.5 Summary

In this chapter, the intrinsic thermal resistance of 14nm FinFETs is simulated using TCAD and extracted from the thermal resistance circuit. To decouple the $R_{th0,FinFET}$, accurate thermal resistance of BEOL is essential. A two-step pseudo isothermal plane model is used to calculate the $R_{th,BEOL}$ based on the non-flat isothermal plane in the BEOL observed in the simulation results. The extracted $R_{th0,FinFET}$ has no substrate thickness dependence and is the intrinsic characteristic of FinFETs.

doi:10.6342/NTU201801788

The free convection of air has a large thermal resistance and should be considered for both face-up (conventional measurement, heat flow from the channel to substrate) and face-down (flip-chip, heat flow from the channel to metal contact) configurations.

The FinFET is not a vertically zygomorphic structure and the heat flow direction affects $R_{th0,FinFET}$. The layout and geometry dependence of $R_{th0,FinFET}$ is summarized by fitted equations. The face-up $R_{th0,FinFET}$ is higher than face-down $R_{th0,FinFET}$.

Device temperature with arbitrary BEOL metal density can be predicted by the proposed $R_{th0,device}$ and $R_{th,BEOL}$ model. The $R_{th0,device}$ model can also be adopted in the circuit simulation for IC design.

2.6 References

[2.1] C.W. Chang, S.E. Liu, B.L. Lin, C.C. Chiu, Y.-H. Lee, and K. Wu, " Thermal Behavior of Self-heating Effect in FinFET Devices Acting on Back-end Interconnects," *Reliability Physics Symposium (IRPS)*, *2015 IEEE International*, pp. 2F.6.1- 2F.6.5.

[2.2] E. Bury, B. Kaczer, D. Linten, L. Witters, H. Mertens, N. Waldron, X. Zhou, N. Collaert, N. Horiguchi, A. Spessot, and G. Groeseneken, " Self-heating in FinFET and GAA-NW using Si, Ge and III/V channels," in *IEEE IEDM Tech. Dig.*, Dec. 2016, pp. 15.6.1-15.6.4.

doi:10.6342/NTU201801788

[2.3] Muhammad Abdul Wahab, SangHoon Shin, and Muhammad Ashraful Alam, " 3D modeling of spatio-temporal heat-transport in III-V gate-all-around transistors allows accurate estimation and optimization of nanowire temperature, " *IEEE Transactions on Electron Devices*. Vol.62, No.11, Nov. 2015, pp.3595-3604.

[2.4] F. Menges, F. Motzfeld, H. Schmid, P. Mensch, M. Dittberner, S. Karg, H. Riel, and B. Gotsmann, " Local thermometry of self-heated nanoscale devices," in *IEEE IEDM Tech. Dig.*, Dec. 2016, pp. 15.8.1-15.8.4.

[2.5] D. Jang, E. Bury, R. Ritzenthaler, M. Garcia Bardon, T. Chiarella, K. Miyaguchi, P. Raghavan, A. Mocuta, G. Groeseneken, A. Mercha, D. Verkest, and A. Thean, " Self-heating on bulk FinFET from 14nm down to 7nm node," in *IEEE IEDM Tech. Dig.*, Dec. 2015, pp. 11.6.1-11.6.4.

[2.6] S. Natarajan, M. Agostinelli, S. Akbar, M. Bost, A. Bowonder, V. Chikarmane, S. Chouksey, A. Dasgupta, K. Fischer, Q. Fu, T. Ghani, M. Giles, S. Govindaraju, R. Grover, W. Han, D. Hanken, E. Haralson, M. Haran, M. Heckscher, R. Heussner, P. Jain, R. James, R. Jhaveri, I. Jin, H. Kam, E. Karl, C. Kenyon, M. Liu, Y. Luo, R. Mehandru, S. Morarka, L. Neiberg, P. Packan, A. Paliwal, C. Parker, P. Patel, R. Patel, C. Pelto, L. Pipes, P. Plekhanov, M. Prince, S. Rajamani, J. Sandford, B. Sell, S. Sivakumar, P. Smith, B. Song, K.

doi:10.6342/NTU201801788

Tone, T. Troeger, J. Wiedemer, M. Yang, K. Zhang, " A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588μm^2 SRAM cell size," in *IEEE IEDM Tech. Dig.*, Dec. 2014, pp. 3.7.1-3.7.4.

[2.7] Ramez Cheaito, John C. Duda, Thomas E. Beechem, Khalid Hattar, Jon F. Ihlefeld, Douglas L. Medlin, Mark A. Rodriguez, Michael J. Campion, Edward S. Piekos, and Patrick E. Hopkins, " Experimental investigation of size effects on the thermal conductivity of silicon-germanium alloy thin films," *Physical Review Letters*, 109, 195901, 2012.

[2.8] Changwook Jeong, Supriyo Datta, and Mark Lundstrom, " Thermal conductivity of bulk and thin-film silicon: A Landauer approach," Journal of Applied Physics, 111.9, 093708, 2012.

[2.9] Hazim B. Awbi, " Calculation of convective heat transfer coefficients of room surfaces for natural convection," Energy and Building, 28. 2, 1998, pp. 219-227.

doi:10.6342/NTU201801788

Chapter 3 Transient Thermal Simulation of Inverters with Capacitive Loading Effects

3.1 Introduction

AC self-heating effect in scaled FinFET circuits retards the performance, i.e. dark silicon. It was reported that about 20% of the chip area will be dark at 22nm node and more than 50% of the chip cannot be utilized for technology nodes beyond 10nm [3.1]. Circuit level thermal management and self-heating awared design rules are the keys to achieve robust circuit design.

In chapter 2, we've pointed out that the embedded SH measurements such as resistance thermometry structure [3.2],[3.3] and heater-sensor structure [3.3] can only provide the average temperature near the DUT instead of device T_j, and the precision of temperature is suffering from the difference of boundary condition between face-up and face-down configurations. To monitor DC SHE in scaled FinFETs, the intrinsic thermal resistance ($R_{th0,FinFET}$) model is proposed in chapter 2.

In this chapter, transient SHE in FinFETs and inverters will be investigated using TCAD simulation. Capacitive loading effects of the inverter on T_j and the temperature distribution in the metal layer of an inverter are investigated with AC input.

doi:10.6342/NTU201801788

3.2 Self-heating Effect in FinFETs with Transient Input

The transient SHE in scaled FinFET is simulated by transient input with different pulse width. **Fig. 3. 1 (a)** shows the simulated cross-sectional temperature contour with different heating times. It is observed that different heating times lead to different hot volumes. With the pulse width of 25ns, the hotspot is localized in the drain extension. It could be even more difficult to measure the SHE with transient input than DC input. The high temperature region extends to the substrate and MEOL with the pulse width of 150ns.

The Cooling time relates to the heat dissipating ability (R_{th}), and thermal capacitance (C_{th}). The C_{th} is proportional to the hot volume in the device. As a result, the short heating time (small hot volume) leads to short cooling time (**Fig. 3. 1 (b)**).

3.3 Electrical Characteristics of the Inverter

Capacitive loading is essential for the mix-mode electro-thermal simulation. An inverter consists of a four-fin-one-finger nFinFET and a four-fin-one-finger pFinFET is experimentally measured by a square wave input with the frequency of 3GHz and the amplitude of 1.1V (**Fig. 3. 2 (a)**). The corresponding output voltage and output current of are shown in **Fig. 3. 2 (a)** and **(b)**, respectively.

doi:10.6342/NTU201801788

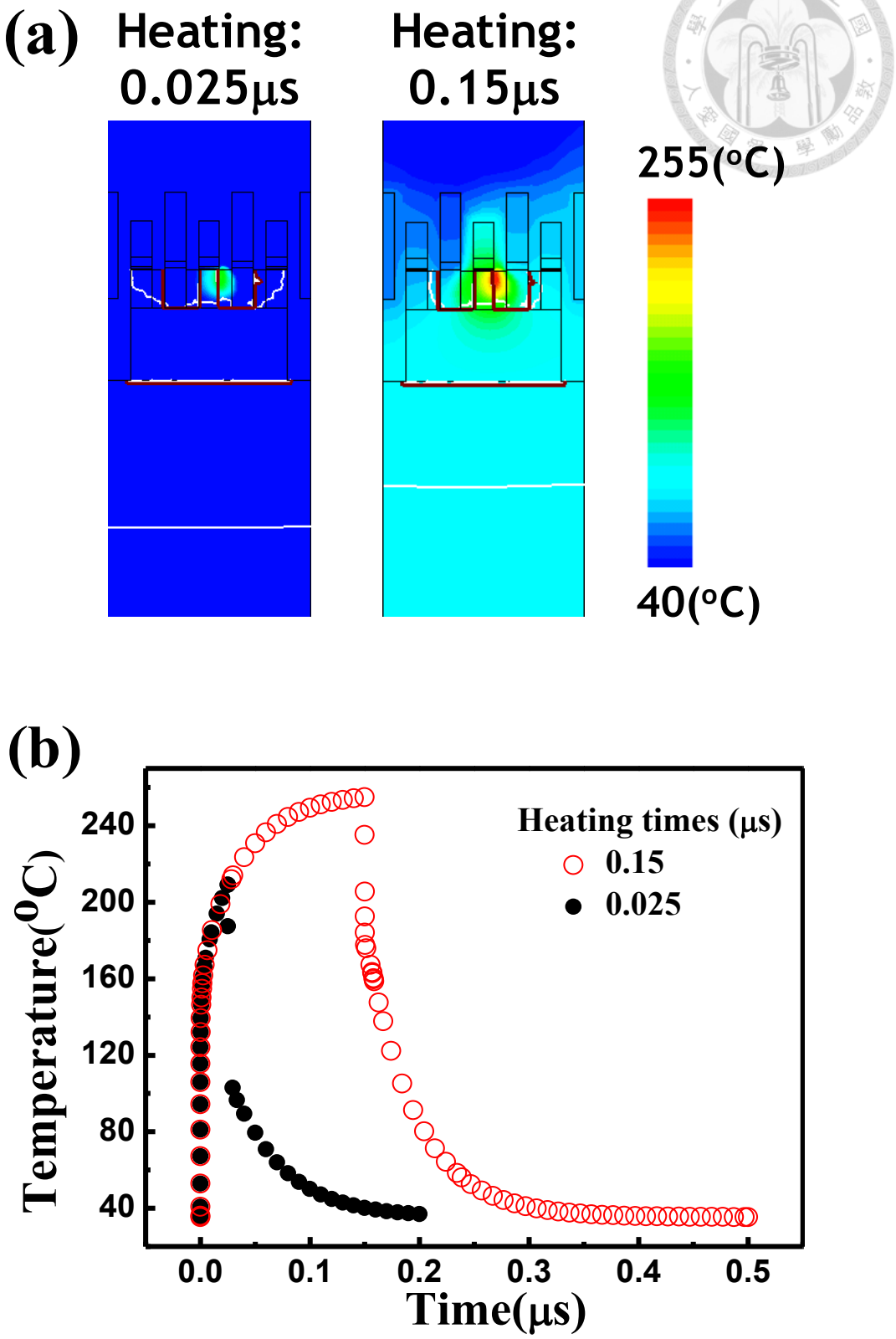


Fig. 3. 1 (a) Cross-sectional temperature distribution and **(b)** transient T_j response with different input pulse width

doi:10.6342/NTU201801788

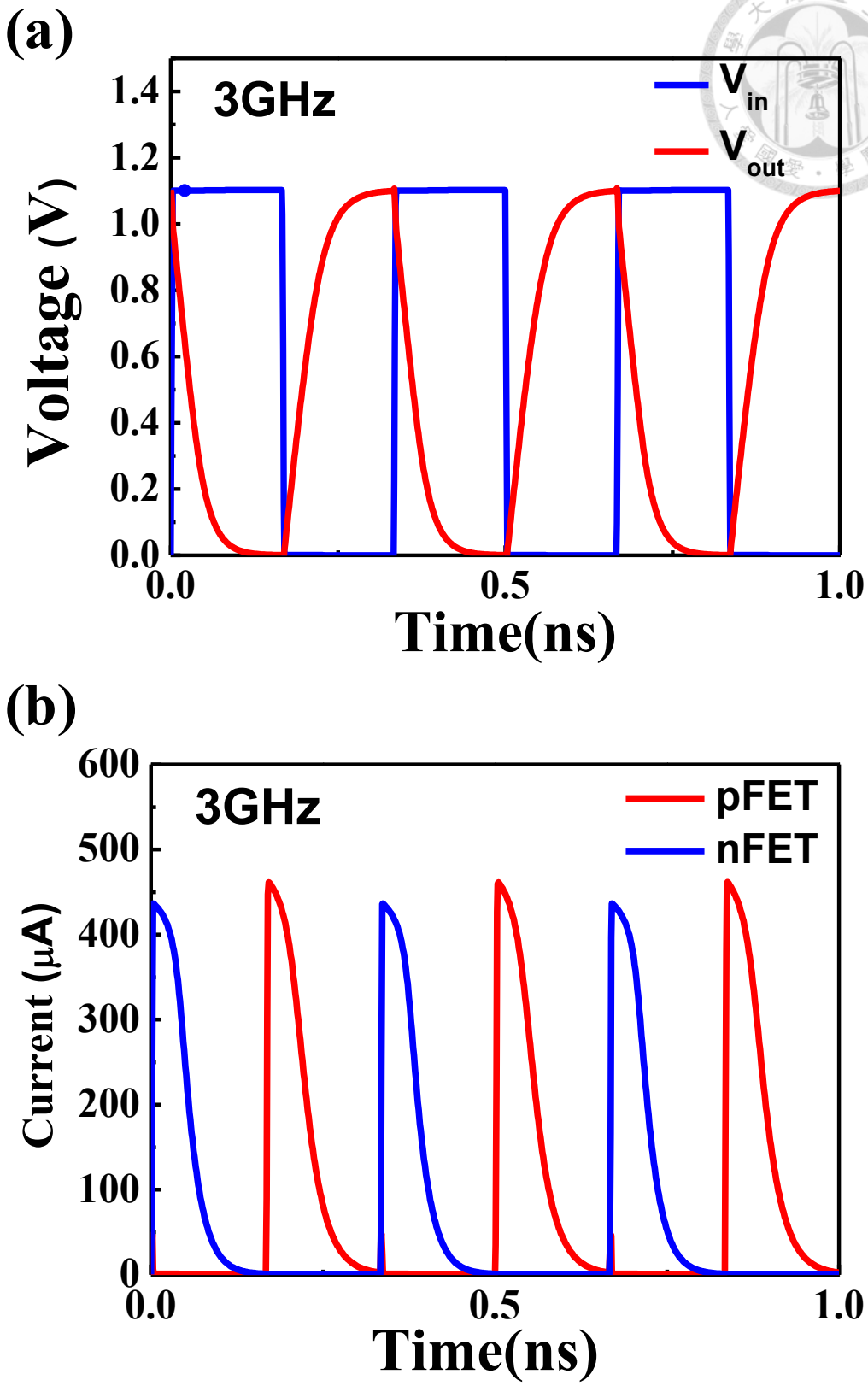


Fig. 3. 2 (a) Measured input and output voltage of an inverter and **(b)** the transient responses of output current and simulated T_j of the inverter

doi:10.6342/NTU201801788

The amount of charge/discharge is obtained by integrate the output current in a period:

$$\mathrm{Q} = \int I(t) \times dt = \ 2.38E-14(C) \qquad (3.1)$$

The output capacitive loading is calculated by:

$$\mathrm{C}_{load} = \frac{Q}{V} = \frac{2.38E-14}{1.1} = 2.164\mathrm{E}-14 = 21.64 \ (fF) \qquad (3.2)$$

3.4 Self-heating and Output Loading Effect in Inverters

The transient response of the T_j in the inverter consist of 4-fin-1-finger n/pFinFETs with output capacitance of 21.64 fF is investigated by mix-mode TCAD simulation.

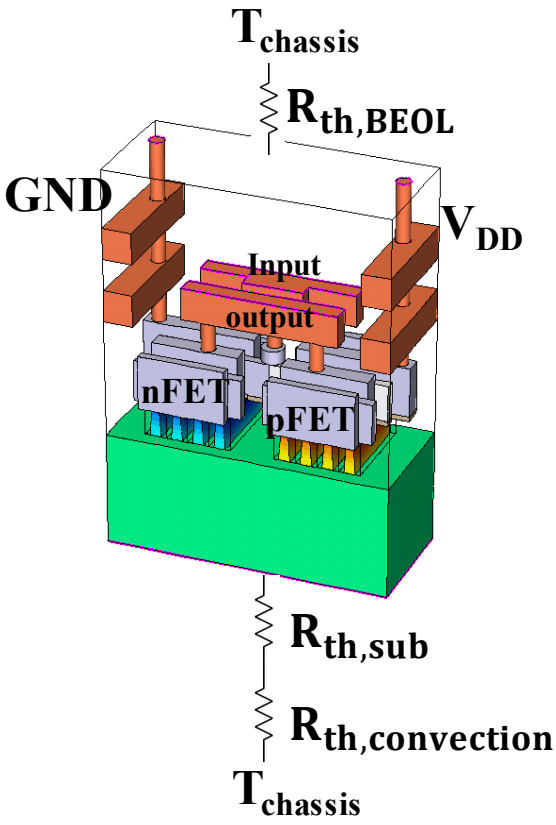


Fig. 3. 3 Simulation structure of the inverter with face-down configuration

doi:10.6342/NTU201801788

The simulation structure is shown in **Fig. 3. 3**, and the simulated T_{max} in the inverter is located at the channel of turned on device (**Fig. 3. 4**). Due to the low thermal conductivity of SiGe S/D, the highest temperature during transient input is with pFET turned on (**Fig. 3. 5**).

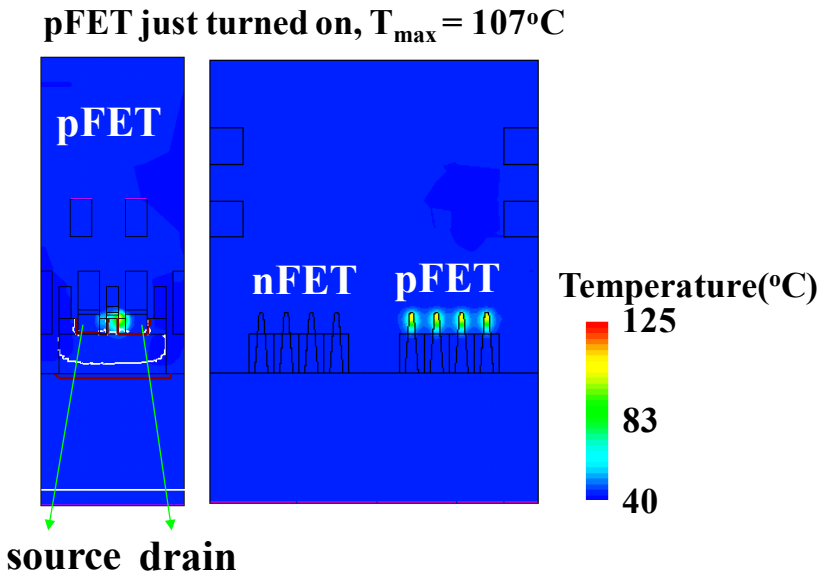


Fig. 3. 4 Cross-sectional temperature distribution in the inverter

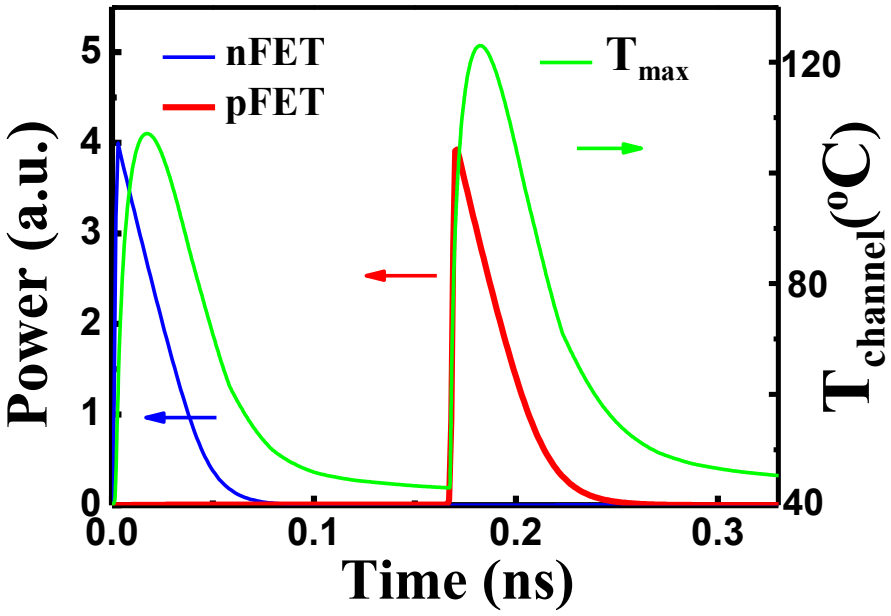


Fig. 3. 5 Transient responses of input power and the T_{max} in the inverter

doi:10.6342/NTU201801788

3.4.1 Output Capacitive Loading Effects on T_j

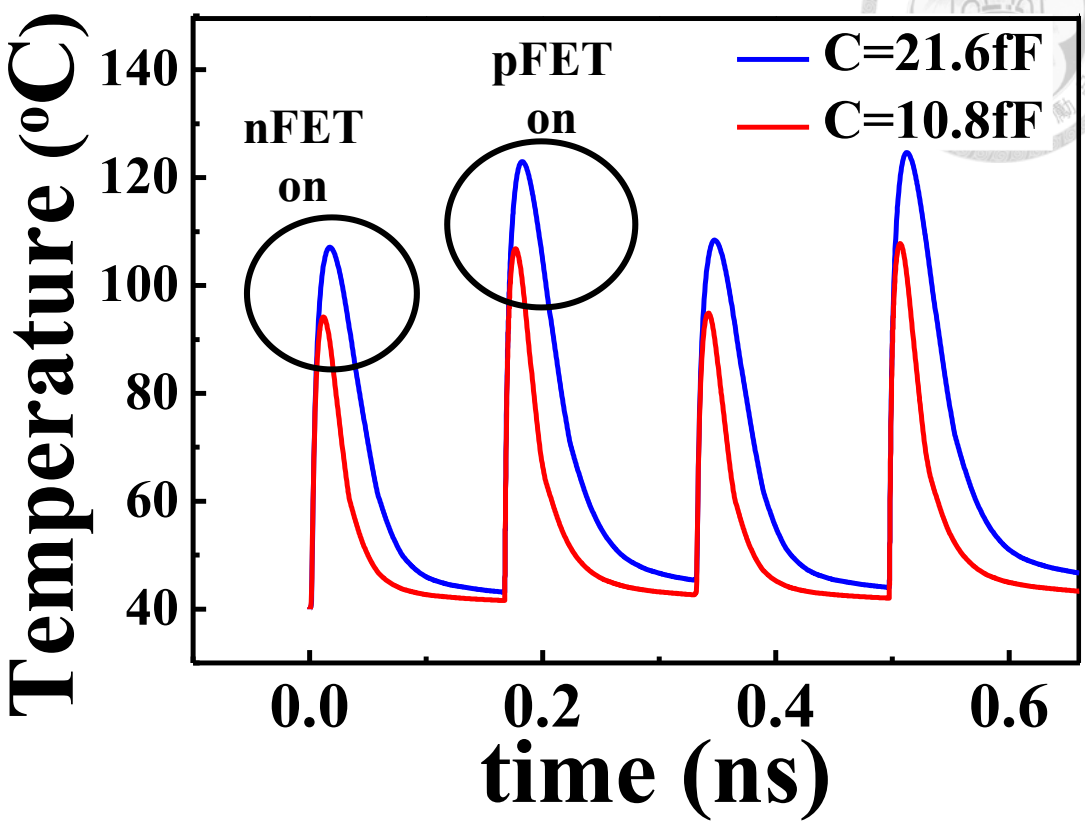


Fig. 3. 6 The device temperature of inverters with different C_{load}

Inverters with 4-fin-1-finger n/pFinFETs are simulated with different capacitive loadings under 3GHz square wave input. **Fig. 3. 6** shows the simulated transient temperature response. With the low output capacitance of 10.8fF, both the T_{max} and the high temperature duration decreases due to the reduction of charge and discharge time. Lowering the parasitic capacitance could help to gain benefits not only from reducing the latency but also from lowering the temperature to maintain the performance and improve the reliability.

doi:10.6342/NTU201801788

3.4.2 Drive Current, T_j, and High Temperature Duration

T_{max} in an inverter consists of 2-fin-1-finger n/pFinFETs with the capacitive loading of 10fF is simulated with different I_{on} by intentionally changing the channel mobility.

Increasing T_{max} is found with the increasing I_{on} due to the increased Joule heating (**Fig. 3. 7 (a)**). The degradation mechanisms could be accelerated by the increased temperature. However, the degradations are also proportional to the duration of stress condition. Decreasing I_{on} can help decreasing the Joule heating and T_{max}, but the charge time is extended which results in an increased hot volume and longer high temperature duration (**Fig. 3. 7 (b)**). The drive current of the inverter affects not only the Joule heat but also the time for charge/discharge, and should be designed carefully from the aspect of reliability.

3.4.3 Layout Size Effects of Inverters

T_{max} decreases with the increasing layout size when charging and discharging a 5fF output loading (**Fig. 3. 8 (a).**) The reduction of T_{max} is due to the reducing $R_{th,BEOL}$ by increasing device area. However, in real circuit, the layout sizes of connected inverters are usually the same, i.e. the output capacitive load is proportional to the layout size, and the charge/ discharge time increases as the output load increases.

doi:10.6342/NTU201801788

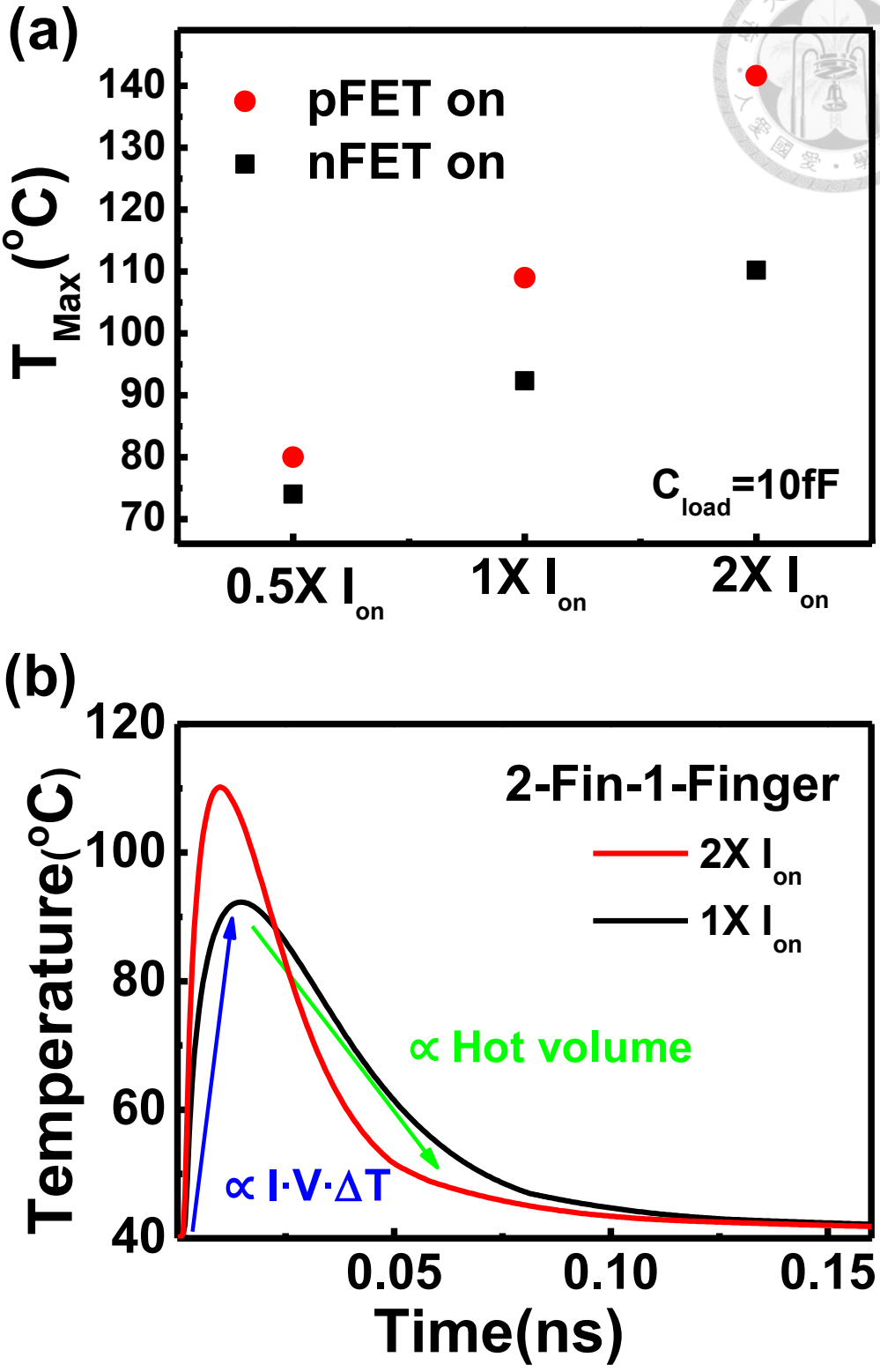


Fig. 3. 7 T_{max} of an inverter (2-fin-1-finger n/pFinFETs) with different I_{on}, and **(b)** the temperature evolution

doi:10.6342/NTU201801788

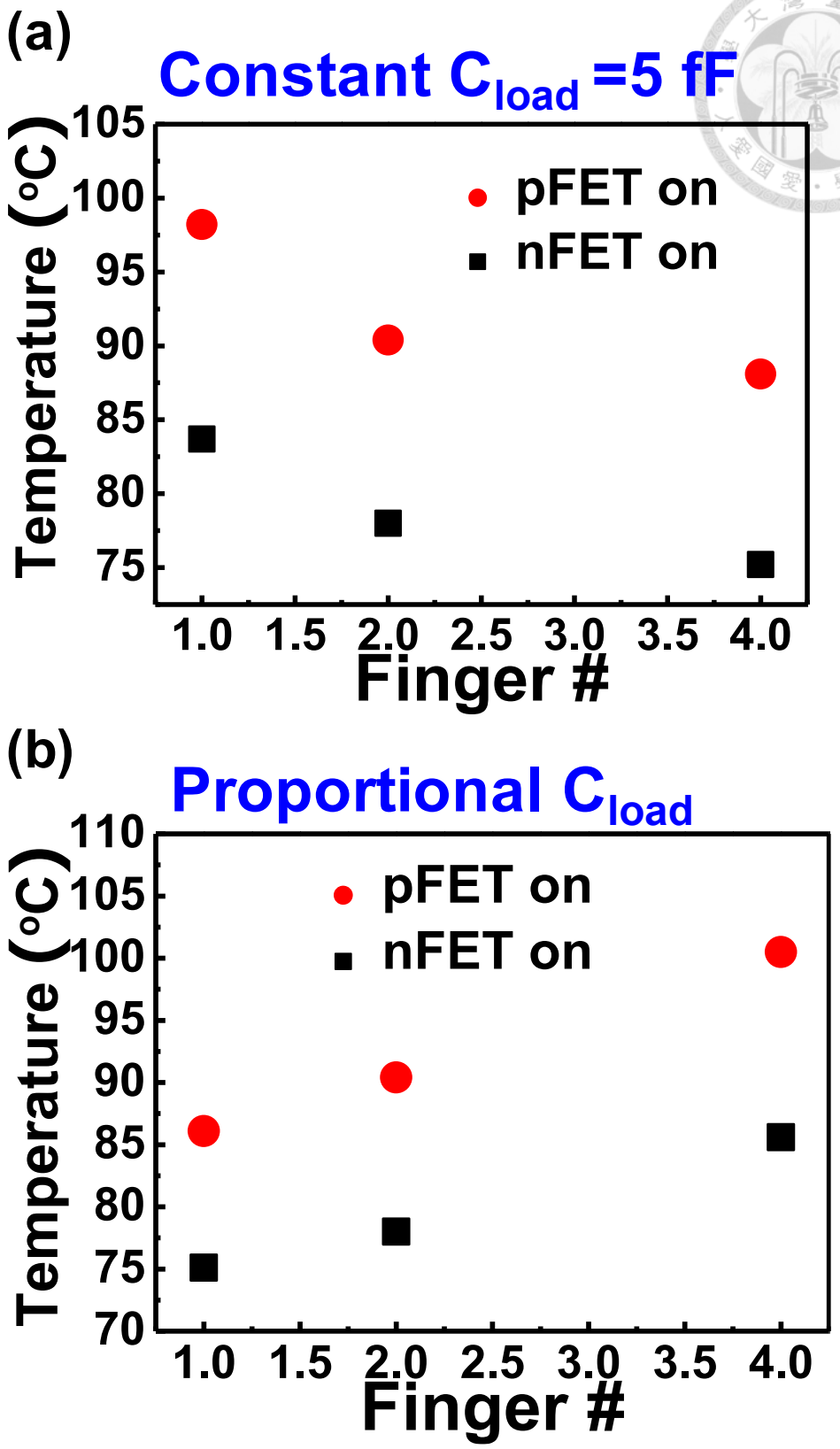


Fig. 3. 8 T_{max} vs finger number with **(a)** the capacitive loading of 5fF and **(b)** the capacitive loading proportional to layout size

doi:10.6342/NTU201801788

Output capacitive loads of 2.5fF, 5fF, and 10fF are applied to inverters with one finger, two fingers, and four fingers, respectively. The simulated T_{max} increases with the increasing output load (**Fig. 3. 8 (b)**), meaning that the increased output load outweighs the reduction of $R_{th,BEOL}$ on T_{max}. This results limit the maximum fin number and finger number of the FinFETs in an inverter.

3.5 SHE of Inverters with AC input

3.5.1 Temperature Distribution on M1 Layer

Snap shots of the temperature distribution on the top surface of M1 layer in one operation are shown in the figures below. At the very beginning, the devices are not turned on yet, and the temperature equals to the chassis temperature (**Fig. 3. 9**)

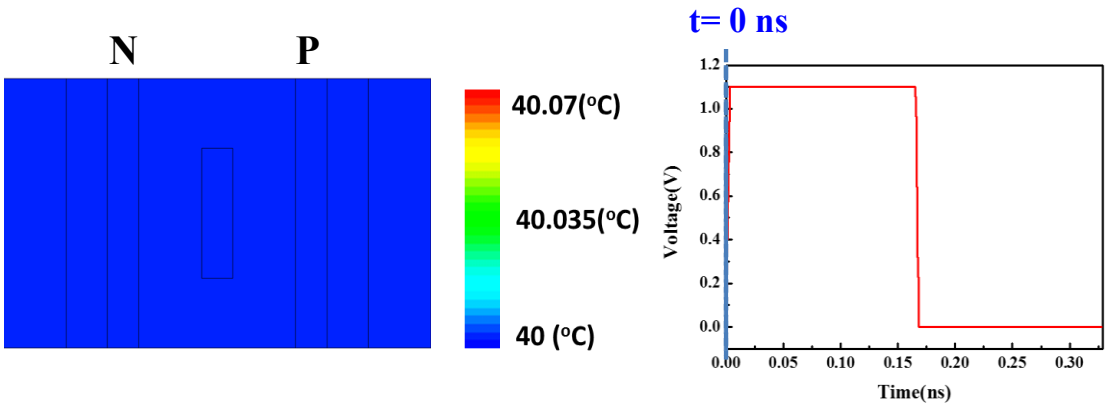


Fig. 3. 9 Snap shot on the M1 layer at t= 0 ns

Fig. 3. 10 shows the temperature distribution in the half of the period. The nFET is just turned off, and the rising of temperature could be observed on drain extension of nFET and the metal gate. However, due to the short heating time, the temperature

doi:10.6342/NTU201801788

difference is only 0.01°C on the metal line.

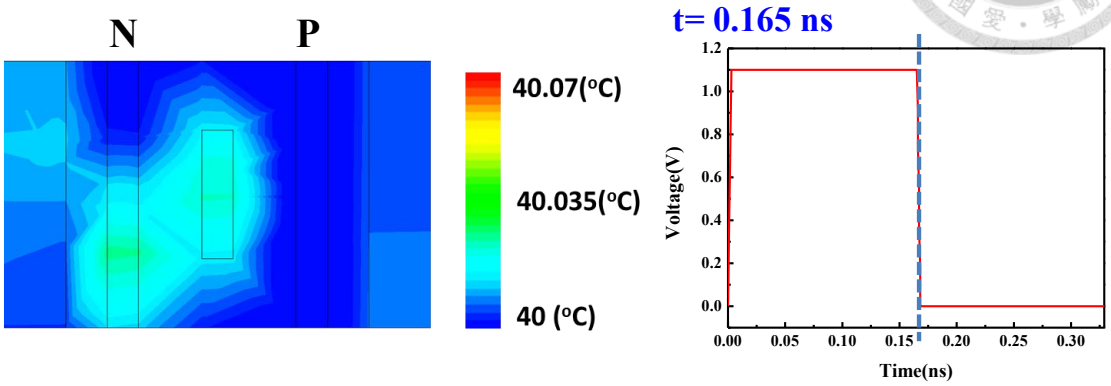


Fig. 3. 10 Snap shot on the M1 layer at t= 0.165 ns

At the end of the operation, the hot spot in M1 layer is found localized, and the ΔT is only 0.07°C (**Fig. 3. 11**).

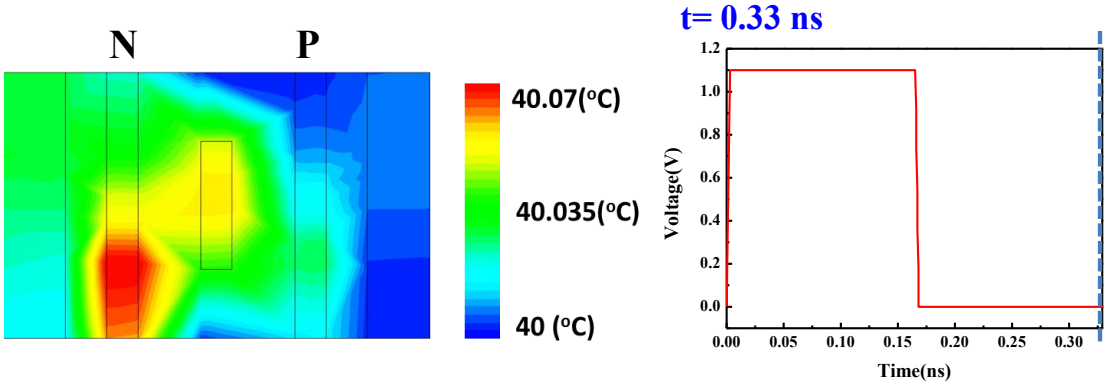


Fig. 3. 11 Snap shot on the M1 layer at t= 0.33 ns

The resistance of a metal line changes with its average temperature. The simulation results show that with a transient AC input, the resistance change of the metal line cannot reflect the temperature evolution in the device.

doi:10.6342/NTU201801788

3.5.2 Residual Temperature in the Channel

AC simulation of 1-fin-1finger n/pFinFETs inverters with the loading capacitance of 2.5fF and 5fF are simulated. Low residual temperature of 13.5°C (**Fig. 3. 12**) and 21°C (**Fig. 3. 13**) is observed in the channel of the inverters after 1500 and 1000 V_g cycles, respectively (square waves with the duty cycle of 50% and the frequency of 3 GHz). In every single cycle, the high T_j over 80°C is observed in both cases. However, the hot volume is small due to the short heating time and most heat dissipates at the end of the period. With AC input, the increased T_j by SHE is hard to be measured.

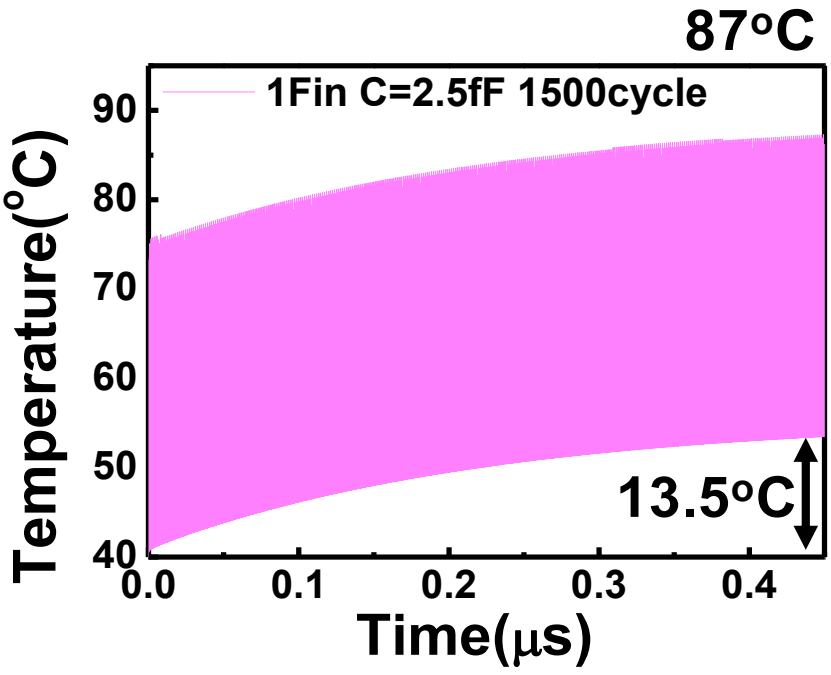


Fig. 3. 12 Residual temperature in the channel of an inverter with C_{load} = 2.5fF after 1500 cycles

doi:10.6342/NTU201801788

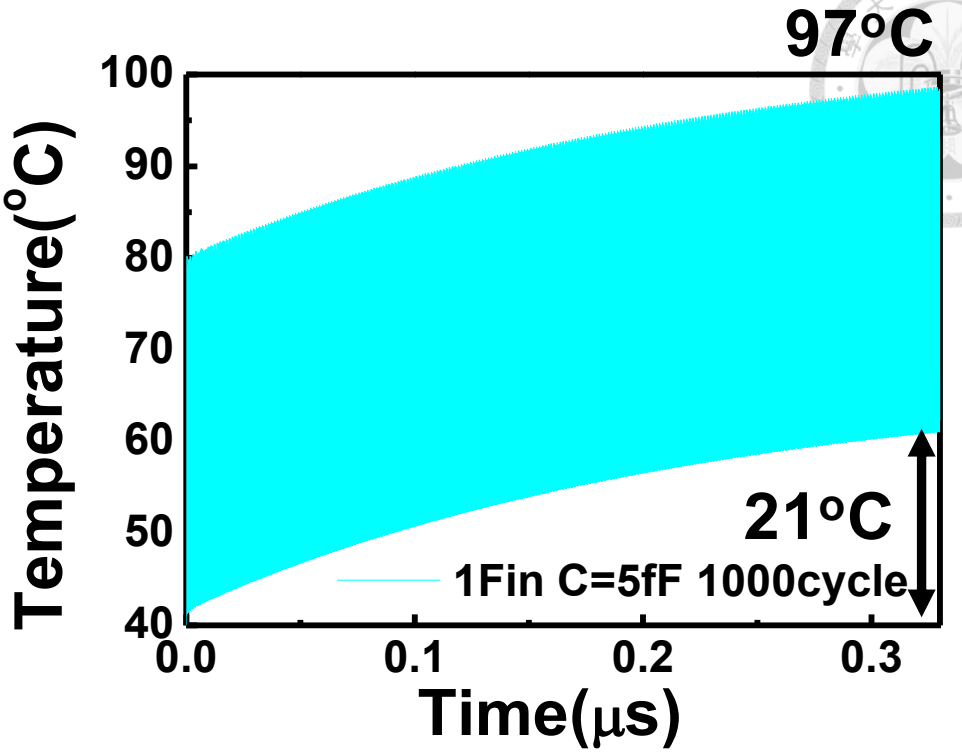


Fig. 3. 13 Residual temperature in the channel of an inverter with C_{load} = 5fF after 1000 cycles

3.6 Summary

In this chapter, SHE of inverters with AC input is simulated. The C_{load} affects the charge/discharge time. C_{load} of an inverter consists of a four-fin-one-finger nFinFET and a four-fin-one-finger pFinFET is extracted from the experimentally measured output current.

The maximum temperature and the high temperature duration can be controlled by output capacitive loading and the drive current of the inverter.

With AC input, the temperature in M1 layer and the residual temperature in the channel are found too low as compared to the device T_j, which may lead to an

doi:10.6342/NTU201801788

underestimation of device temperature if the M1 is temperature indicator.

Both the maximum temperature and the high temperature duration are important for device reliability with transient input and capacitive loading. The temperature evolution can only be obtained by mix-mode electro-thermal simulation.

3.7 References

[3.1] Hadi Esmaeilzadeh, Emily Blem, Rene´e St. Amant, Karthikeyan Sankaralingam, and Doug Burger, " Dark Silicon and the End of Multicore Scaling," *IEEE Micro, vol.32, no.3,* May/June 2012, pp.122-134.

[3.2] C.W. Chang, S.E. Liu, B.L. Lin, C.C. Chiu, Y.-H. Lee, and K. Wu, " Thermal Behavior of Self-heating Effect in FinFET Devices Acting on Back-end Interconnects," *Reliability Physics Symposium (IRPS)*, *2015 IEEE International*, pp. 2F.6.1- 2F.6.5.

[3.3] E. Bury, B. Kaczer, D. Linten, L. Witters, H. Mertens, N. Waldron, X. Zhou, N. Collaert, N. Horiguchi, A. Spessot, and G. Groeseneken, " Self-heating in FinFET and GAA-NW using Si, Ge and III/V channels," in *IEEE IEDM Tech. Dig.*, Dec. 2016, pp. 15.6.1-15.6.4.

doi:10.6342/NTU201801788

Chapter 4 Comprehensive Thermal SPICE Modeling of FinFETs and Back-end-of-line

4.1 Introduction

Precise AC response of spatial and temporal temperature distribution is essential for accurate reliability lifetime projection for devices and BEOL. Several methods have been reported for SHE measuring, but none of them can provide sufficient resolution of spatial and temporal temperature [4.1]. Therefore, the device simulation is the only way to know the exact temperature distribution and time evolution.

In the previous chapters, device simulation is adopted to evaluate the SHE in scaled FinFET. Intrinsic thermal resistance of FinFETs ($R_{th0,FinFET}$) is decoupled to model the DC SHE. Transient device simulation is demonstrated with inverters. However, the computational cost is the shortcoming of TCAD which makes it unsuitable for circuit design. For example, it took us one month to simulate the transient temperature response of an inverter to 1500 cycles using a PC with the 8^{th} generation intel core-i7 processor.

Instead of device simulation, compact thermal model (CTM) is usually used with their high computational efficiency. CTMs consist of R_{th} networks [4.2]-[4.5] and lumped R_{th}-C_{th} [4.6]- [4.8] were reported to simulate the steady state and transient SHEs,

doi:10.6342/NTU201801788

respectively. However, the element size, which is essential for transient simulation, is decided not precisely either by the heat loss in a material (e.g. characteristic thermal length=30nm for the channel [4.5]) or by arbitrarily discretizing the physical dimensions of the device. Note that CTMs with large element size could miss the details of SHE, particularly for AC input.

In this chapter, a SPICE model is proposed and validated by 3-D electro-thermal TCAD. Guideline for determining the element size is reported using the frequency dependence of $\tau_{hotspot}$. Multi-fin FinFETs can be assembled easily with our modularized model, no additional device simulation is needed. Device-device interaction and interfacial thermal resistance can be considered in the thermal SPICE model, different from the CTM with lumped elements.

4.2 Frequency Dependent Thermal Time Constant in FinFETs

4.2.1 Issues of Lumped R_{th}-C_{th} models

The lumped R_{th}-C_{th} model (one τ_c model) used in BSIM underestimates the device T_j with AC input. A two τ_c model [4.8] was reported to deal the AC SHE in scaled FinFET by adding an additional lumped R_{th}-C_{th} with a small thermal time constant (~ps) to the one τ_c model of BSIM. However, the ΔT_{MAX} predicted by two τ_c model has no frequency dependence in the range of interest **(Fig. 4. 1)**, different from our

doi:10.6342/NTU201801788

TCAD/SPICE results.

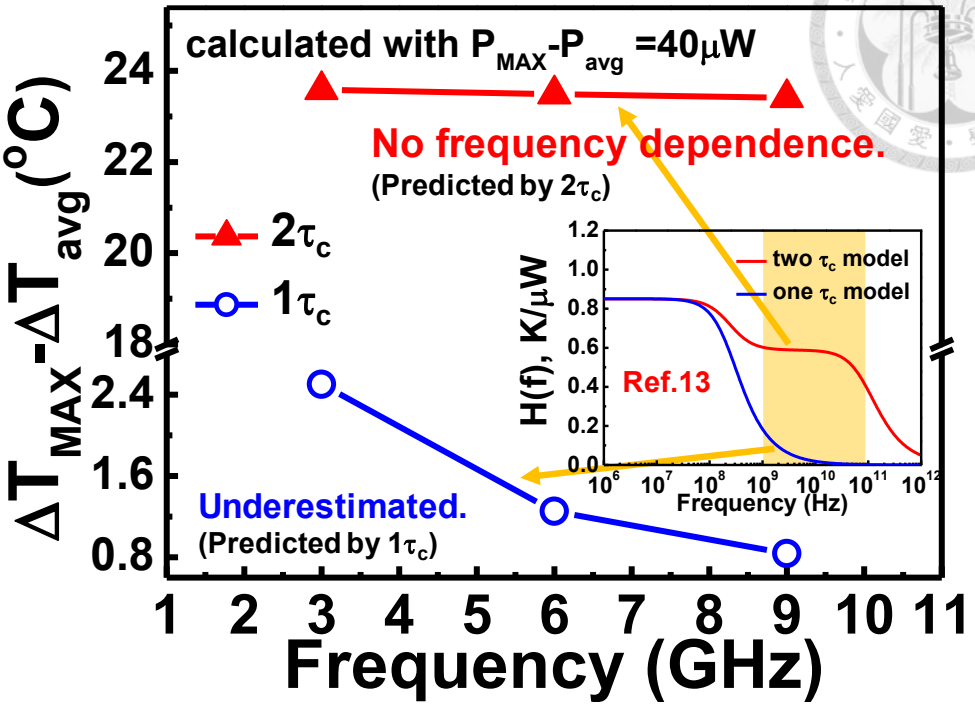


Fig. 4. 1 ΔT_{MAX}-ΔT_{avg} calculated by H(f) (differential R_{th}) times (P_{MAX}-P_{avg}) versus frequency. **(inset)** *H(f)* of a FinFET by lumped R_{th}-C_{th} [4.8]

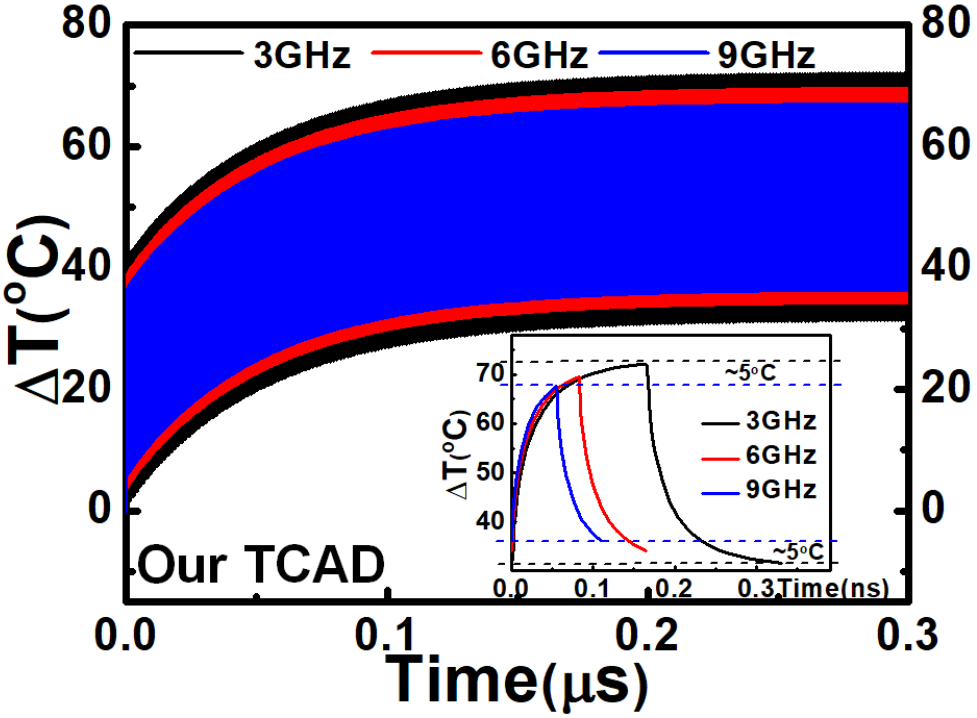


Fig. 4. 2 AC thermal response of a 14nm FinFET **(inset)** AC response in one period

doi:10.6342/NTU201801788

The transient temperature evolution of a 14nm FinFET is simulated by TCAD with 1.1V square input at 3, 6, and 9 GHz under face-down configuration. Both the envelope of the temperature evolution (**Fig. 4. 2**) and the device T_j in one period (**inset** of **Fig. 4. 2**) show clear frequency dependency. The penalties of using lumped R_{th}-C_{th} models are the loss of accuracy and the time consumption on FEM simulations for different device geometries and circuit layouts, (e.g. different isolation spacing between n/pFETs in an inverter).

4.2.2 Thermal time constant of scaled FinFETs

A FinFET with MEOL, $R_{th,BEOL}$, and $R_{th,Sub}$ (**inset** of **Fig. 4. 3**) is simulated by TCAD. The $R_{th,BEOL}$ model has been discussed and verified in chapter 2. With the 1.1V square wave input at 3GHz, localized hot spot is observed in the simulated temperature distribution (**inset** of **Fig. 4. 3**).

The $\tau_{hotspot}$, which is calculated by the time when ΔT_j decreases in value to e^{-1} in one cycle, is inversely proportional to the input frequency in the range interested (**Fig. 4. 3**). The fitted slope of $\tau_{hotspot}$ versus input frequency is -0.71 ps/GHz. To have the SPICE model at an input frequency resolution of 1GHz, the thermal time constant of an element ($\tau_{element}$) should be smaller than 0.71ps.

doi:10.6342/NTU201801788

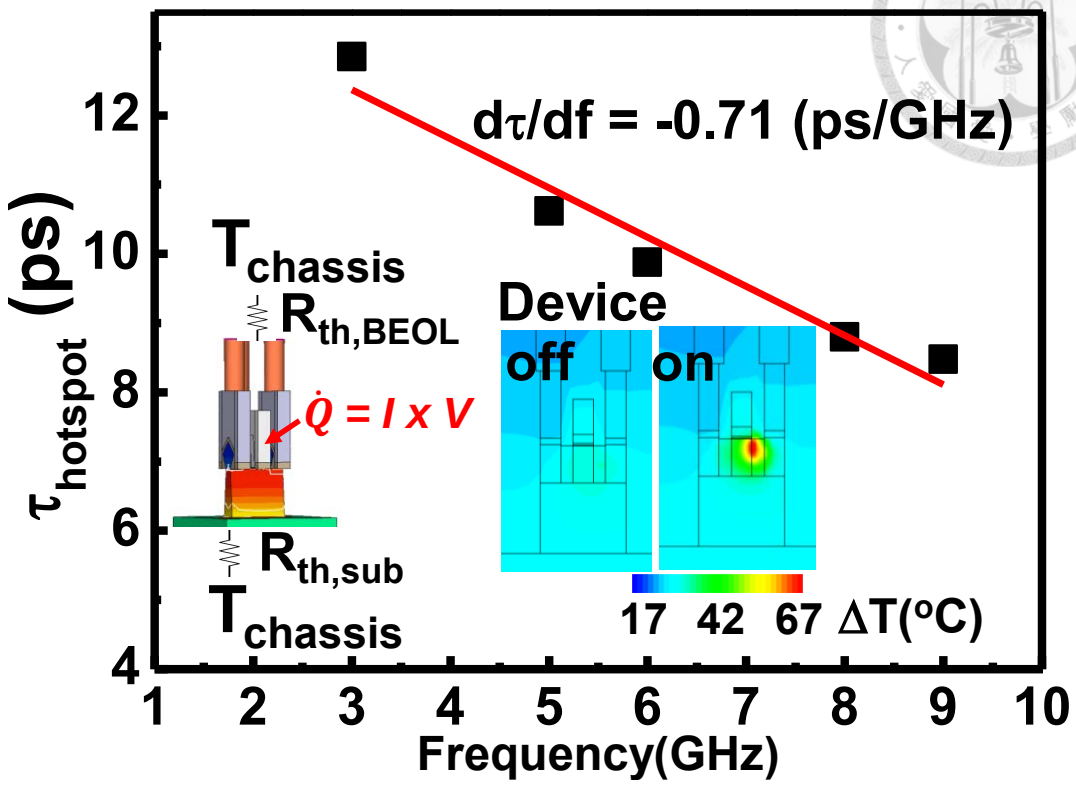


Fig. 4. 3 $\tau_{hotspot}$ versus input frequency. **(inset)** Simulated FinFET with corresponding thermal resistances and the localized hotspot in AC simulation

4.3 Thermal SPICE Modeling of FinFETs

4.3.1 Equivalent Circuit Models in SPICE

To achieve precise temperature mapping in device channel and interconnects, 3D heat flow is considered in the equivalent circuit model of fin and metal (**Fig. 4. 4**). The in-plane heat flow is considered by $R_{th,vertical}$ and $R_{th,horizontal}$, and the out-of-plane heat flow is considered by the connection with neighboring regions, e.g. channel to HKMG.

On the other hand, to reduce the number of element and simulation time. 2D heat

doi:10.6342/NTU201801788

flow is adopted for STI and IMD (**Fig. 4. 5**). In a multi-fin FinFET, The STI and IMD are sandwiched by fins and metals, respectively. The heat flows from hot fins or hot metals are modeled by the equivalent model **(Fig. 4. 5)**. The thermal conductivity of SiO_2 is low as compared to Si fin and metal, and the heat flows in the STI and IMD are ignored.

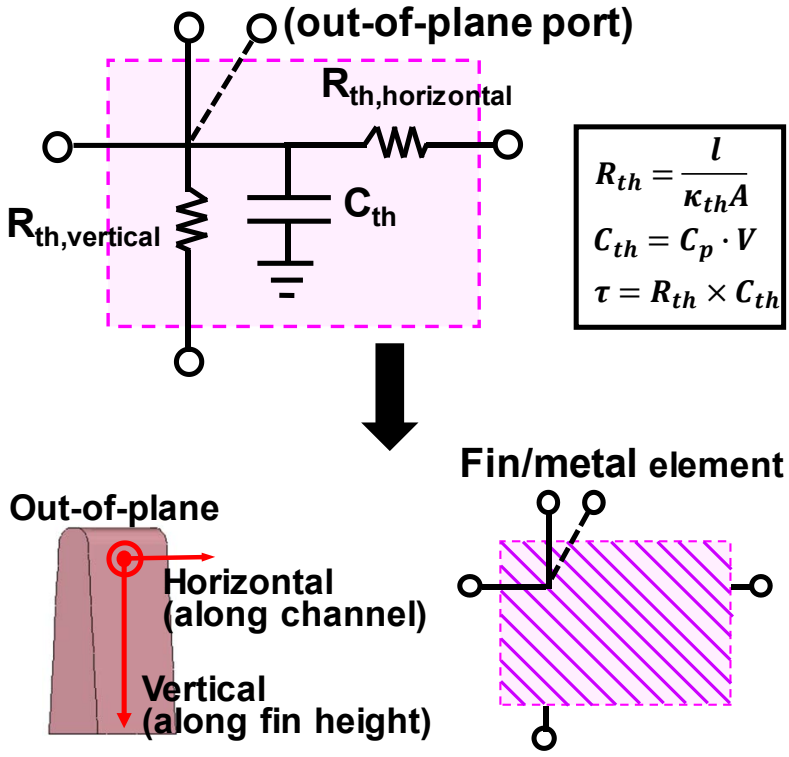


Fig. 4. 4 Equivalent circuit models and simplified color symbols of fin or metal

Distributed R_{th}-C_{th} network is implemented in the SPICE. The R_{th} and C_{th} are calculated by the equations in **Fig. 4. 6**. The element size in the channel region is 2nm x 5nm x W_{fin} which has a calculated $\tau_{element} \approx 0.5$ps

doi:10.6342/NTU201801788

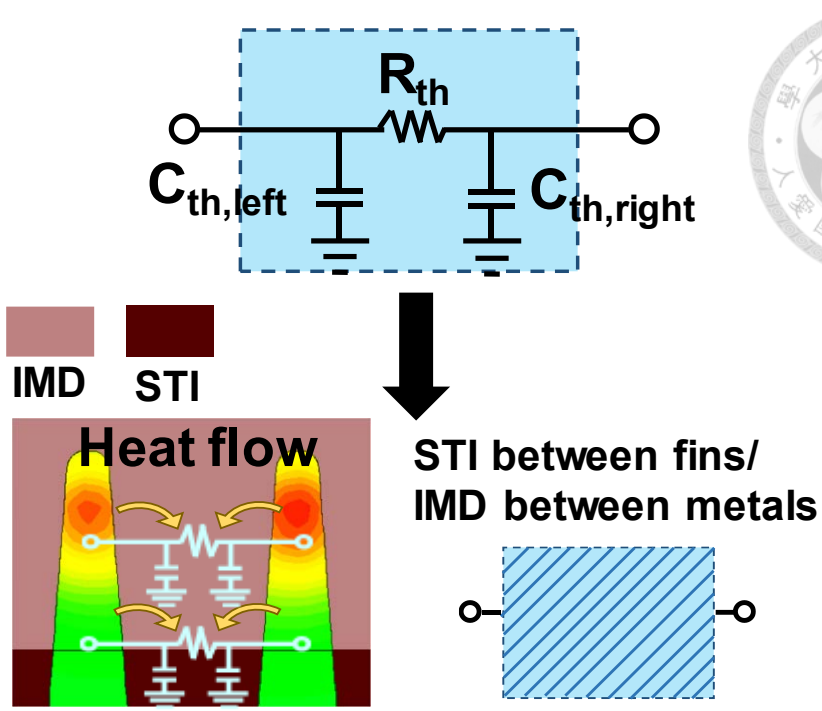


Fig. 4. 5 Equivalent circuit models and simplified color symbols of STI between fins or IMD between metals

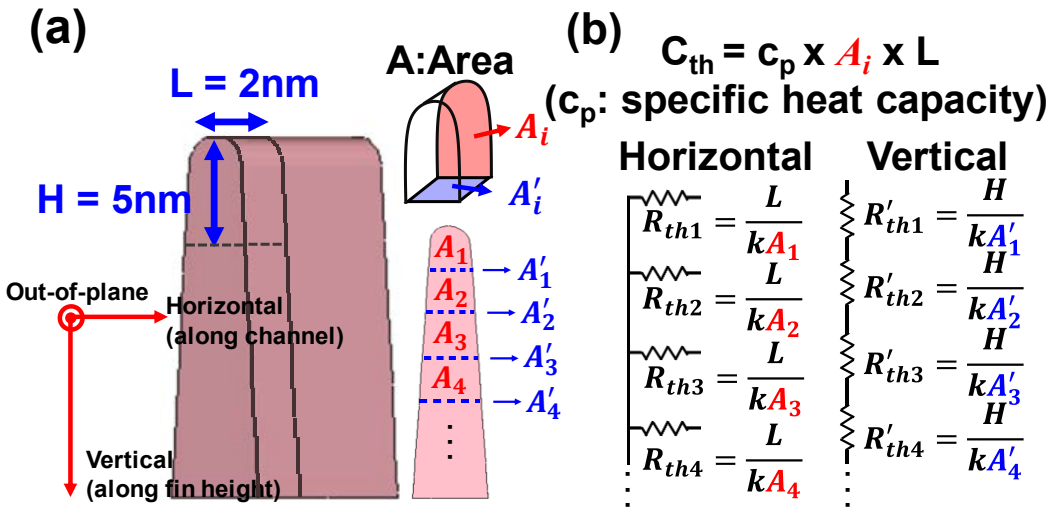


Fig. 4. 6 (a) Element size in a FinFET channel and **(b)** the thermal capacitance, horizontal, and vertical thermal resistances to form a distributed R_{th}-C_{th} model

doi:10.6342/NTU201801788

Fig. 4. 7 (a) schematically shows the SPICE module in the MEOL. The picking up of heat flow which depends on the position of via0s is considered. Gate extension on the STI is implemented by the discrete elements based on the physical geometry.

Fig. 4. 7 (b) shows the TCAD counterparts of the MEOL interconnects.

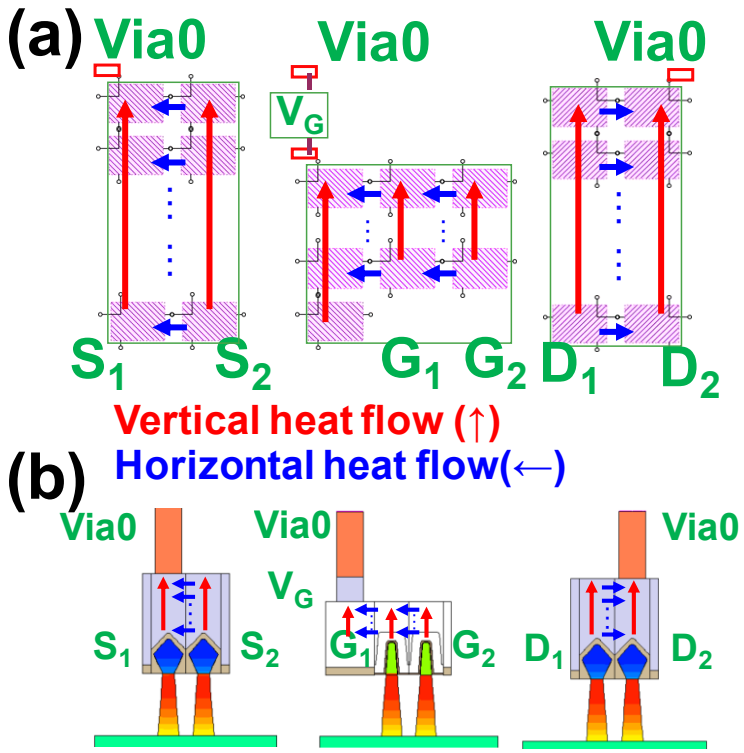


Fig. 4. 7 (a) SPICE modules of source, gate, and drain as well as corresponding MEOLs, and **(b)** TCAD counterparts

4.3.2 Modularized Fin and Power Distribution Matrix

The frequency dependence of $\tau_{hotspot}$ is implemented by charging different amounts of elements with different input frequencies **(Fig. 4. 8 (a))**. The discretized fin is

doi:10.6342/NTU201801788

modeled with device geometry and material thermal conductivities **(Fig. 4. 8 (b))**. The completely modularized fin **(Fig. 4. 8 (c))** is then connected to the MEOL and IMD to form a FinFET.

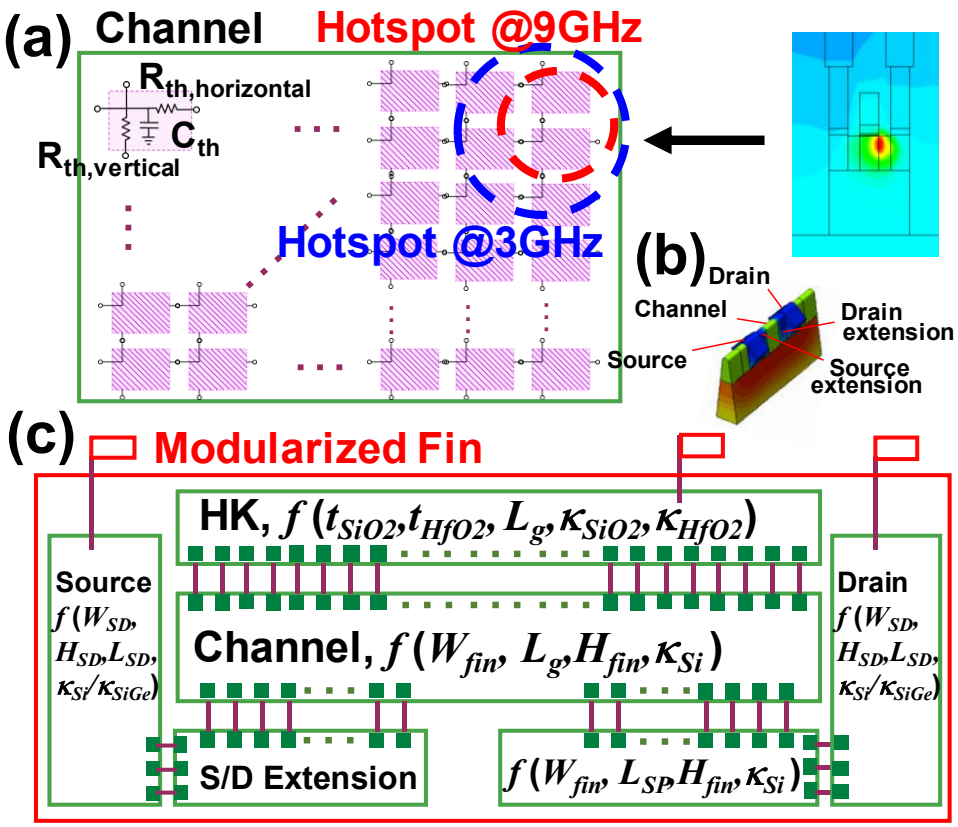


Fig. 4. 8 (a) R_{th}-C_{th} network of channel region with schematically drawn hotspot at 3GHz and 9GHz **(b)** Geometry of fin including the SiGe S/D, and **(c)** the geometry based modularized fin

The Joule heat in device channel is also distributed instead of a point power source. The distribution depends on the applied voltage. However, the Joule heat at V_G=0.5V (**Fig. 4. 9 (a)**) is 0.1X of the Joule heat at V_G=0.7V (**Fig. 4. 9 (b)**).

doi:10.6342/NTU201801788

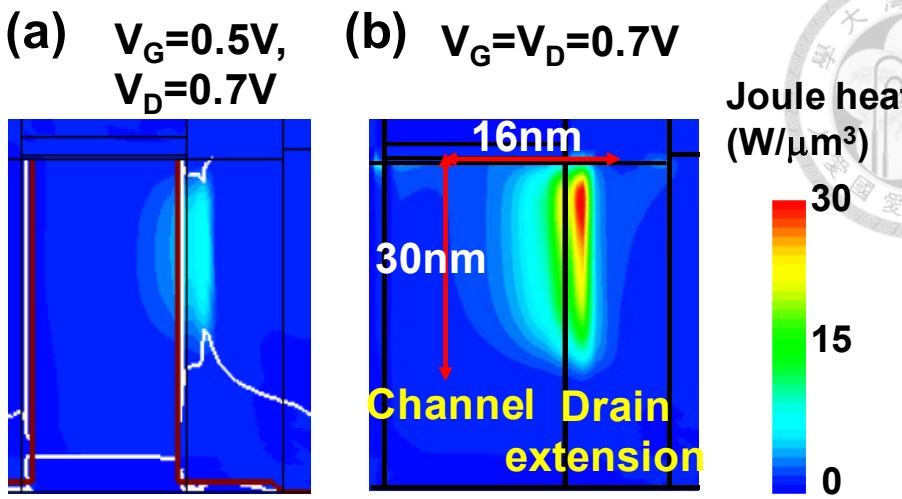


Fig. 4. 9 Joule heat distribution by TCAD at **(a)** V_G=0.5V and V_D=0.7V, and **(b)** V_G=V_D=0.7V

Table 4. 1 shows the extracted power distribution matrix at V_G=V_D=0.7V from TCAD. The total Joule heat is distributed to the nodes in a region with the area of 16nm x 30nm at the interface of channel and drain extension.

Table 4. 1 Extracted nodal power distribution for SPICE

16nm

0.09	0.52	0.85	1.22	2.92	3.77	0.49	0.05
0.21	1.26	2.05	2.93	7.02	9.07	1.17	0.12
0.19	1.13	1.85	2.64	6.33	8.18	1.05	0.11
0.18	1.04	1.69	2.41	5.79	7.48	0.96	0.10
0.15	0.88	1.43	2.04	4.90	6.33	0.82	0.09
0.04	0.26	0.43	0.61	1.46	1.89	0.24	0.03

30nm

doi:10.6342/NTU201801788

4.4 DC Validation of Thermal SPICE Model

4.4.1 Distributed R_{th} Verification

The FinFET SPICE model is verified by TCAD, and the comparison between SPICE and TCAD is shown in **Fig. 4. 10**. The steady state nodal temperature distribution is extracted from a 1-fin-1-finger 14nm pFinFET under face-down configuration with DC bias of V_G=V_D=0.7V. The root-mean-square temperature difference is ΔT_{TCAD}-ΔT_{SPICE}= 1.73°C, and the corresponding error = $\frac{\Delta T_{TCAD}-\Delta T_{SPICE}}{\Delta T_{TCAD}}$ is less than two percent.

The correctness of the R_{th} values in the distributed R_{th}-C_{th} network is validated by the small error.

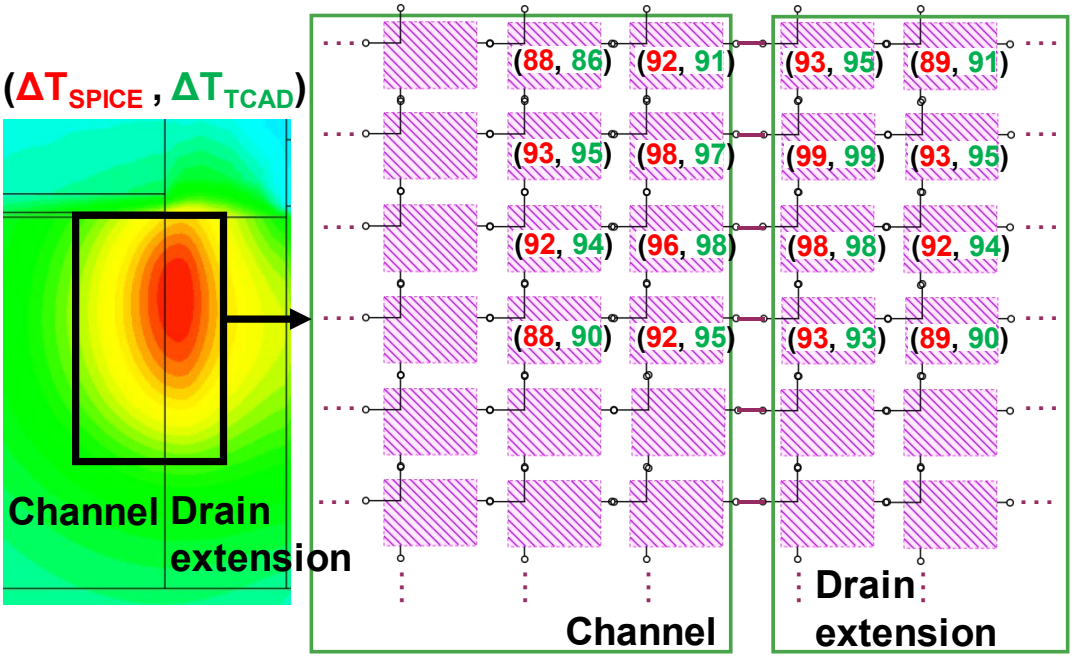


Fig. 4. 10 Nodal temperature near the hot spot with the comparison between SPICE and TCAD

doi:10.6342/NTU201801788

4.4.2 Temperature on MEOL

2D heat flow is adopted for STI and IMD to reduce the number of element and simulation time. The connection between fins and MEOL metals is schematic drawn in **Fig. 4. 11** to form a 2-fin-1-finger pFinFET.

Small temperature difference on the metal line is observed to make sure that ignoring the heat flow in STI and IMD won't affect the precision of temperature mapping both in T_j and interconnects.

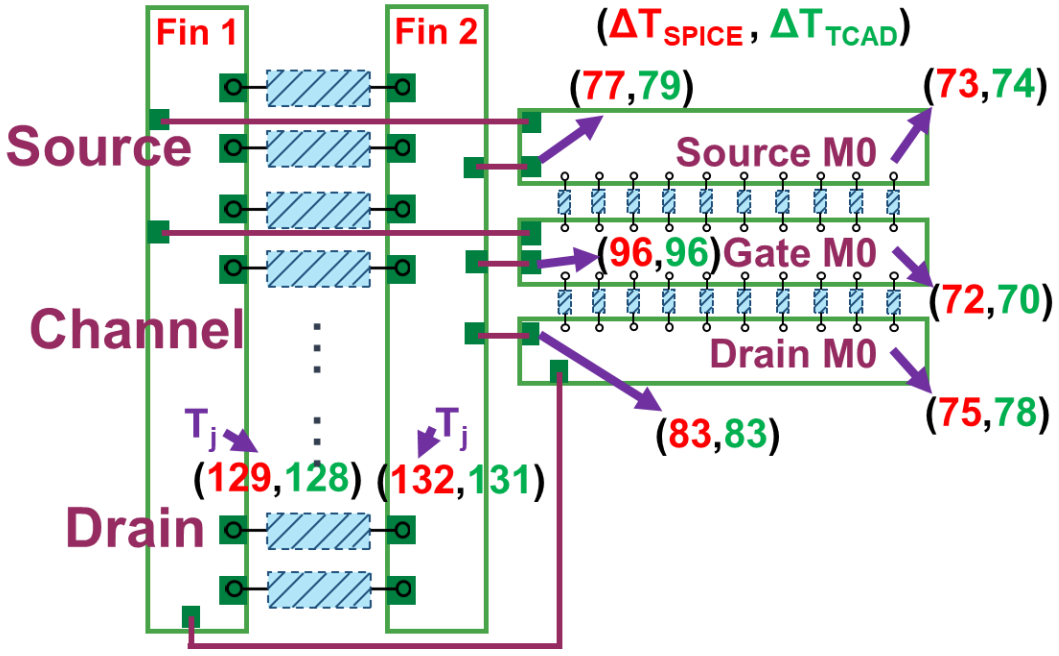


Fig. 4. 11 Nodal temperature on MEOL in the complete SPICE model with the comparison of TCAD

doi:10.6342/NTU201801788

4.4.3 Layout Flexibility of Thermal SPICE Model

To demonstrate the layout flexibility of our thermal SPICE model, multi-fin FinFETs are implemented by using the modularized fins, MEOL, STI, and IMDs based on physical geometry of the FinFETs.

$R_{th0,FinFET}$ is extracted for different fin numbers for both SPICE and TCAD. The extracted intrinsic thermal resistance can be fitted by the same equation. By using the proposed SPICE model with discrete elements for different device layouts, there is no additional TCAD simulation needed for the calibration.

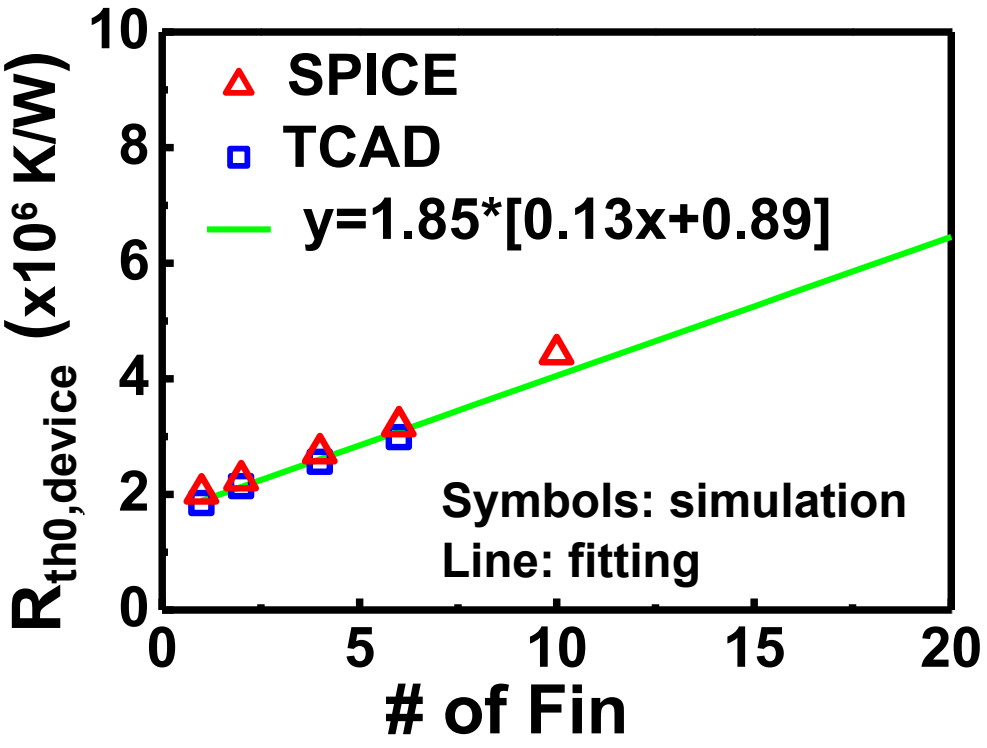


Fig. 4. 12 SPICE and TCAD yielding the similar $R_{th0,FinFET}$ with the same fitting equation for face-down configuration

doi:10.6342/NTU201801788

4.5 Including Interfacial Thermal Resistance into SPICE

The performance of advanced transistors is boosted by high-k materials, SiGe stressor, and silicides…etc. However, phonon mismatch and scattering at material interfaces lead to additional interfacial thermal resistance (ITR).

Table 4. 2 Interfacial thermal conductivity at material interface

Interfacial Thermal Conductivity	
Material interface	**h (GW/K/m^2)**
Si/Ge (Si/SiGe)	**0.33**[4.9]
TiN/W	**0.40**[4.10]
Si/TiN (SiGe/TiN)	**1.10**[4.10]
TiN/SiO_2	**1.03**[4.10]
W/SiO_2	**0.65**[4.10]
Si/SiO_2	**1.10**[4.11]
SiO_2/TiN (HfO_2/TiN)	**1.03**[4.10]

doi:10.6342/NTU201801788

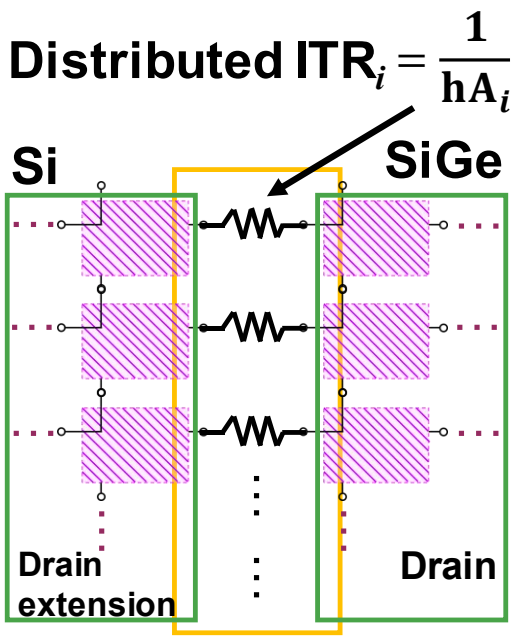


Fig. 4. 13 Distributed interfacial thermal resistance between two different materials in SPICE model

Interfacial thermal conductivities for material interfaces are listed in **Table 4. 2**. The additional ΔT appears at the interface of two different materials due to ITR is adopted into our SPICE model using the value calculated with discretized cross-sectional area, A_i (**Fig. 4. 13**).

ITR effects on the temperature will be discussed in the next chapter using AC input.

4.6 Summary

In this chapter, a SPICE model is proposed and validated by 3-D electro-thermal TCAD. Guideline for determining the element size is provided using the frequency dependence of $\tau_{hotspot}$. To have the SPICE model at an input frequency resolution of

doi:10.6342/NTU201801788

1GHz, the thermal time constant of an element ($\tau_{element}$) should be smaller than 0.71ps. The SPICE model has the layout flexibility to implement FinFETs with different number of fin and number of finger. 3D heat flow and 2D heat flow is considered for fin/metal elements and STI/IMD elements, respectively, to achieve accurate temperature mapping in device channel and to speed up the simulation simultaneously. The ITRs lead to abrupt ΔT at material interfaces, and are considered distributedly in our SPICE.

4.7 References

[4.1] Muhammad Abdul Wahab, SangHoon Shin, and Muhammad Ashraful Alam," 3D modeling of spatio-temporal heat-transport in III-V gate-all-around transistors allows accurate estimation and optimization of nanowire temperature, " *IEEE Transactions on Electron Devices*, Vol. 62, No. 11, Nov. 2015, pp.3595-3604.

[4.2] Brian Swahn and Soha Hassoun, "Electro-Thermal Analysis of Multi-Fin Devices," *IEEE Transactions on very large scale integration (VLSI) systems*, Vol. 16, No. 7, July 2008, pp.816-829.

doi:10.6342/NTU201801788

[4.3] Eric Pop, Robert Dutton, and Kenneth Goodson," Thermal Analysis of Ultra-Thin Body Device Scaling," in *IEEE IEDM Tech. Dig.*, Dec. 2003, pp. 36.6.1-36.6.4.

[4.4] Mayank Shrivastava, Manish Agrawal, Sunny Mahajan, Harald Gossner, Thomas Schulz, Dinesh Kumar Sharma, and V. Ramgopal Rao," Physical Insight Toward Heat Transport and an Improved Electrothermal Modeling Framework for FinFET Architectures," *IEEE Transactions on Electron Devices*, Vol. 59, No. 5, May 2012, pp.1353-1363.

[4.5] Ming-Cheng Cheng, Jeffrey A. Smith, Wangkun Jia, and Ryan Coleman," An Effective Thermal Model for FinFET Structure," *IEEE Transactions on Electron Devices*, Vol. 61, No. 1, Jan. 2014, pp.202-206.

[4.6] BSIM-CMG v110.0.0 manual, p.71, 2015.

[4.7] H. Jiang, L. Shen, S.H. Shin, N. Xu, G. Du, B.-Y. Nguyen, O. Faynot, M.A. Alam, X. Zhang, X. Y. Liu," Unified Self-Heating Effect Model for Advanced Digital and Analog Technology and Thermal-Aware Lifetime Prediction Methodology," in *Proc. Symposium on VLSI Technology*, June 2017, pp. T136–T137.

[4.8] Woojin Ahn, Chunsheng Jiang, Jun Xu, and Muhammad Ashraful Alam, " A New Framework of Physics-Based Compact Model Predicts Reliability of Self-

doi:10.6342/NTU201801788

Heated Modern ICs: FinFET, NWFET, NSHFET Comparison," in *IEEE IEDM Tech. Dig.*, Dec. 2017, pp. 13.6.1-13.6.4.

[4.9] E. S. Landry and A. J. H. McGaughey," Thermal boundary resistance predictions from molecular dynamics simulations and theoretical calculations," *Phys. Rev. B* **80**, 165304 (2009)

[4.10] National institute for material science, Japan. (http://interface.nims.go.jp)

[4.11] Jie Chen, Gang Zhang, and Baowen Li," Thermal contact resistance across nanoscale silicon dioxide and silicon interface, " *Journal of Applied Physics* **112**, 064319 (2012).

doi:10.6342/NTU201801788

Chapter 5 Circuit Level Reliability Evaluation of Ring Oscillators using 10nm FinFETs

5.1 Introduction

Circuit level reliability evaluation and self-heating aware design can be achieved with accurate temperature mapping in devices and interconnects with AC input. The transient SH behavior of scaled FinFETs can only be modeled with distributed R_{th}-C_{th} elements with thermal time constant small enough. The R_{th} values in the distributed R_{th}-C_{th} network of the proposed CTM of FinFETs has been verified in chapter 4.

In this chapter, AC input will be applied to the CTM to check the C_{th} settings. Precise model of $C_{th,BEOL}$ is needed to simulate the transient behavior of the entire circuit. The two-step isothermal plane model for C_{th} will be compared with TCAD results. Transient SHE on ring oscillator (RO) and BEOL will be simulate by our modularized SPICE model.

Cobalt interconnect has the intrinsic benefit of long electro-migration (EM) mean time to failure (MTTF) as compared to Cu. The circuit level Co effects on SHE and corresponding reliability prediction will be demonstrated.

With the accurate SHE simulation and modeling for device and BEOL, innovations for improving the heat dissipation and lowering the SH are proposed.

doi:10.6342/NTU201801788

5.2 AC Validation of Thermal SPICE Model and ITR Effects on Temperature

The validity of C_{th} in the SPICE is ensured by similar transient T_j responses in the first period with the AC input at 3 GHz, 6 GHz, and 9 GHz by TCAD **(Fig. 5. 1)**.

The thermal time constant of an element in the FinFET channel is with the volume of 2nm x 5nm x W_{fin}, which results in a $\tau_{element}$ ~ 0.5nm. The $\tau_{element}$ is confirmed to be small enough to present the frequency dependence of $\tau_{hotspot}$.

Fig. 5. 1 (d) shows the similar envelope of transient T_j evolution at 3GHz by TCAD and our SPICE. The ramping of the lower bound of the envelope is decided by the total C_{th} value of the distributed C_{th} in every element. The correctness of C_{th} from FinFET to MEOL is verified by the simulation results.

Fig. 5. 2 shows the ITR effects on the SH in a 2-fin-1-finger pFinFET with 0.7V square wave input. With ITR, significant ΔT_j of 42°C is observed. Using lumped R_{th}-C_{th} models and device simulation without considering the ITR effect will lead to underestimated device temperature and overestimated lifetime.

doi:10.6342/NTU201801788

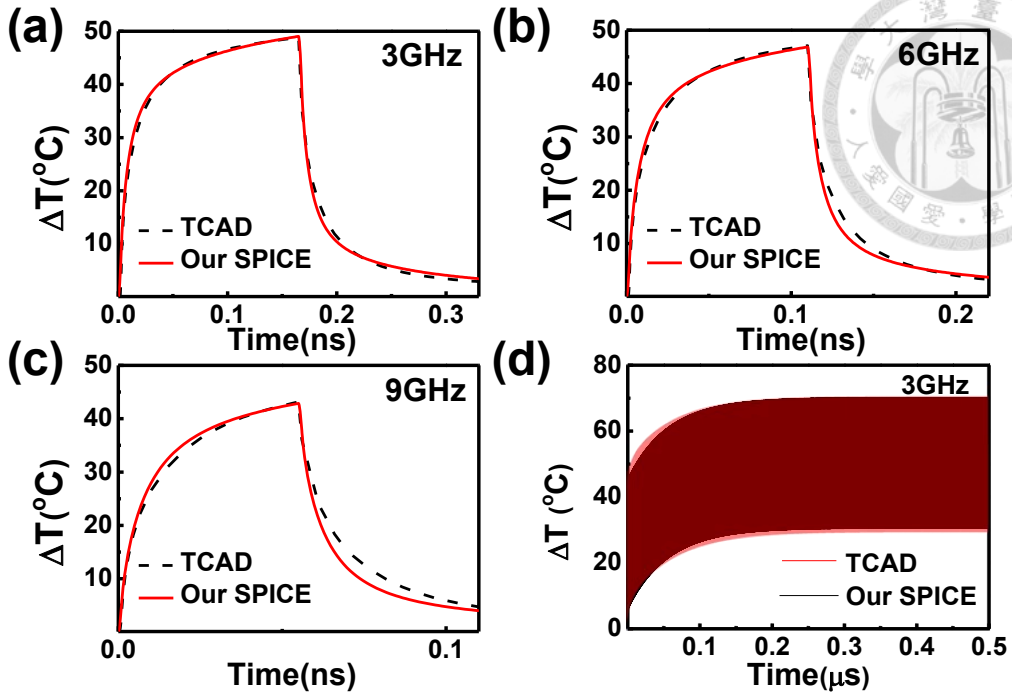


Fig. 5. 1 Transient T_j of a 1-fin-1finger pFinFET in the first period with 0.7V square wave input at (a) 3GHz, (b) 6GHz, (c) 9GHz, and (d) envelope of transient T_j evolution at 3GHz by TCAD and our SPICE

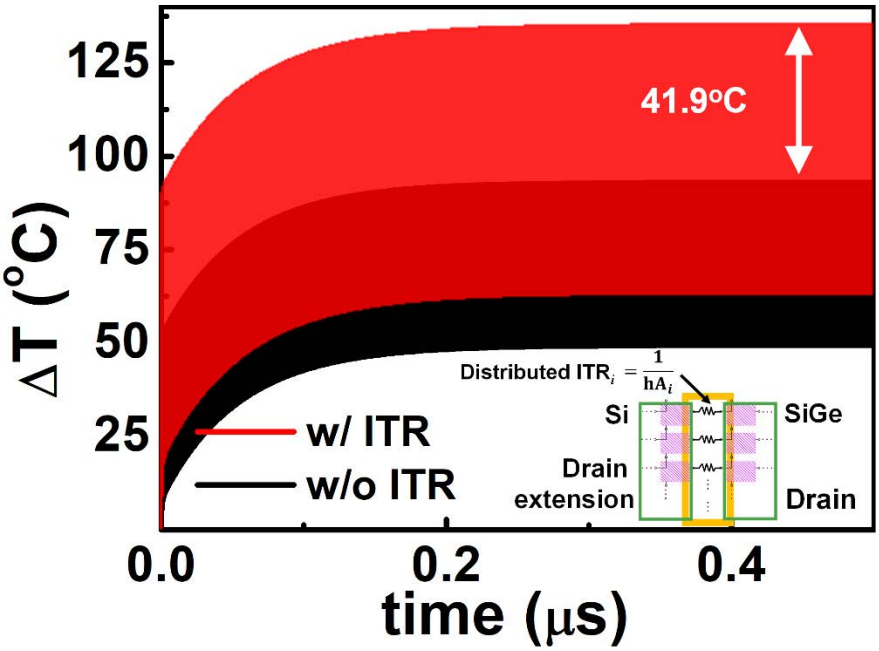


Fig. 5. 2 Envelope of transient T_j evolution at 3GHz with and without ITR in the SPICE model

doi:10.6342/NTU201801788

5.3 Transient Self-heating Effects in FinFETs

5.3.1 Thermal Capacitance Modeling of BEOL

Fig. 5. 3 shows the effective thermal resistance/capacitance circuit including the FinFET, MEOL, BEOL, and substrate. To simulate the transient SHE in FinFET circuits, accurate model of BEOL C_{th} is as important as using CTM consists of distributed R_{th}-C_{th} network for FinFET+MEOL.

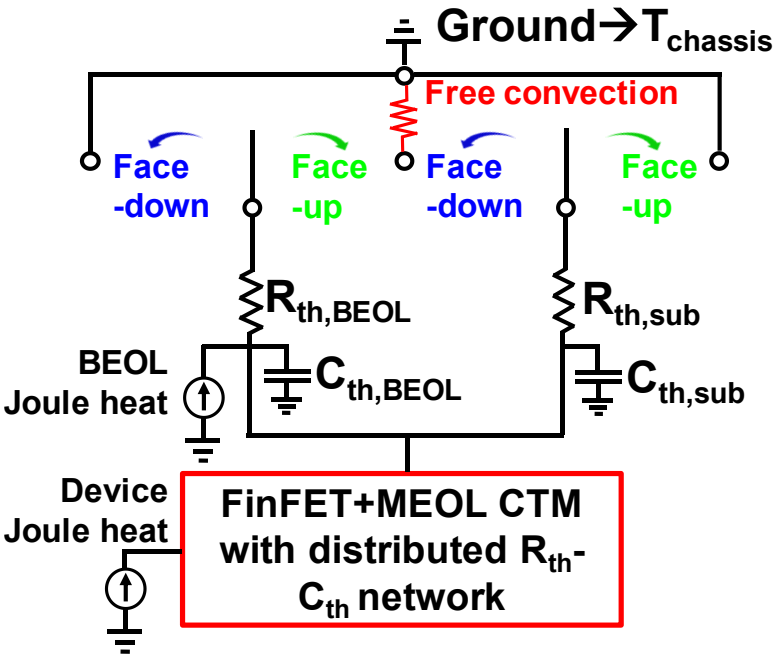


Fig. 5. 3 R_{th}-C_{th} circuit of the FinFET, MEOL, BEOL, and substrate

The effect of non-flat isothermal plane on the $R_{th,BEOL}$ have been modeled using the two-step pseudo isothermal plane in chapter 2. It is extended by including the $C_{th,via}$ and $C_{th,metal}$ calculated layer by layer to form the R_{th}-C_{th} network in BEOL.

doi:10.6342/NTU201801788

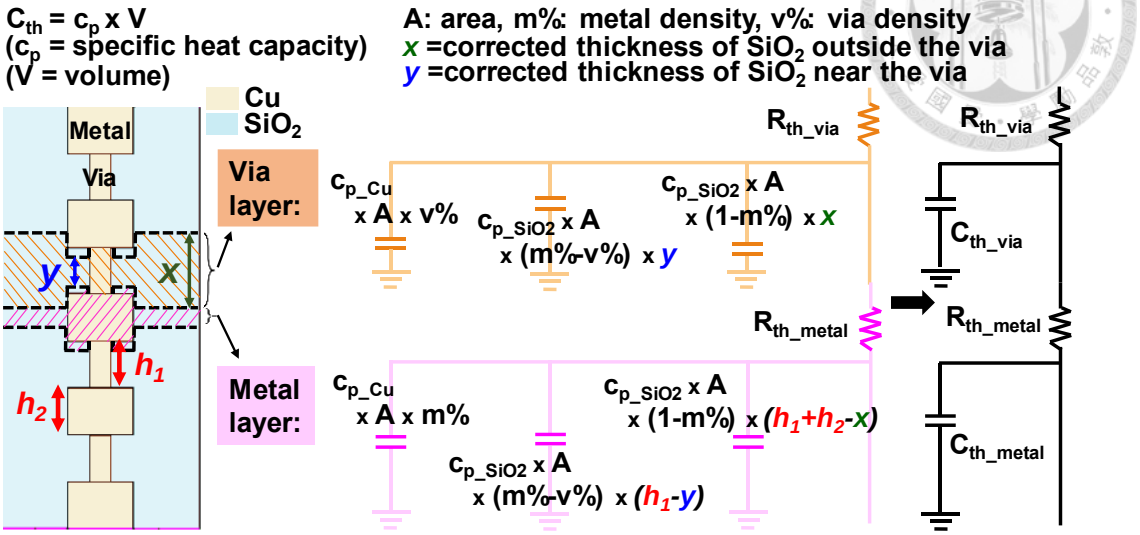


Fig. 5. 4 Extended two-step pseudo isothermal plane model of $R_{th,BEOL}$ and $C_{th,BEOL}$ with the corrected thicknesses of SiO_2

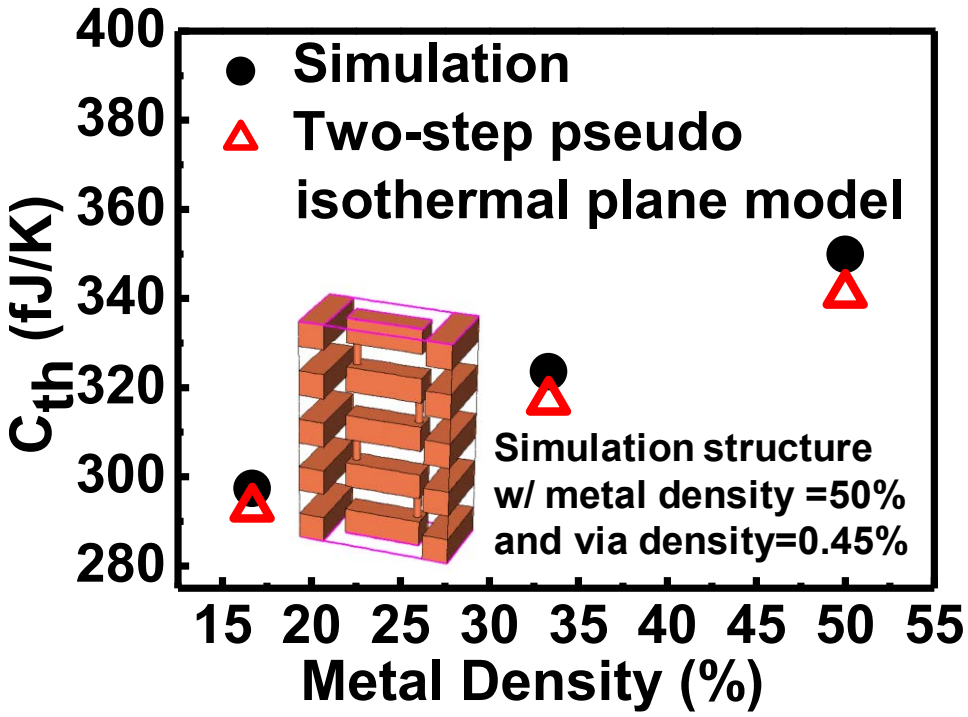


Fig. 5. 5 Modeling $C_{th,BEOL}$ with the comparison of simulation results

doi:10.6342/NTU201801788

Five metal layers with the via density of 0.45% is simulated with metal densities from 17%-50%. The modeled $C_{th,BEOL}$ has been verified by TCAD simulation up to the metal density of 50% (**Fig. 5. 5**) using the same correct oxide thickness of x=0.14μm, and y=0.098μm with the $R_{th,BEOL}$ modeling in 2.3.2.

5.3.2 Thermal Time Constants in FinFET, MEOL, BEOL, and Substrate

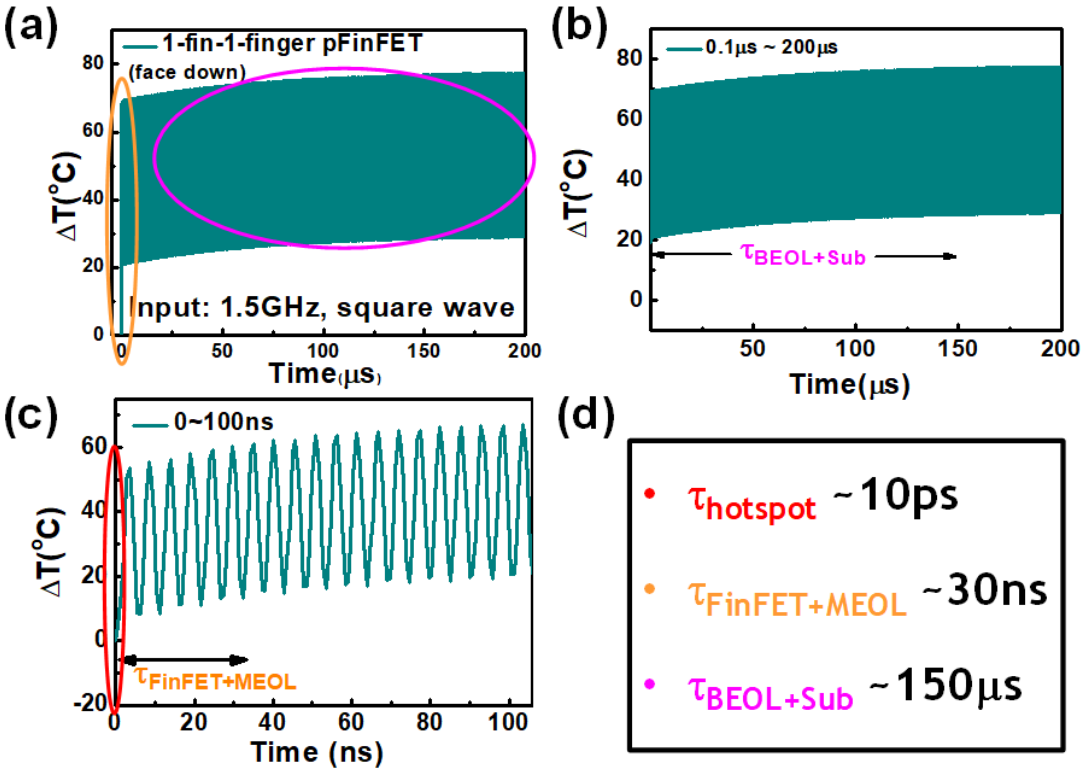


Fig. 5. 6 (a) Envelope of transient T_j evolution at 1.5GHz including the BEOL and substrate to extract thermal time constants of **(b)** BEOL+substrate **(c)** hotspot and FinFET. **(d)** Summarized thermal time constants

doi:10.6342/NTU201801788

Complete R_{th}-C_{th} model of a 1-fin-1-finger pFinFET, MEOL, BEOL, and substrate is simulated with 0.7V square wave input at 1.5 GHz with face-down configuration (**Fig. 5. 6**). A fast ramping of T_j is observed in the beginning as compared to the slow increasing lower bound of the envelope of temperature evolution (**Fig. 5. 6 (a)**). The slow increasing of the lower bound from 0.1μs~200μs reveals the thermal time constant of the BEOL and substrate with the large volume (**Fig. 5. 6 (b)**). The thermal time constant of the hotspot and FinFET+MEOL can be extracted from the 1st period and the fast ramping, respectively (**Fig. 5. 6 (c)**). **Fig. 5. 6 (d)** summarized the extracted thermal time constants. The $\tau_{BEOL+Sub}$ of a 1-fin-1finger FinFET for face-down configuration is ~150 μs, 7 order of magnitude larger than $\tau_{hotspot}$ and 4 order larger than $\tau_{FinFET+MEOL}$ due to the difference in volume.

5.4 Transient Self-heating Effects on Ring Oscillator

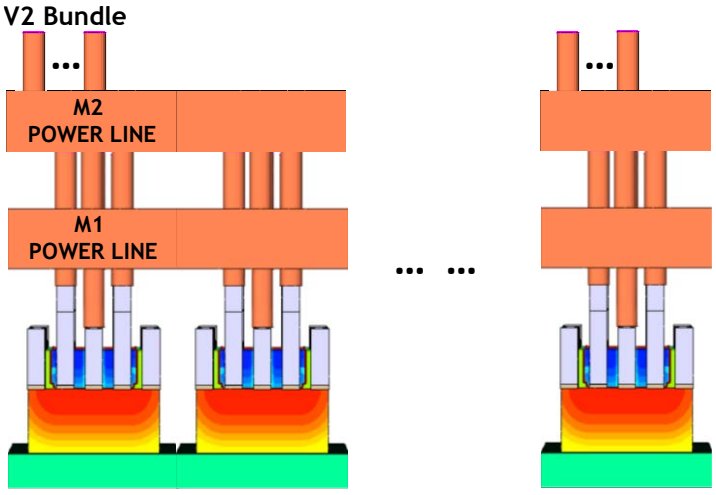


Fig. 5. 7 RO with stacked metal line and V2 bundle

doi:10.6342/NTU201801788

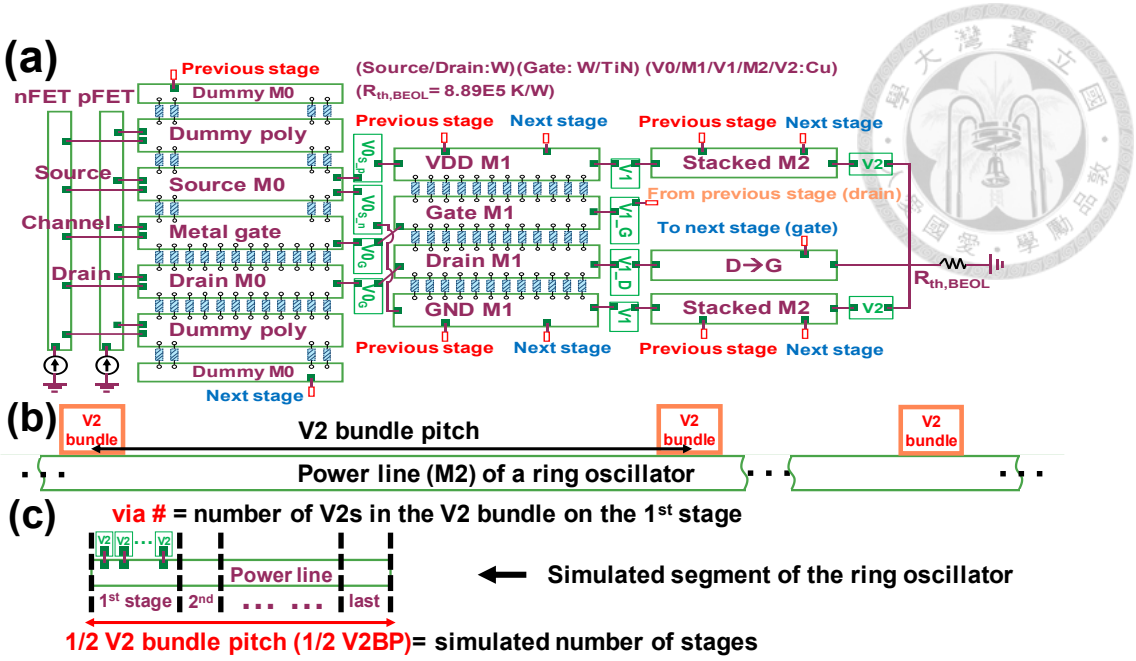


Fig. 5. 8 (a) One stage of the ring oscillator with stacked power line, **(b)** the V2 bundles on the power line of a ring oscillator, and **(c)** simulated segment of the ring oscillator and the definitions of via # and **1/2 V2 bundle pitch (1/2 V2BP)**. Only the first stage has the V2s (see the left end) on the power line as shown in (c)

SPICE module of a stage in a ring oscillator (RO) is shown in **Fig. 5. 8 (a)**. Stacked power lines (M1 and M2, **Fig. 5. 7**) and the inter-stage heat flows are considered. The V2 bundles on the M2 power line of a RO **(Fig. 5. 8 (b))** and the simulated segment of the RO **(Fig. 5. 8 (c))** are schematically drawn with the definitions of via # and 1/2 V2 bundle pitch (1/2 V2BP). A RO can have hundreds of stages. However, only the stages in the 1/2 V2BP are simulated with the consideration of the symmetry of heat flow across the boundaries of the segment.

doi:10.6342/NTU201801788

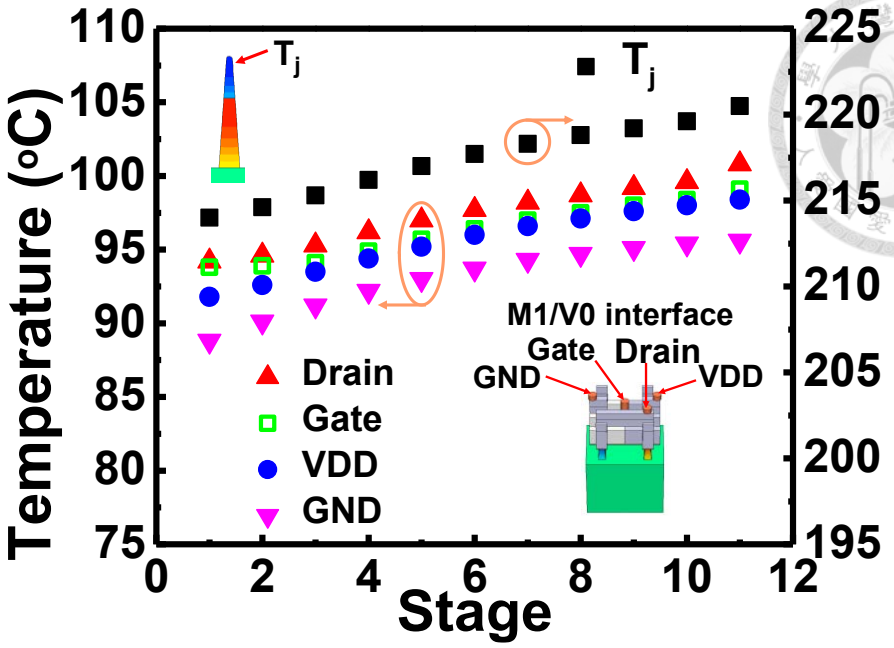


Fig. 5. 9 T_j and the temperature at M1/V0 interfaces for a segment of the RO with 1/2 V2BP of 11 and via # of 1

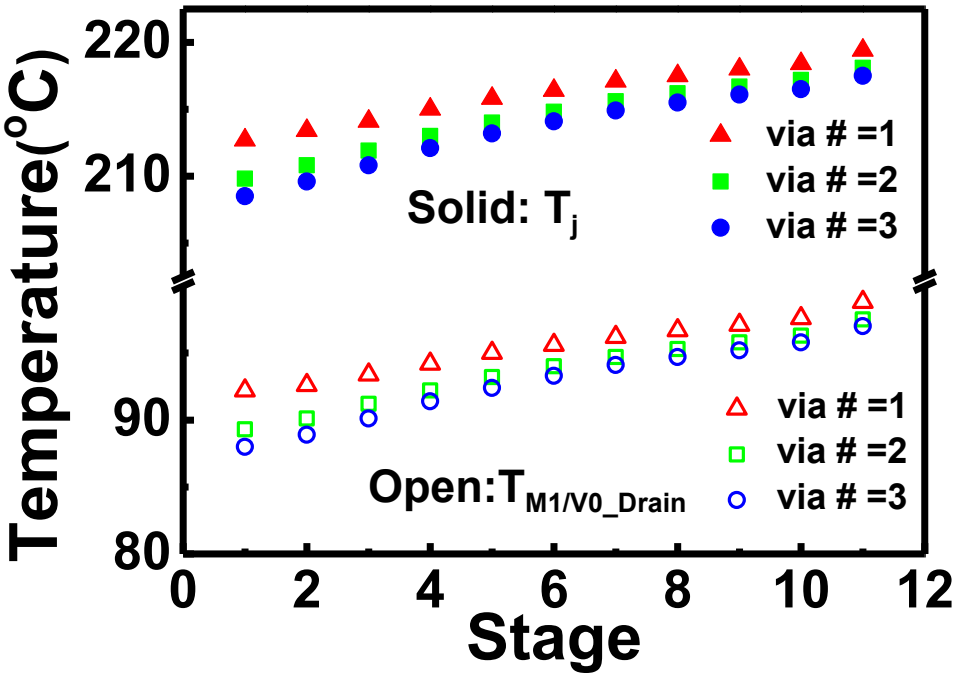


Fig. 5. 10 T_j and the temperature at drain M1/V0 interface for a segment of the RO with 1/2 V2BP of 11 and different via numbers

doi:10.6342/NTU201801788

The highest T_j of 219.4°C is found in the last stage of the simulated segment of the RO (1/2 V2BP =11, via # =1) with $T_{chassis}$=40°C **(Fig. 5. 10)**. The highest nodal temperature in the BEOL is 98.8°C at the M1/V0 interface of the drain **(Fig. 5. 10)**.

SHEs can be eased with the designs of V2 placement (**Fig. 5. 11**). T_j and $T_{M1/V0\_drain}$ in the segment of the RO with 1/2 V2BP =11 are reduced by ~4°C at the 1st stage with the increasing via # from 1 to 2 and 3 in the V2 bundle. However, the nodal temperatures after the 6th stage merely change (<2°C) due to the long distance to the V2 bundle.

5.5 Cobalt Interconnect and Electro-migration Mean Time to Failure Prediction

The segment of a RO (1/2 V2BP =11, via # =1) using pure Co MEOL/V0/M1 is simulated and compared to W MEOL and Cu V0/M1 **(inset of Fig. 5. 11)**. The SHE is deteriorated due to the low κ_{th_Co}. $\Delta T_{j,\ last\ stage}$ =+9.2°C and $\Delta T_{M1/V0\_drain,\ last\ stage}$ = +5.2°C is observed for Co case (**Fig. 5. 11**). The intrinsic EM improvement of Co interconnect as compared to Cu is countervailed.

Reducing the 1/2 V2BP and increasing the via # are effective to improve the worsened SHE **(Fig. 5. 12)**. $\Delta T_{M1/V0\_drain,\ last\ stage}$= -8°C is achieved by the reduction of 1/2 V2BP from 11 to 5. $T_{M1/V0\_drain,\ last\ stage}$ can be further improved by -2.6°C with the increasing via # from 1 to 2.

doi:10.6342/NTU201801788

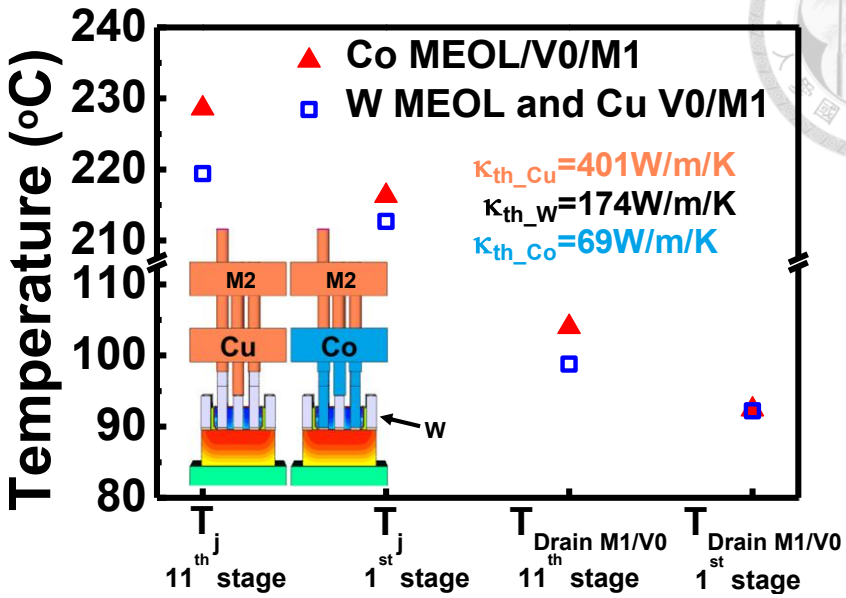


Fig. 5. 11 Comparison of T_j and the temperature at drain M1/V0 interface for pure Co or W MEOL and Cu M1

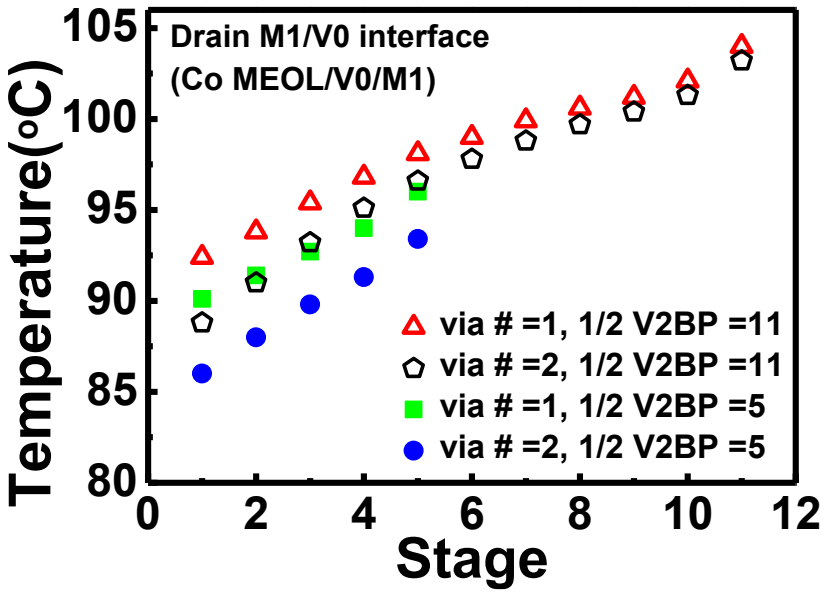


Fig. 5. 12 Temperature at drain M1/V0 interface for segments of the ROs using Co MEOL and M1 with different stage numbers and via numbers

doi:10.6342/NTU201801788

The EM MTTF is modeled by Black's empirical equation [5.1]:

$$\text{EM MTTF} = \text{A}/J^n \times \exp(E_a/kT) \quad (5.1)$$

The EM MTTF of Co technology is reportedly at least 5X longer than Cu technology [5.2]. E_{a_Co} is calculated to be 0.9416eV with respect to E_{a_Cu}=0.9eV for EM MTTF prediction. The intrinsic EM benefits of Co interconnects is retarded to be 2.44X by the increasing temperature in BEOL as compared to W MEOL and Cu V0/M1 **(Fig. 5. 13)**. The EM MTTF can be improved by circuit design. By the reduction of 1/2 V2BP from 11 to 5, the MTTF is extended to 4.58X. An EM MTTF improvement of 5.64X with respect to W MEOL and Cu V0/M1 is achieved by increasing the via # from 1 to 2 with SHE considered.

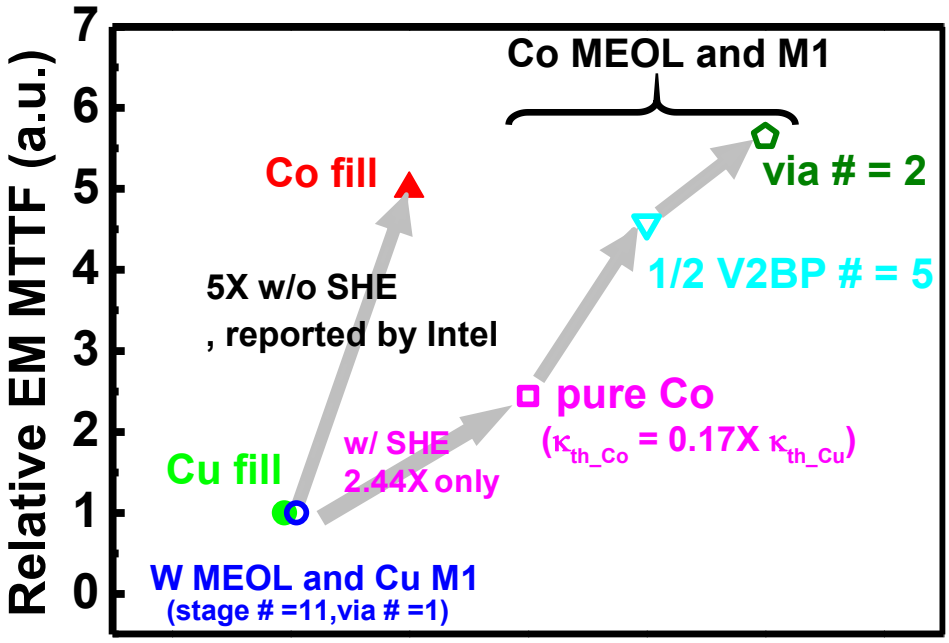


Fig. 5. 13 EM MTTF prediction based on different designs as compared to Cu technology

doi:10.6342/NTU201801788

5.6 Innovations to Improve Heat Dissipation of Device and BEOL

The improvement of EM MTTF achieved by circuit design is demonstrated in 5.5. Lowering the temperature in device and interconnect can slow down the degradation and extend the lifetime. Two designs of device and BEOL are proposed to enhance the heat dissipation in the following parts.

5.6.1 Additional Via0 to Lower Intrinsic Thermal Resistance

The $R_{th0,FinFET}$ is found linearly dependent to the number of fin and saturates with the increasing number of finger in multi-fin and multi-finger FinFETs, respectively.

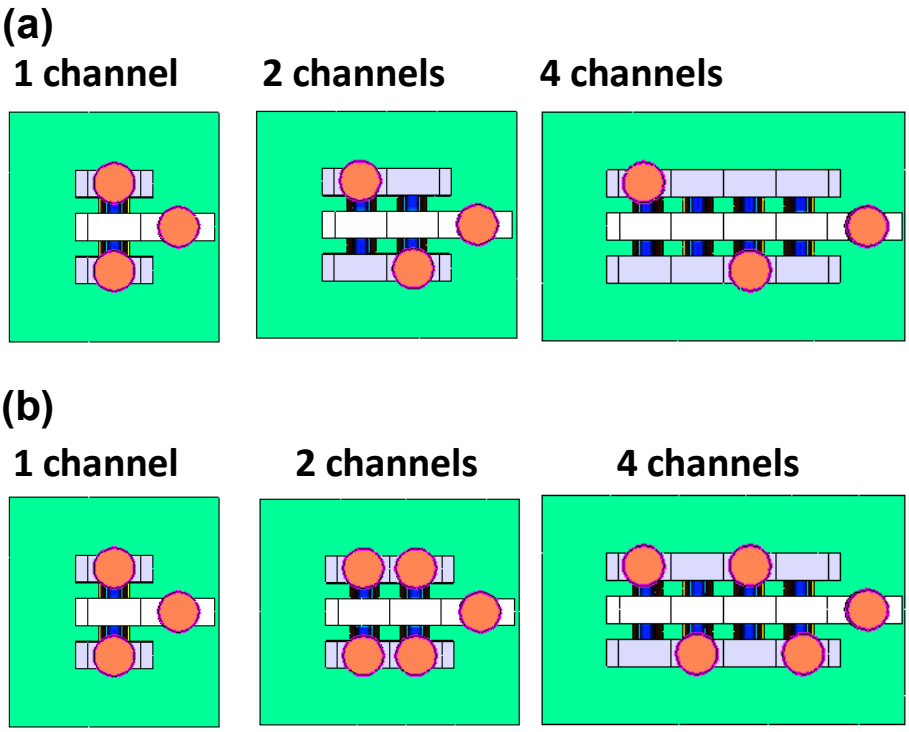


Fig. 5. 14 Top view of multi-fin FinFETs (a) remain three via0s and (b) adding additional via0s

doi:10.6342/NTU201801788

Fig. 5. 14 (a) shows the top views of multi-fin FinFETs. The contact via0 remains three with the increasing number of fin, different from multi-finger FinFETs. The total heat increases with the increasing fin number, and the via0s are the bottleneck of heat dissipation. It is possible to add additional via0s on the source contact and drain contact without the violation of design rules (**Fig. 5. 14 (b)**).

Fig. 5. 15 shows the $R_{th0,FinFET}$ with and without additional via0s. By adding additional via0s, the $R_{th0,FinFET}$ of multi-fin FinFETs saturates with the increasing number of fin, similar to the $R_{th0,FinFET}$ of multi-finger FinFETs with the increasing finger number.

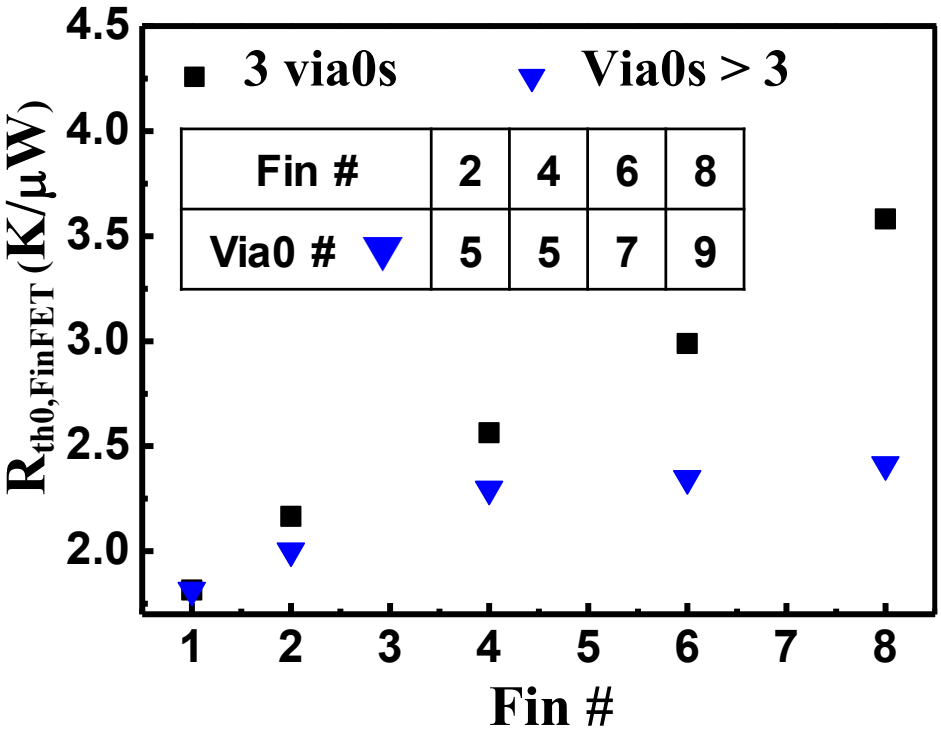


Fig. 5. 15 Reduced $R_{th0,FinFET}$ with the increasing fin # by additional via0s

doi:10.6342/NTU201801788

5.6.2 Highly Thermally Conductive Via in BEOL

In face-down configuration, most of the heat dissipates through the BEOL. The via layers are the bottleneck of heat dissipation, due to its low metal density as compared to metal layers.

Fig. 5. 16 shows the calculated $R_{th,BEOL}$ with different via density using two-step pseudo isothermal plane modeling. 40% $R_{th,BEOL}$ reduction is achieved by increasing the via density from 2% to 5%. However, it will violate the design rules to further increase the via density above 5%. The constrain of via density is due to the difference of material CTE between Cu and SiO_2 (**Table 5. 1**). The stress induced by Cu vias leads to reliability issues in BEOL.

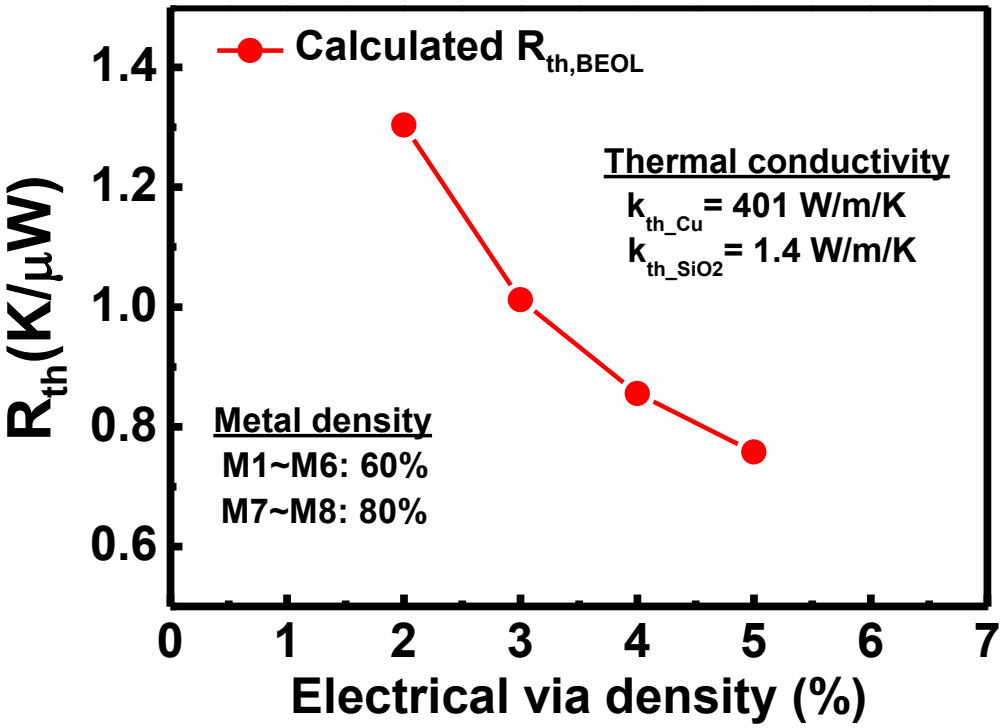


Fig. 5. 16 $R_{th,BEOL}$ with different via densities

doi:10.6342/NTU201801788

Table 5. 1 Thermal conductivity and coefficient of thermal expansion of materials

Material	k_{th}(W/m/K)	CTE(ppm/K)
Cu	401	18
SiO_2	1.4	0.6
Si	148	2.8
Diamond	2000	1.0
AlN	285	5.3
BeO	330	7.4~8.9

Dielectrics such as the diamond, AlN, and BeO are highly thermally conductive (**Table 5. 1**). CTE of these materials are smaller than that of Cu and will add less stress if they're implemented in the BEOL.

Fig. 5. 17 shows the reduced $R_{th,BEOL}$ with the incorporation of BeO thermal vias. The $R_{th,BEOL}$ with 2% Cu via can be reduced by 0.42X with 10% thermal vias. The simulated ΔT_j of a 4-fin-1-finger pFinFET can be dramatically reduced from 245°C to 155°C.

doi:10.6342/NTU201801788

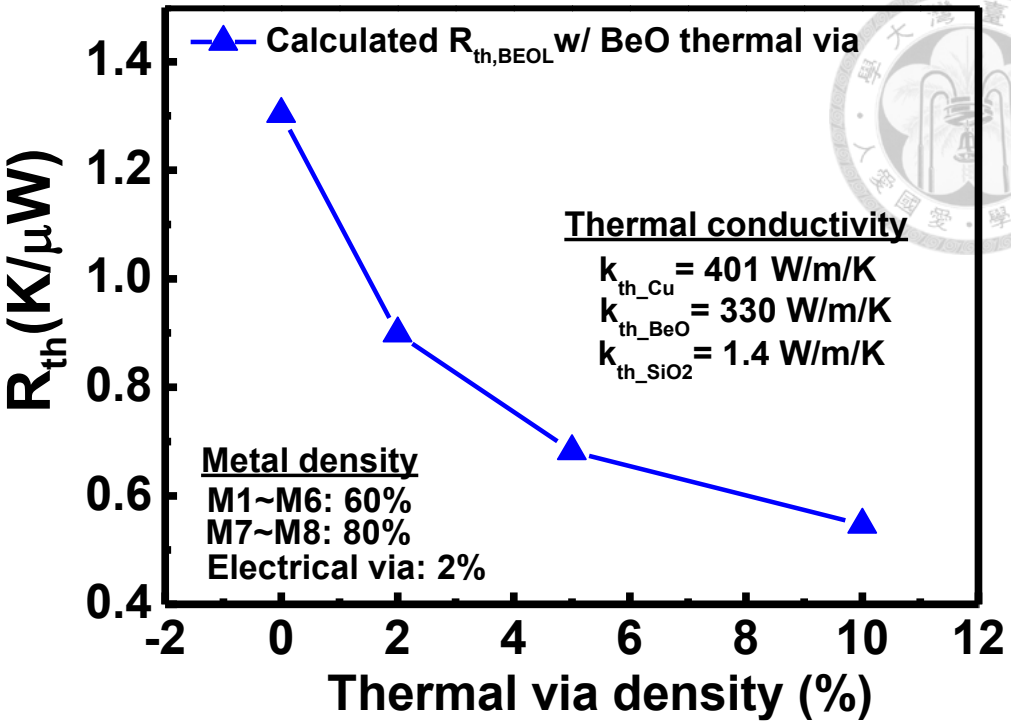


Fig. 5. 17 $R_{th,BEOL}$ with different thermal via densities

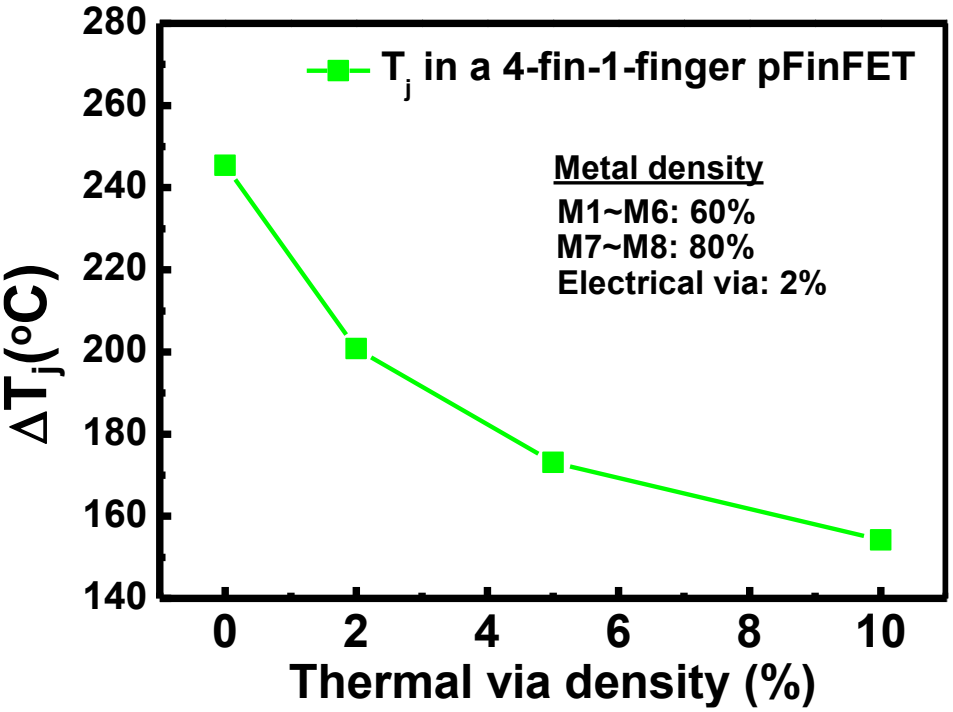


Fig. 5. 18 Reduced T_j with increasing thermal via densities

doi:10.6342/NTU201801788

5.7 Summary

The C_{th} values in the SPICE affect the precision of transient SHE simulation in FinFETs. The correctness has been verified by the similar results with device simulation. ITR at material interfaces lead to an abrupt temperature change, and must be included in the SPICE in case of underestimating the T_j. Accurate $C_{th,BEOL}$ modeling is also essential for the transient SHE simulation, and the validity of extended two-step pseudo isothermal plane model is confirmed by TCAD. The BEOL and substrate has a large thermal time constant as compared to FEOL FinFETs due to the large volume.

The intrinsic electro-migration (EM) improvement of Co interconnect (5X) is countervailed (5X→2.44X) by the increasing T_{metal} due to the low thermal conductivity of Co. Different V2 placements on the power line of a RO are proposed to lower both the T_j (FinFET) and T_{metal}. The predicted EM MTTF of Co interconnect with the additional heat dissipation by V2 insertion is ~5.65X of W/Cu interconnect.

Adding additional via0s for multi-fin FinFETs and applying highly thermally conductive vias in BEOL can help to dissipate the Joule heat from devices. The reliability can be improved by lowering the temperature. SHE-aware circuit designs can be conducted with our SPICE to efficiently achieve the robust reliability from the device to circuit level.

doi:10.6342/NTU201801788

5.8 References

[5.1] James R. Black, " Current Limitations of thin Film Conductors," *Reliability Physics Symposium (IRPS)*, *1982 IEEE International*, pp. 300- 306.

[5.2] C. Auth, A. Aliyarukunju, M. Asoro, D. Bergstrom, V. Bhagwat, J. Birdsall, N. Bisnik, M. Buehler, V. Chikarmane, G. Ding, Q. Fu, H. Gomez, W. Han, D. Hanken, M. Haran, M. Hattendorf, R. Heussner, H. Hiramatsu, B. Ho, S. Jaloviar, I. Jin, S. Joshi, S. Kirby, S. Kosaraju, H. Kothari, G. Leatherman, K. Lee, J. Leib, A. Madhavan, K. Marla, H. Meyer, T. Mule, C. Parker, S. Parthasarathy, C. Pelto, L. Pipes, I. Post, M. Prince, A. Rahman, S. Rajamani, A. Saha, J. Dacuna Santos, M. Sharma, V. Sharma, J. Shin, P. Sinha, P. Smith, M. Sprinkle, A. St. Amour, C. Staus, R. Suri, D. Towner, A. Tripathi, A. Tura, C. Ward, and A. Yeoh, " A 10nm High Performance and Low-Power CMOS Technology Featuring 3rd Generation FinFET Transistors, Self-Aligned Quad Patterning, Contact over Active Gate and Cobalt Local Interconnects," in *IEEE IEDM Tech. Dig.*, Dec. 2016, pp. 29.1.1-29.1.4.

doi:10.6342/NTU201801788

Chapter 6 Asymmetric Keep-out Zone Modeling of Through-Silicon Via with Experimental Verification

6.1 Introduction

The coefficient of thermal expansion (CTE) mismatch of the materials in through-silicon via (TSV) leads to a stress field on silicon after process. An additional stress field is introduced, and adds on strain in nearby transistor and changes the device current. The effect of via-middle TSV proximity on different CMOS technology nodes were reported [6.1] - [6.5]. No keep-out zone (KOZ) was observed [6.1] - [6.4] for both nFETs and pFETs under the criterion of $\Delta I_{on} < 5\%$, probably due to the small TSV diameters. The TSV proximity reported in [6.5] induced a 30% I_{on} change for pFETs only. The impact of via-last TSV was investigated using ring-oscillator measurement [6.6], a KOZ size of 2μm was reported. Analytic model of symmetric KOZ, i.e. the absolute value of radial stress equals to tangential stress ($|\sigma_r| = |\sigma_\theta|$), has been proposed [6.7] based on the Lame equation.

In this chapter, TSV induced stress field and device performance variation are investigated using N28 silicon data. Back-end-of-line (BEOL) dielectrics affects the stress field of via-last TSV. Asymmetric KOZ is observed, i.e. $|\sigma_r|$ is different from that of $|\sigma_\theta|$. 3D finite-element analysis (3D-FEA) is used to simulate the stress distribution. The physics behind the asymmetry is described. A KOZ model considering the asymmetric stress field is proposed and verified by experiment data.

doi:10.6342/NTU201801788

6.2 TSV and Device Layout

The via-last TSVs are fabricated with the Cu diameter of 5μm, oxide thickness of 0.5μm, and the height of 50μm (**Fig. 6. 1**). The fabricated device channels are along [110] direction. The devices with channel along the radial direction and tangential direction of TSV are measured. The distance from device to the center of TSV varies from 9.5μm to 19.5μm

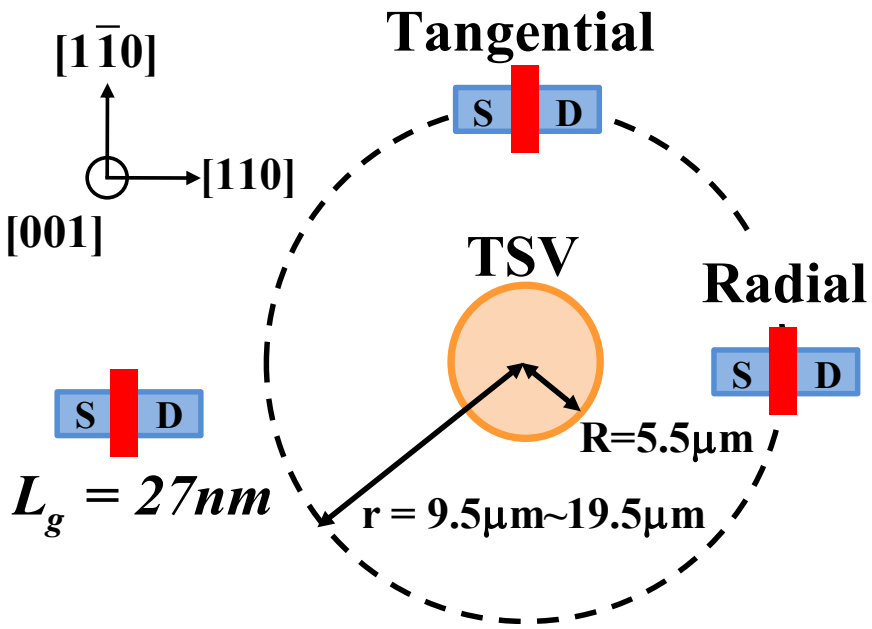


Fig. 6. 1 TSV and device placement

The active regions are surrounded by shallow trench isolation (STI), and the area ratio of the active region to the total area is about 15% (**Fig. 6. 2**). The active regions are located beside the TSV and formed an STI pattern. The TSV is fabricated after the BEOL, known as the via-last process. Measurements are conducted at room temperature.

doi:10.6342/NTU201801788

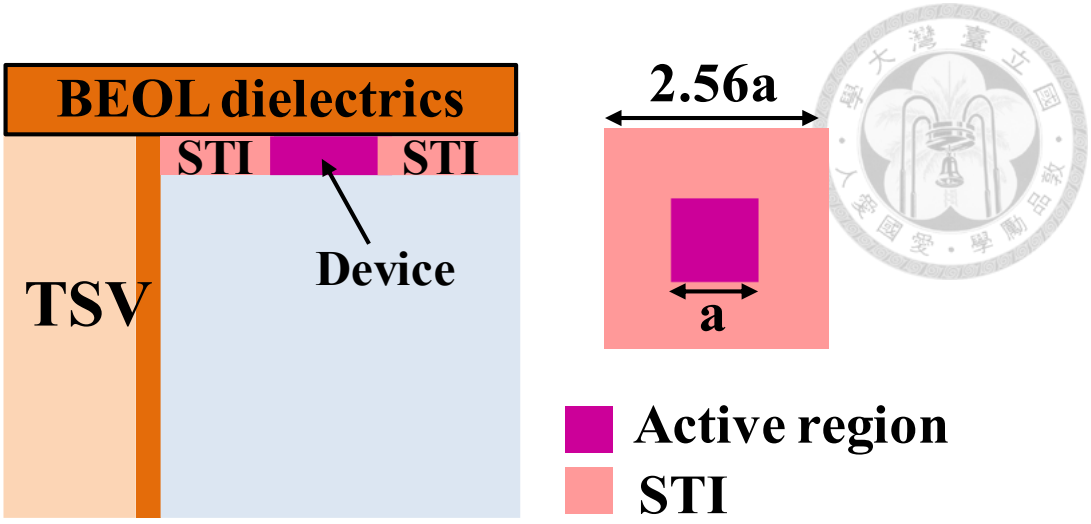


Fig. 6. 2 Side view of TSV and device, and area ratio of active region

6.3 3D Finite-element Analysis Simulation

6.3.1 Simulation Settings

It is important to consider the 3D effects [6.8],[6.9] on Si surface to achieve accurate stress field. 3D FEA tool, Ansys., is adopted to simulate the stress around the TSVs. **Fig. 6. 3 (b)** shows the simulated structure. The TSV structure is with some symmetric planes (**Fig. 6. 3 (a)**), i.e. zero displacement along the normal vector during the ramping of temperature. Only one-eighth of the structure is simulated based on the symmetry, and the efficiency of simulation is enhanced.

Material properties used in Ansys. are listed in **Table 6. 1**. The copper and SiO_2 are considered to be isotropic, the material CTE, Young's modulus, shear modulus (k) and Poisson's ratio (v) were reported in [6.8]. The silicon anisotropy of silicon is

doi:10.6342/NTU201801788

considered. Stiffness of Si are 165.7, 63.9, and 79.6 GPa for C11, C12, and C44, respectively. To represents the stress from the BEOL, a 1μm thick dielectric (SiO_2) is set on the top surface of TSV, Si and STI. The temperature ramps from 200°C to 27 °C. The mechanical stresses induced by the mismatch of material CTEs are simulated at 27 °C.

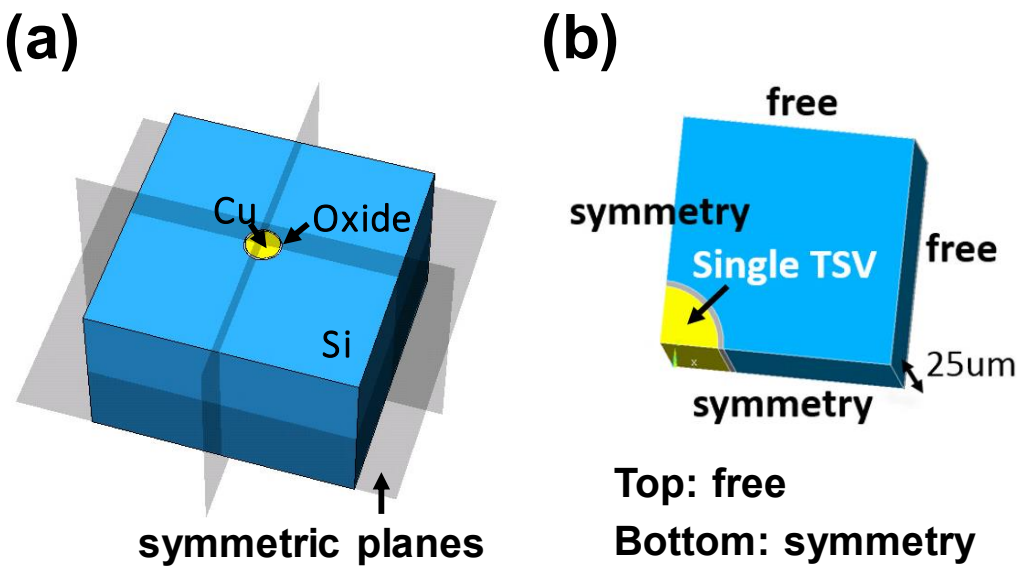


Fig. 6. 3 (a) Structure of single TSV with symmetric planes **(b)** simplified simulation structure with regard to the symmetric planes

Table 6. 1 Material properties used in Ansys.

Material	Poisson's Ratio (v)	Shear Modulus (k)	Young's Modulus	CTE (ppm/°C)
Silicon	**Anisotropic**			**2.8**
SiO_2	**0.16**	**30 GPa**	**69.6 GPa**	**0.6**
Copper	**0.34**	**26 GPa**	**120 GPa**	**18.0**

doi:10.6342/NTU201801788

6.3.2 STI Pattern Effects

Fig. 6. 4 schematically shows the STI pattern, and is included in simulation. A compressive stress is induced by STI in the active regions. This stress should be eliminated both in the experiments and simulations. Devices with the same STI pattern and without the TSV are fabricated and measured. The STI stress is deducted from the simulated stress of Ansys..

To confirm the layout uniformity, TSV with the STI pattern but without the BEOL dielectrics is simulated. The simulated stress is extracted from different directions (0°, 45°, and 90°). The extracted stress in active region can be fitted by one analytic solution for different direction (**Fig. 6. 5**), verifying there is no directional dependence of the STI pattern.

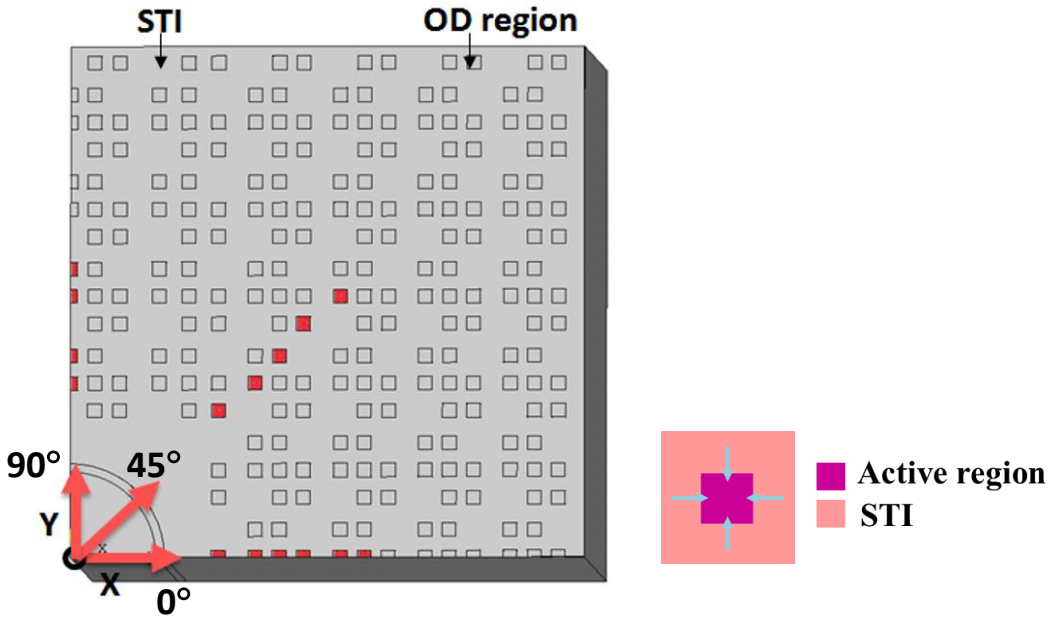


Fig. 6. 4 STI pattern and devices in different direction

doi:10.6342/NTU201801788

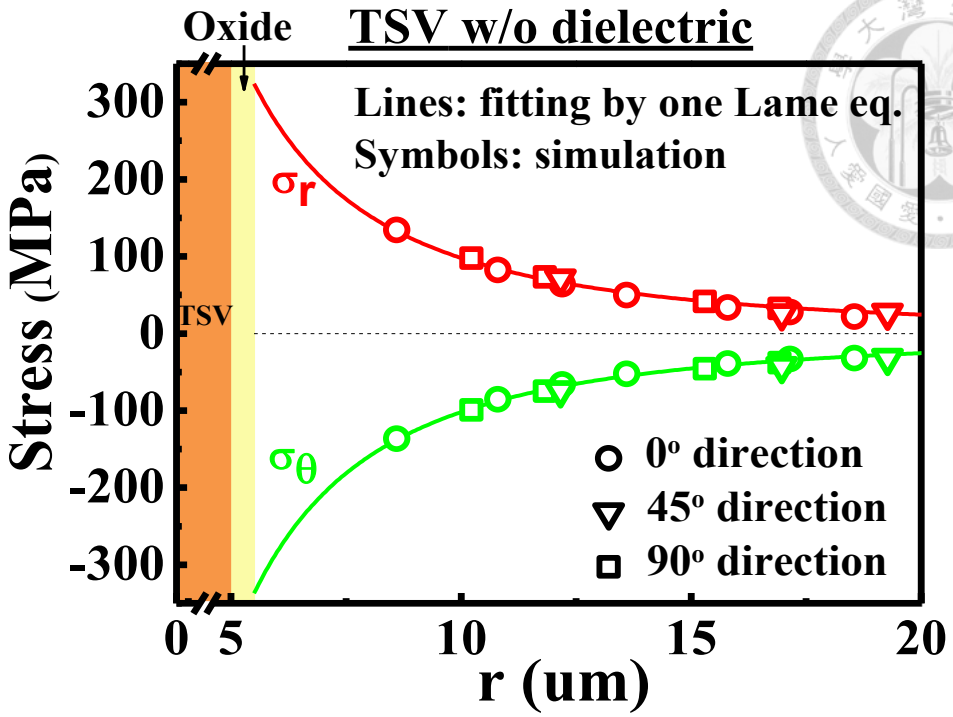


Fig. 6. 5 Simulated stress in devices which induced by a TSV w/ STI pattern but w/o the BEOL dielectrics

6.4 Modeling of Stress Field on Silicon Surface

6.4.1 BEOL Dielectric Effects on the Stress Field

An asymmetric stress field ($|\sigma_r| \neq |\sigma_\theta|$) on silicon is observed (**Fig. 6. 6**). The Cu shrinkage from 200°C to 27°C is retarded by 1 μm BEOL dielectrics on the TSV (inset of **Fig. 6. 6**), leading to the reduction of TSV proximity as compared to the stress field without the BEOL dielectric.

Note that the $|\sigma_\theta|$ is larger than $|\sigma_r|$, the origin of the asymmetric stress field is discussed below.

doi:10.6342/NTU201801788

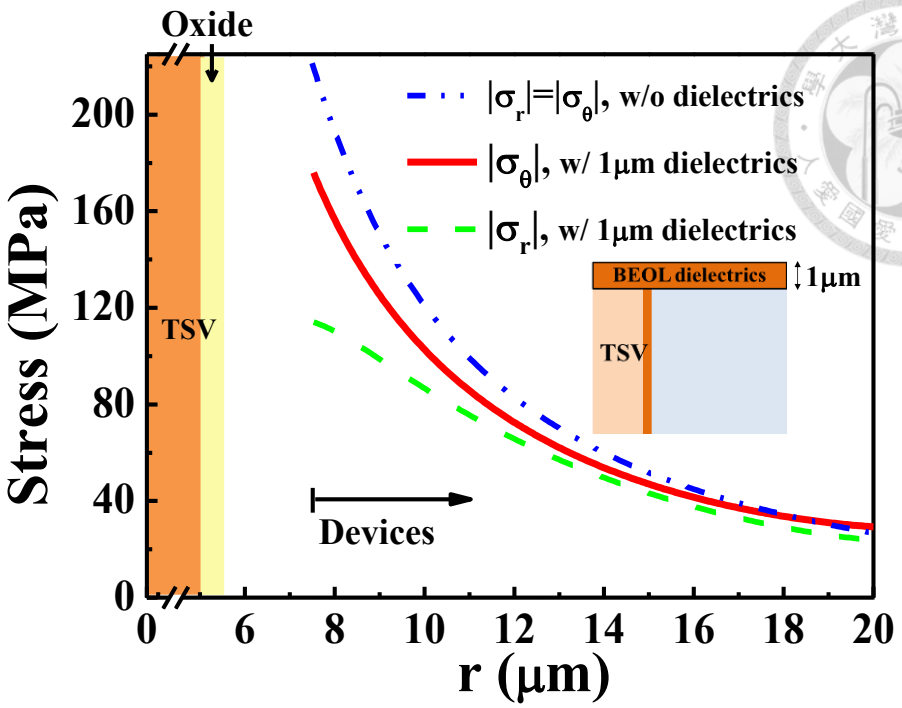


Fig. 6. 6 The asymmetric stress field by BEOL dielectrics **(inset)** simulated TSV with the BEOL dielectric

6.4.2 Modeling the asymmetric stress field

The stress on Si can be obtained by the radial displacement vector ($u_{r,\,Si}$).

The dominant term of $u_{r,\,Si}$ can be simplified as:

$$u_{r,Si}(r) \approx a \times r^{n},\ a = constant \tag{6.1}$$

The corresponding elastic strain (e) can be obtained as the following.

$$e_r = \frac{du}{dr} \approx na \times r^{n-1}, e_\theta = \frac{u}{r} \approx a \times r^{n-1} \tag{6.2}$$

The corresponding stresses are:

$$\sigma_r = 2ke_r + \lambda \times (e_r + e_\theta) \approx 2ka \times (1.64n + 0.64) \times r^{n-1} \tag{6.3}$$

and

$$\sigma_\theta = 2ke_\theta + \lambda \times (e_r + e_\theta) \approx 2ka \times (0.64n + 1.64) \times r^{n-1} \tag{6.4}$$

doi:10.6342/NTU201801788

for radial direction and tangential directions, respectively.

Where $\lambda = 2k \times \nu \times (1-2\nu)^{-1} \approx 0.64 \times 2k$ for silicon.

The value of n equals -1 for a TSV w/o overlaying BEOL dielectrics [6.7], and $|\sigma_r| = |\sigma_\theta|$ is derived.

The value of n is extracted to be -0.7 for a TSV with 1 μm overlaying dielectrics on the top by FEA simulation (**Fig. 6. 7**).

The corresponding stresses are:

$$\sigma_r \approx -0.51 \times 2ka \times r^{-1.7} \tag{6.5}$$

$$\sigma_\theta \approx 1.19 \times 2ka \times r^{-1.7} \tag{6.6}$$

The $|\sigma_r| < |\sigma_\theta|$ is consistent with the simulated asymmetric stress field.

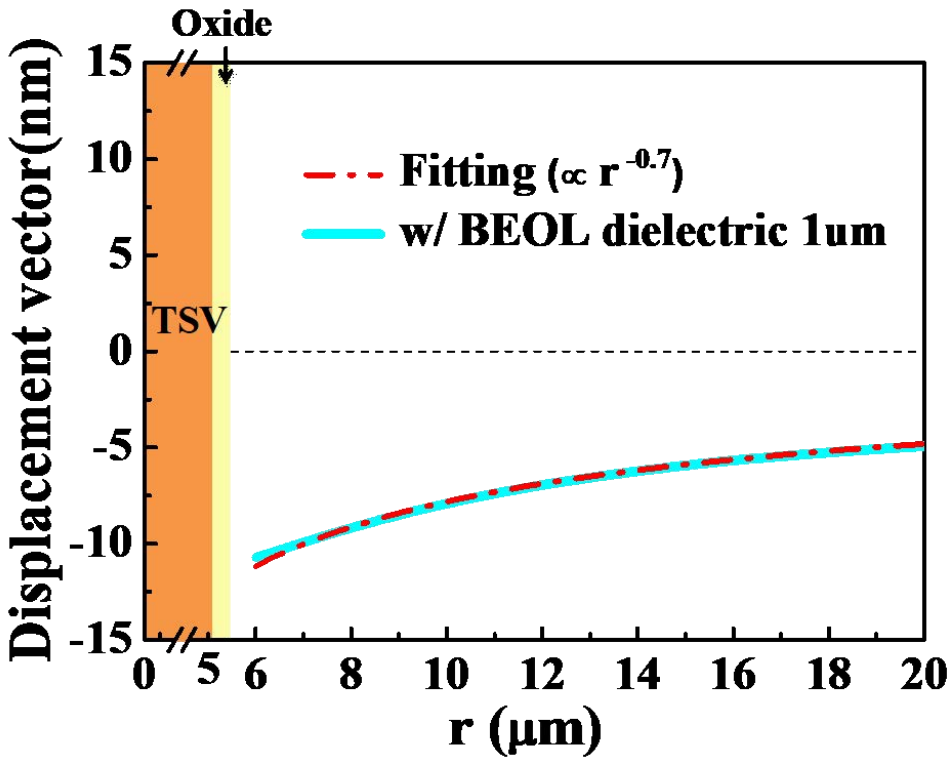


Fig. 6. 7 Simulated displacement vector of a TSV with 1μm BEOL dielectrics

doi:10.6342/NTU201801788

The stress induced by STI in the active regions has been removed. However, SiO_2 with the CTE smaller than Si might retard the shrinkage of Cu pillar on the surface, and the simulated stress is about 8% lower with STI pattern as compared to that without STI.

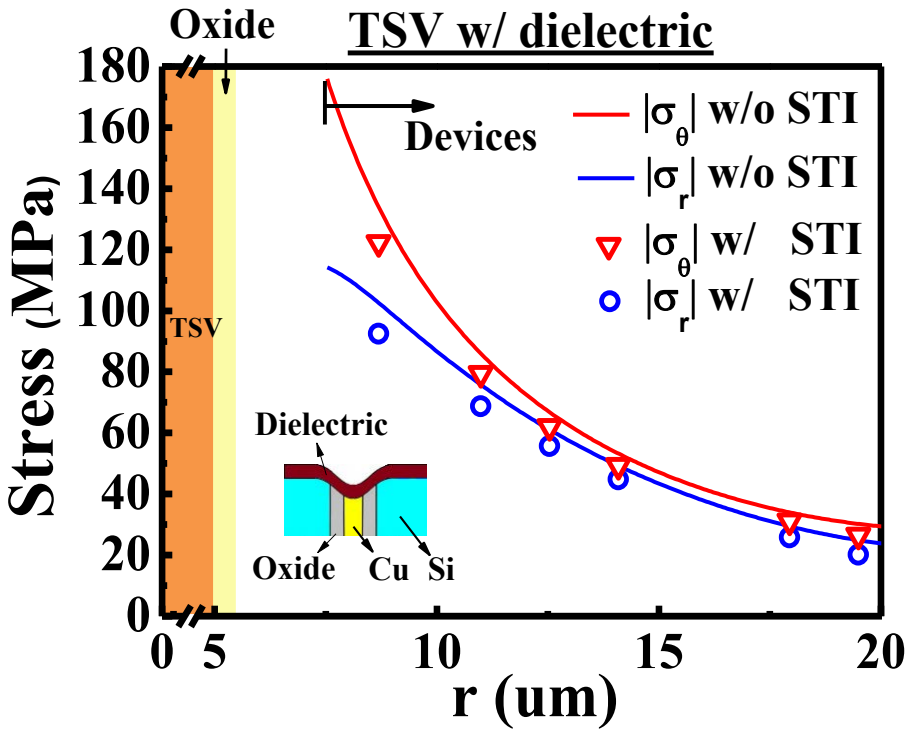


Fig. 6. 8 STI effect on the asymmetric stress field

6.5 Asymmetric Keep-out Zone of Via-last TSV

6.5.1 Analytic Model of KOZ

To model the KOZ based on the device layout, the stress field in cylindrical coordinates are transformed into Cartesian coordinates (**Fig. 6. 9**).

$$\sigma_{xx} = \cos^2\theta\,\sigma_r + \sin^2\theta\,\sigma_\theta \tag{6.7}$$

$$\sigma_{yy} = \cos^2\theta\,\sigma_\theta + \sin^2\theta\,\sigma_r \tag{6.8}$$

$$\sigma_{zz} = \sigma_{zz} \tag{6.9}$$

doi:10.6342/NTU201801788

The piezo-resistances along [100] current direction (π_{ij}) are transformed to [110] direction (π'_{ij}).

$$\pi'_{11} = \frac{1}{2} \cdot (\pi_{11} + \pi_{12} + \pi_{44}) \tag{6.10}$$

$$\pi'_{12} = \frac{1}{2} \cdot (\pi_{11} + \pi_{12} - \pi_{44}) \tag{6.11}$$

$$\pi'_{13} = \pi_{12} \tag{6.12}$$

By assuming that the change of I_{on} is totally caused by the change in carrier mobility (μ), the analytic model of KOZ is derived:

$$-\frac{\Delta\mu_{110}}{\mu_{110}} = \pi'_{11}\sigma_{xx} + \pi'_{12}\sigma_{yy} + \pi'_{13}\sigma_{zz} \tag{6.13}$$

$$\Rightarrow \delta I_{on} = \frac{\Delta I_{on}}{I_{on}} \approx \frac{K_1}{r^{1-n}}\cos^2\theta + \frac{K_2}{r^{1-n}}\sin^2\theta + \frac{K_3}{r^{1-n}} \tag{6.14}$$

$$K_1 = -(\pi'_{11}\sigma_r + \pi'_{12}\sigma_\theta)\big|_{r=R}\ R^{1-n} \tag{6.15}$$

$$K_2 = -(\pi'_{11}\sigma_\theta + \pi'_{12}\sigma_r)\big|_{r=R}\ R^{1-n} \tag{6.16}$$

$$K_3 = -\pi_{12}\sigma_{zz}\big|_{r=R}\ R^{1-n} \tag{6.17}$$

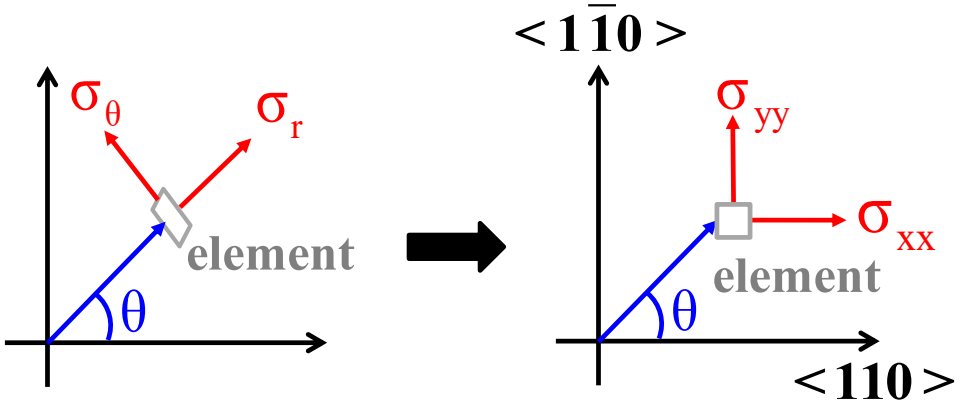


Fig. 6. 9 Coordinate transformations into Cartesian coordinates

doi:10.6342/NTU201801788

6.5.2 KOZs of nFET and pFET

Our model (equation 6.14) shows that $|K_1|$ and $|K_2|$ are different if the stress field is asymmetric. With the statistical I_{on} measurement of the radial and tangential devices across the 12-inch wafer, the K_1, K_2 values are determined, respectively. Note that the K_3 term could be neglected with the small vertical stress.

The asymmetry of KOZ is verified by the fitted K_1, K_2 values for both 28nm nFETs (**Fig. 6. 10**) and pFETs (**Fig. 6. 11**).

	K_1 (% · um^2)	K_2 (% · um^2)
nFET	123	-260

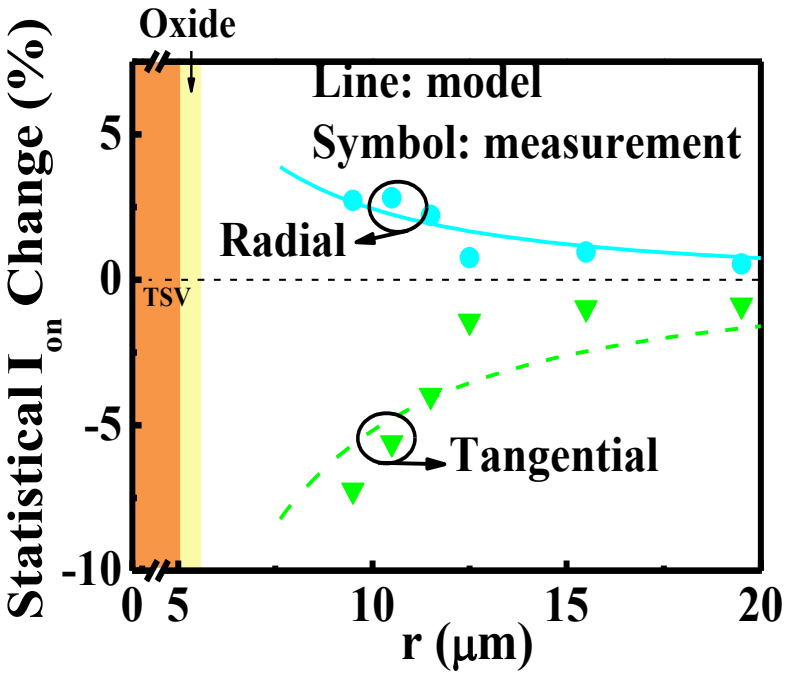


Fig. 6. 10 Fitted K_1, K_2 values with measured I_{on} of nFETs

doi:10.6342/NTU201801788

	K_1 (% · um^2)	K_2 (% · um^2)
pFET	-200	291

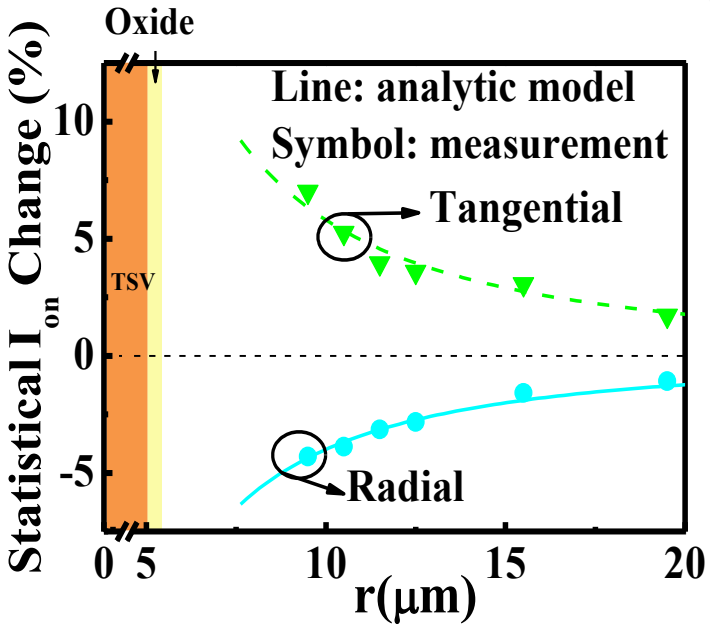


Fig. 6. 11 Fitted K_1, K_2 values with measured I_{on} of pFETs

The KOZ of nFETs is observed to be more asymmetric than that of pFETs (**Fig. 6. 12**). The lengths from the lobe edge of radial and tangential KOZ lobes to the TSV center (**Fig. 6. 12 (a)**) are directly proportional to $|K_1|^{1/1.7}$ and $|K_2|^{1/1.7}$, respectively. The level of asymmetry can be described by the value of $\left(\frac{|K_1|}{|K_2|}\right)^{1/1.7}$, which is 0.72 for nFET and 0.95 for pFET using the piezoresistances [10] and the simulated stress values. The asymmetry is consistent with our experimental data, qualitatively (0.68 and 0.83 for nFET and pFET, respectively). However, the numerical difference between simulation and data may be due to the internal stress of transistors, since the piezoresistances in the calculation are only obtained from the zero stress condition.

doi:10.6342/NTU201801788

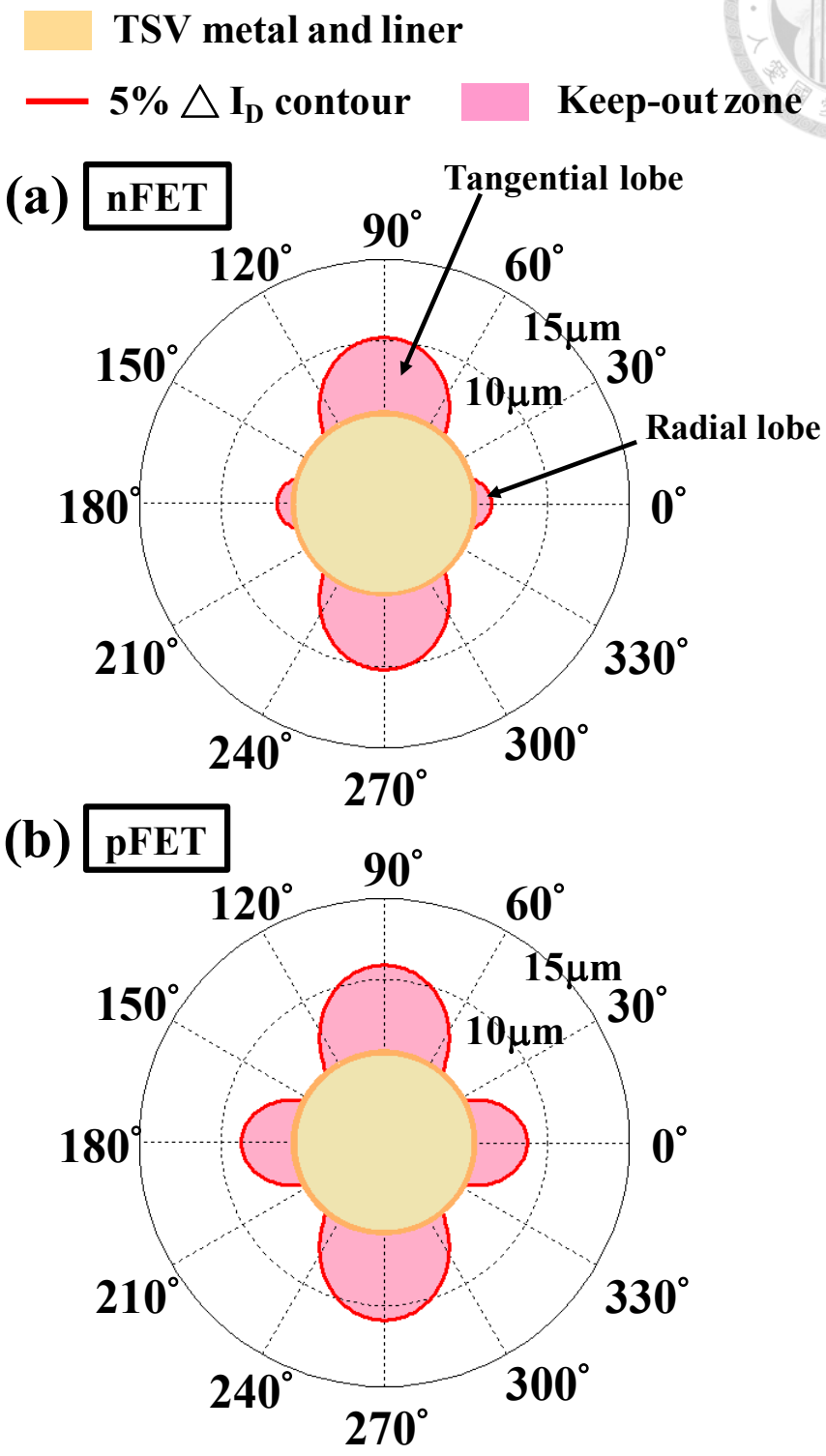


Fig. 6. 12 Graphical expression of **(a)** nFET KOZ and **(b)** pFET KOZ

doi:10.6342/NTU201801788

Comparable KOZ size is obtained from the measurement of nFETs and pFETs.

Fig. 6. 13 conceptually indicates the internal stress in transistors. The mobility change ($\Delta\mu/\mu$) corresponding to KOZ would be much different between nFETs and pFETs if the internal stresses were the same. For example, at 0.75GPa internal stress for nFETs and pFETs, the $\Delta\mu/\mu$ is 0.2 and 0.59, respectively. Similar KOZ implies that the pFET should have a larger internal stress, which can decrease $\Delta\mu/\mu$ to be ~0.2 from 0.59.

Table 6. 2 Piezoresistances from zero internal stress condition [6.10]

	π'_{11}	π'_{12}
nFET	-35.5	-14.5
pFET	71.7	-33.8

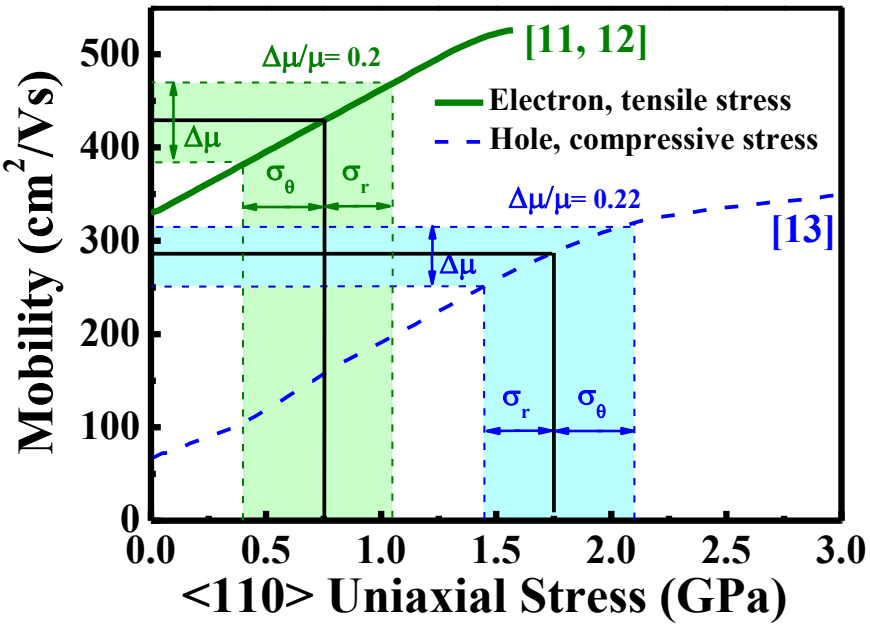


Fig. 6. 13 Electron and hole mobility versus <110> uniaxial stress with E_{eff} = 0.7MV/cm. The internal stress is conceptually drawn for nFET and pFET.

doi:10.6342/NTU201801788

6.6 Summary

In this chapter, an analytical KOZ model based on the asymmetric stress field caused by BEOL dielectrics on the top of TSV is experimentally verified by 28nm technology node. The KOZ decreases due to the reduction of TSV induced stress with the STI pattern. The origin of $|\sigma_r| \neq |\sigma_\theta|$ is the affectet radial displacement ($\sim r^{-0.7}$) on the Si surface by the BEOL dielectrics. The BEOL dielectrics also reduce the TSV proximity. nFETs has the measured KOZ more asymmetric than that of pFETs, which is similar to calculated KOZ using the simulated stress and the piezoresistance values. The similar mobility changes ($\Delta\mu/\mu$) leads to the comparable KOZ for nFETs and pFETs. The high internal stress is predicted to reduce the KOZ, since the mobility enhancement starts to saturate at high stress,

6.7 Reference

[6.1] R. Weerasekera, H. Y. Li, L. W. Yi, H. Sanming, J. Shi, J. Minkyu, and K. H. Teo, "On the impact of Through-Silicon-Via induced stress on 65-nm CMOS devices", *IEEE Electron Device Letters*, Vol. 34, No. 1, pp. 18-20, 2013.

[6.2] J. West, Y. S. Choi, and C. Vartuli, "Practical implications of via-middle Cu TSV-induced stress in a 28nm CMOS technology for wide-IO logic memory interconnects," in *Proc. VLSI Symp. Technol.*, 2012, pp. 101–102.

doi:10.6342/NTU201801788

[6.3] Y. Yang, G. Katti, R. Labie, Y. Travaly, B. Verlinden, and I. De Wolf, "Electrical evaluation of 130-nm MOSFETs with TSV proximity in 3D-SIC structure," in *Proc. IEEE IITC*, 2010, pp. 1–3.

[6.4] S. Cho, S. Kang, K. Park, J. Kim, K. Yun, K. Bae, W. S. Lee, S. Ji, E. Kim, J. Kim, Y. L. Park, and E. S. Jung, "Impact of TSV proximity on 45nm CMOS devices in wafer Level," in *Proc. IEEE IITC*, 2011, pp. 1–3.

[6.5] A. Mercha, G. Van der Plas, V. Moroz, I. De Wolf, P. Asimakopoulos, N. Minas, S. Domae, D. Perry, M. Choi, A. Redolfi, C. Okoro, Y. Yang, J. Van Olmen, S. Thangaraju, D. Sabuncuoglu Tezcan, P. Soussan, J.H. Cho, A. Yakovlev, P. Marchal, Y. Travaly, E. Beyne, S. Biesemans, and B. Swinnen, "Comprehensive analysis of the impact of single and arrays of through silicon vias induced stress on high-k/metal gate CMOS performance," *IEDM Tech. Dig.*, 2010, pp. 2.2.1-2.2.4.

[6.6] M. Aoki, F. Furuta, K. Hozawa, Y. Hanaoka, H. Kikuchi, A. Yanagisawa, T. Mitsuhashi, and K. Takeda, "Fabricating 3D integrated CMOS devices by using wafer stacking and via-last TSV technologies," *IEDM Tech. Dig.*, 2013, pp. 29.5.1-29.5.4.

[6.7] W. Guo, V. Moroz, G. Van der Plas, M. Choi, A. Redolfi, L. Smith, G. Eneman, S. Van Huylenbroeck, P. D. Su, A. Ivankovic, B. De Wachter, I. Debusschere,

doi:10.6342/NTU201801788

K. Croes, I. De Wolf, A. Mercha, G. Beyer, B. Swinnen, and E. Beyne, “Copper through silicon via induced keep out zone for 10nm node bulk FinFET CMOS technology,” *IEDM Tech. Dig.*, 2013, pp. 12.8.1-12.8.4.

[6.8] S.-R. Jan, T.-P. Chou, C.-Y. Yeh, C. W. Liu, R. V. Goldstein, V. A. Gorodtsov, and P. S. Shushpannikov, “A compact analytic model of the strain field induced by through silicon vias,” *IEEE Trans. Electron Devices*, Vol. 59, No. 3, pp. 777–782, Mar. 2012.

[6.9] M.-Y. Tsai, P.-S. Huang, and P.-C. Lin, “A study of overlaying dielectric layer and its local geometry effects on TSV-induced KOZ in 3-D IC,” *IEEE Trans. on Electron Devices*, Vol. 61, No. 9, pp.3090-3095, 2014.

[6.10] S. E. Thompson, S. Guangyu, Y. S. Choi, and T. Nishida, “Uniaxial process induced strained-Si: Extending the CMOS roadmap,” *IEEE Trans. on Electron Devices*, Vol. 53, No. 5, pp. 1010–1020, May 2006.

[6.11] H.-S. Lan, S.-T. Chan, T.-H. Cheng, C.-Y. Chen, S.-R. Jan, and C. W. Liu, “Biaxial tensile strain effects on photoluminescence of different orientated Ge wafers," *Appl. Phys. Lett.* **98**, 101106 (2011).

[6.12] S. Suthram, J. C. Ziegert, T. Nishida, and S. E. Thompson, "Piezoresistance coefficients of (100) silicon nMOSFETs measured at low and high (~ 1.5 GPa) channel stress," *IEEE Electron Device Letters*, Vol. 28, No. 1, pp. 58-61, 2007.

doi:10.6342/NTU201801788

[6.13] G. Sun, Y. Sun, T. Nishida, and S. E. Thompson "Hole mobility in silicon inversion layers: Stress and surface orientation." *Journal of Applied Physics* 102.8 (2007): 084501-084501.

doi:10.6342/NTU201801788

Chapter 7 Summary and Future Work

7.1 Summary

SHE of 14nm FinFETs is simulated using TCAD to model the intrinsic thermal resistance. The proposed two-step pseudo isothermal plane model can provide accurate BEOL thermal resistance value with different metal densities. The large thermal resistance of the free convection of air ($R_{th,convection}$) affects the direction of heat flow. It is important to extract the intrinsic thermal resistance from the overall thermal resistance. The overall thermal resistance is affected by substrate thickness and the heat dissipating scenario of the chip. Device temperature can be predicted for the IC with flip-chip packaging by the proposed $R_{th0,device}$ and $R_{th,BEOL}$ model under face-down configuration.

The capacitive loading of the inverter with 4-fin-1-finger FinFETs is calculated to be 21.6fF using the output current measured experimentally. In the transient SHE simulation of an inverter, T_j in pFETs is higher than in nFETs due to the low thermal conductivity of SiGe S/D. It is merely possible to monitor the device temperature with transient AC input by the change of metal line resistance. The simulated residual T_j is lower than 20°C after 1500 cycles operation and the simulate ΔT_{M1} is about 0.07°C after the first period.

doi:10.6342/NTU201801788

The thermal time constant of the hotspot ($\tau_{hotspot}$) in scaled FinFETs is linearly dependent to input frequenct with the fitted slope of -0.71ps/GHz. Modularized thermal SPICE Modeling of FinFETs and BEOL with distributed R_{th}-C_{th} network has been verified both with DC and AC inputs. Devices with different layouts can be easily implemented using the proposed SPICE model. Interfacial thermal resistance leads to siginificantly increased T_j, and was not considered in lumped R_{th}-C_{th} models and device simulation. The EM MTTF predictions for the incorporation of Co interconnect and different circuit designs are demonstrated using the RO. Co with low thermal conductivity as compared to Cu and W leads to the increasing temperature and poisions reliability. Layout design such as reducing the distance between V2 bundles and placing additional V2s can lower the raised temperature and extend the life time. Additional thermal via in the BEOL and via0s on the multi-fin FinFET are proposed to decrease the $R_{th,BEOL}$ and $R_{th0,FinFET}$, respectively.

The stress field near a TSV fabricated by a via-last process is affected by the BEOL dielectrics. The simulated stress is reduced by 8% with the STI pattern. The observed $|\sigma_r| \neq |\sigma_\theta|$ leads to the asymmetric keep-out zone (KOZ) for both nFETs and pFETs. The proposed ssymmetric KOZ modeling of devices can provide the graphical KOZ. The comparible KOZ sizes of nFETs and pFETs is explained by the large internal strain in pFETs.

doi:10.6342/NTU201801788

7.2 Future Work

1. FinFET structure will continue scaling for 7nm and 5nm technology nodes. The scaling effects on the intrinsic thermal resistances should be further studied.
2. Gate-all-around devices such as nanowire, nanosheet, and stacked nanosheets are promising candidates to achieve good electrostatic control. The self-heating effect in GAAFETs should be simulated to provide the design guideline of device geometry.
3. The proposed compact thermal model of FinFET should be applied to scaled FinFETs with the geometry of 7nm or 5nm technology node, and compared with TCAD simulation results.
4. The width of interconnect scales down together with FEOL devices, and the thermal conductivity of metals will be lowered by the increasing boundary scattering. The metal thermal conductivity with confined geometry should be implemented into the SPICE to provide accurate T_{metal}.
5. With presise T_j and T_{metal} mapping from device to circuit level, the temperature effects on device and BEOL reliability lifetime should be further investigated.

E-mail: F01943063@ntu.edu.tw

*This dissertation was typed by the author.

doi:10.6342/NTU201801788