Received 4 June 2018; revised 13 July 2018 and 19 July 2018; accepted 21 July 2018. Date of publication 24 July 2018; date of current version 14 August 2018. The review of this paper was arranged by Editor P. Pavan.

Digital Object Identifier 10.1109/JEDS.2018.2859276

An Experimental Approach to Characterizing the Channel Local Temperature Induced by Self-Heating Effect in FinFET

E RAY HSIEH[1] (Member, IEEE), MENG-RU JIANG[1], JIAN-LI LIN[1], STEVE S. CHUNG[1] (Fellow, IEEE), TSE PU CHEN[2], SHIH AN HUANG[2], TAI-JU CHEN[2], AND OSBERT CHENG[2] (Member, IEEE)

1 Department of Electronics Engineering and Institute of Electronics, National Chiao Tung University, Hsinchu 300, Taiwan
2 United Microelectronics Corporation, Hsinchu 300, Taiwan

CORRESPONDING AUTHOR: STEVE S. CHUNG (e-mail: schung@cc.nctu.edu.tw)

This work was support by the Ministry of Science and Technology, Taiwan, under Contract MOST 104-2221-E-009-005 and Contract MOST 107-2633-E-009-003.

ABSTRACT In this paper, we have developed a methodology of a lateral profiling technique of the channel local temperature in 14 nm FinFET, incurred by the self-heating effect (SHE). As SHE happens, the thermal source generated near the drain will dissipate toward the source side. Since the interaction between RTN trap and channel carriers is very sensitive to the temperature, the channel local temperature can be extracted through this interaction process between random-telegraph-noise (RTN) trap and carriers, and the position of the channel local temperature can be obtained from the RTN trap position. The results show that the highest temperature happens at the drain edge during SHE and pFinFET exhibits a much higher temperature than that of nFinFET. Furthermore, the distribution of channel local temperature can be described by the Fourier's law of thermal conduction. Averaged channel temperature can be used to extract the thermal resistance, R_{th}, which increases rapidly as the channel length is scaled down to 20 nm, further degrading the SHE, in terms of a significant short channel effect. We also found that the incremental channel resistance is proportional to the incremental channel local temperature, whose slope indicates the degree of SHE, and the slope of pFinFET is larger than that of nFinFET. Finally, SHE will cause 10% and 14% degradation of $I_d V_{ds}$ for n- and pFinFET respectively. This can be reasonably explained by the decay of saturation velocity in high temperature. The results obtained based on this methodology will help us on the understanding of the SHE impact on a nano-scaled FinFET device.

INDEX TERMS FinFET, self-heating effect, random telegraph noise.

I. INTRODUCTION

The self-heating effect has been an important issue since the development of the semiconductor-on-insulator (SOI) architecture in 1990s because the thermal conductance of buried SiO_2 in SOI is much smaller than that of silicon. When CMOS devices are forced by a driving voltage, channel minority carriers with a high speed transport in the cramped-volume channel between the gate-oxide and buried oxide. During this process, the phonon scattering event increases, which continues to heat up the silicon lattice. Since it is difficult to dissipate the heat through the buried oxide, the thermal is accumulated in the channel. As a result, the channel local temperature arises, known as "self-heating effect" (SHE) [1]. As MOSFETs are further scaled following the G. E. Moore's prediction [2], the device structure has been evolved into a three-dimensional (3D) FinFET to overcome severe short-channel effect. In the first generation of FinFET, it was fabricated on SOI [3]; however, the cost on SOI is much higher than that on bulk wafers, and hence the mainstream of FinFET has been fabricated on a bulk wafer since 22nm node [4]. In a bulk-type FinFET, there are two kinds of the shallow-trench –isolation (STI), including the isolated STI and inner-Fin STI. The former is used to isolate n- and p-well, which should be wider and deeper; the latter is inserted between fins in a multi-fin structure to define the fin-pitch, which should be narrower and shallower, Fig. 1(a).

2168-6734 © 2018 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

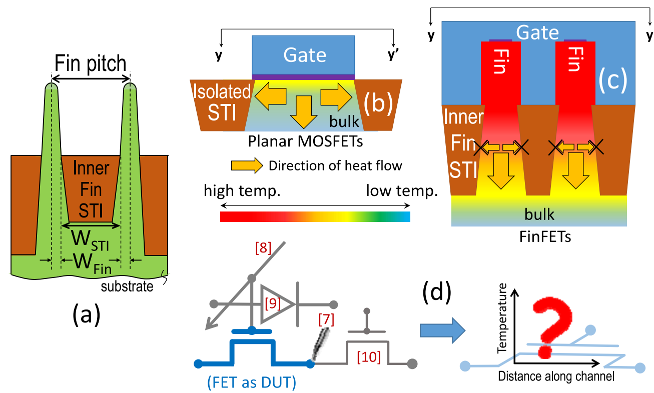

FIGURE 1. (a) Definition of Fin-pitch. (b) In planar bulk MOSFET, heat can be dissipated not only to the lateral direction but also to well direction. (c) However, in bulk FinFET, the path of dissipation has been stopped by inner-Fin STI in the lateral direction, which decreases efficiency of heat dissipation. (d) Different approaches have been developed to characterize device temperature generated by SHE but failed to provide the details of channel local temperature from S to D.

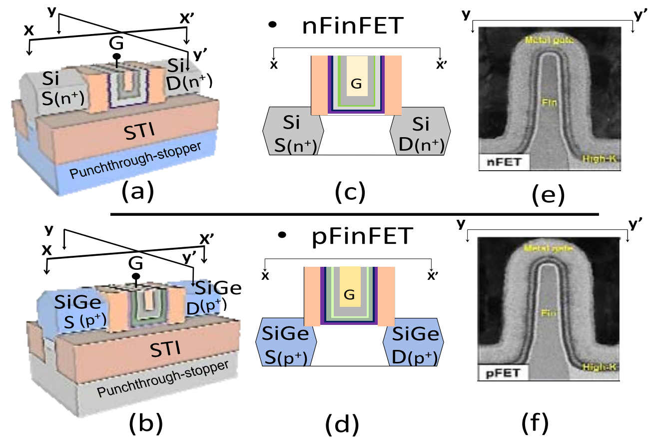

FIGURE 2. (a), (b) are device structures of n- and p-FinFET, respectively; (c), (d) are the cross sections in the channel direction, respectively; (e), (f) are the cross sections in the width direction, respectively.

As FinFET has been developed from 22nm, 16/14nm [5], to 10/7nm [6], the fin width becomes narrower, and the fin height is taller. As a result, SHE-generated heat cannot propagate through inner-fin STI in width direction and only can dissipate into the well. Efficiency of thermal dissipation in bulk FinFETs becomes worse, compared to that of planar CMOS devices, in Fig. 1(b) and 1(c). Eventually, the SHE happens in bulk FinFET again. SHE causes the reliability degradation of FinFET, such as PBTI and NBTI [6]. As far as SHE is concerned, the most important physical parameter of FinFET is the channel temperature. In the past, researches have been dedicated to develop experimental approaches to characterizing this parameter. Menges *et al.* [7] utilized Atomic force Microscope (AFM) to scene tiny delta-current-signals induced by the temperature differences from the device surface; Takahashi *et al.* [8] used the four-point probes to measure the channel temperature from a four-terminal poly-liner on the DUT(device under test); Franco *et al.* [9] used solid-state devices as temperature-sensors, such as diode or MOSFET [10] near the DUT to detect its thermal radiation during operations, as shown in Fig. 1 (d). However, the previous works only provided an average value of the device temperature, which is measured from a relative longer distance to the DUT but not the exact local temperature inside the DUT from SHE. The *uniqueness* in this work is to develop an experimental measurement to characterize the channel local temperature of a FinFET by the using of random-telegraph-noise (RTN) approach so as to profile the channel local temperature along the channel.

This paper is organized as follows. Section II describes the fabrication of devices. Section III will develop a methodology and its experimental approach to characterize the channel local temperature of FinFET. Section IV will give the experimental results of the channel local temperature and its lateral profile from the source to the drain. Moreover, the impacts of SHE on the FinFET will be discussed as well. Finally, summary and conclusions will be given.

II. THE FABRICATION OF 14NM FINFET

Fig. 2(a) and (b) show device structures of n- and p- FinFET, respectively. Fig. 2(c) and (d) show the cross-section views of n- and p-FinFET from source to drain, respectively. Fig. 2(e) and (f) show the TEM cross-section of n- and p-FinFET in the width direction, respectively. Si Fin-bodies have been defined with an assistance of inner-fin STI formation. Punch-through stopper has been implanted deeply into the bottom of the Fin-body to prevent the leakage current from the source-to-drain. Next, the epitaxial Si and SiGe S/D with in-situ doped impurities for n- and pFinFET have been formed and attached on the channel Fin after the sacrificial poly-gate deposition on a 0.8nm ultra-thin SiO_2 thermal oxide. After removal of the poly-gate, 2nm atomic-layer-deposition (ALD) HfO_2 was grown on 0.8m SiO_2 film. Then, ALD TiN film was deposited on HfO_2 layer, followed by the work-function metal. Finally, the post-metal-anneal at 600°C for 5 minutes has been carried out. E.O.T. (Equivalent Oxide Thickness) is equivalent to 1.2nm (2nm HfO_2/0.8nm SiO_2), formed by HK-last and metal-last process. Different dimensions of devices with various fin numbers have also been prepared.

III. THE METHODOLOGY OF CHANNEL LOCAL TEMPERATURE PROFILING BY RTN MEASUREMENT

Random-telegraph-noise (RTN) is in the form of digital waveform with two levels, which is actually caused by the traps generated in the gate dielectric [11]–[14]. Using n-channel MOSFET as an example, when an RTN trap captures the electron, this trap has been negatively-charged, and hence V_{th} of a transistor increases, while I_d decreases accordingly. On the contrary, when the trap emits the trapped electron, V_{th} is brought back to the low value, thereby I_d jumps to a higher current level, Fig. 3(a). Fig. 3(b) shows a switching waveform of RTN signals. The low level of RTN signals

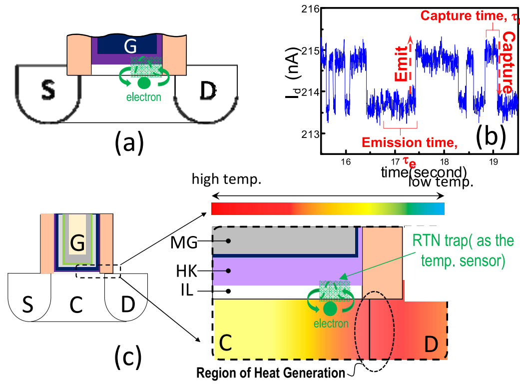

FIGURE 3. (a) RTN are signals induced by the interaction of traps in gate dielectric and electrons in the channel, which forms a digital-like waveform, composed of τ_c and τ_e, (b). (c) Since RTN traps are very close to the heat generating source during SHE, they can be used to sense the channel local temperature.

is defined as τ_e, which is the time-to-emit of the trapped electron from the RTN trap; the high level of RTN signals is defined as τ_c, which is the time-to-be-captured for an electron into the RTN trap from the channel. RTN traps can significantly impact the degradation of device reliabilities when CMOS devices are suffered from the bias-temperature-instability [15]–[19] or breakdown [20]. One can extract the values of τ_e and τ_c to calculate the depth of RTN traps in the dielectrics [21]. When the self-heating effect (SHE) occurs in a FinFET, the thermal source will be generated in the Si Fin near the drain and beneath the spacer [22], as shown in Fig. 3(c). If an RTN trap in the gate dielectric is located in proximity to the thermal source, the capturing or emitting electrons of RTN trap will be affected by the temperature gradient from the thermal source. As a result, we came up with an idea by using the RTN trap *to measure the channel temperature*. Fig. 4(a) shows the comparisons of the time to emit, τ_e, of nFinFET in logarithm scale, Ln|τ_e|, against the drain-to-source bias, V_{ds}, and the chuck temperature, respectively. Both values of Ln|τ_e| are functions of V_{ds} or the chuck temperature which show the same trend, that is, the higher the V_{ds} or the chuck temperature is, the smaller the value of Ln|τ_e| becomes. As the chuck temperature is fixed, the V_{ds} goes high, the channel-to-drain depletion region is expanded, and transporting electrons in the channel will be accelerated by the built-in electric field under this depletion region, which become much hotter and reach the channel region near the drain. As a result, the local channel temperature increases when V_{ds} is raised up. This increasing temperature intensifies the interaction between electrons and RTN traps, and the channel electrons are captured by traps more easily, making a decreasing time-to-emit (τ_e) of electrons in the RTN trap. The green line in Fig. 4(a) is a two-slope curve. Before V_{ds} = 0.51 V, the slope is slow changing but after that the slope changes fast. It is believed that the local temperature around this RTN trap is gradually increased when $V_{ds} >$ 0.51V, which gives rise to an increase of the slope. Therefore,

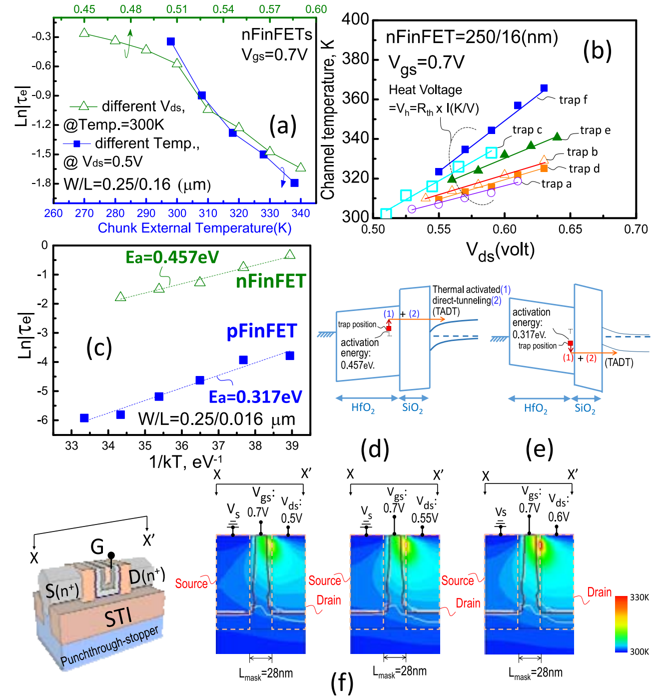

FIGURE 4. (a) The V_{ds}-induced SHE causes the raise of channel local temperature, which exhibits a similar trend of τ_e as function of the chuck temperature. Thus, (b) the chuck temperature is correlated to V_{ds} through τ_e to transfer this external temperature to the channel local temperature caused by V_{ds}. (c) The extraction of activation energy, and the thermal activation plays an important role in the interaction of the RTN trap and electrons, (d) and (e). (f) shows the simulated lattice temperature of nFinFETs under different bias conditions. The hot spot happens at the underlap region between the gate and drain. Meanwhile, the temperature is gradually increasing from 300K to around 330K as the Vds raises to 0.6V.

if we keep V_{ds} constant and raise the chuck(external) temperature to directly heat the RTN trap, we should observe a similar trend since the interaction between electrons and RTN traps obeys the same principle as the previous case. Furthermore, we can directly match the V_{ds} and the chuck temperature through τ_e in Fig. 4(a) to plot the temperature as a function of the V_{ds}, as shown in Fig. 4(b), whose slope is equivalent to 380 in unit of Kelvin divided by Volt (K/V). K/V is defined as the heat-voltage (V_h), which can be derived as the following,

$$V_h = R_{th} \cdot I = \frac{T}{P} \cdot I \equiv \frac{T}{V \cdot I} \cdot I = \frac{T}{V}. \text{ Unit: K/Volt} \quad (1)$$

Since the temperature divided by voltage can be written as thermal resistance multiplied by the current, this physical quantity is reasonably treated as the heat-voltage (V_h) in perspective of a thermal circuitry. Moreover, V_h can be used to efficiently quantify how much temperature can be increased in a bounded space by the external voltage during SHE. Then, one may ask how the interaction between RTN trap and electrons is affected by the channel temperature? In Fig. 4(c), if we take a plot of Ln|τ_e| against the inverse of $K_B T$, in which K_B is Boltzmann constant, a linear fitting slope can be obtained. This slope represents the

activation energy of an electron in the RTN trap, as derived in [23], i.e.,

$$\tau_e = \frac{\exp\left[(E_F - E_T)/k_B T\right]}{g\sigma v_{th} n}; \tag{2.1}$$

$$\ln|\tau_e| = \frac{E_F - E_T}{k_B T} - \ln|v_{th}| - \ln|g\sigma n|, \tag{2.2}$$

where E_F-E_T is defined as the activation energy, i.e., the energy difference from the trap level to the Fermi-level of the channel; g is a degeneracy factor, as unity of trapped electrons; s is the cross section of the trap; v_{th} is the thermal velocity; n is electron concentration of the channel. In our experiments, g is equivalent to unit since we only deal with two-level signals of RTN; s is a feature only relied on the RTN trap, which can be treated as a constant; n will be saturated as the minority carriers in channel are strongly inverted. Therefore, we neglect dependency of Ln $|\tau_e|$ on the last term of (2.2). Since v_{th} is a logarithm function, it shows much weaker dependency on Ln $|\tau_e|$ than (E_F-E_T)/ K_BT. As a result, T is the only dominant factor dependent on varying of Ln $|\tau_e|$. In other words, the captured electrons need to be activated by the environment temperature from the energy level of the trap and jump to the activation energy level by tunneling through the energy barrier of the gate dielectrics to the channel, in Fig. 4(d) and (e). Although the process of capture and emission for RTN traps should involve the mechanism of trap-assisted tunneling(TAT) [24], [25], we can still use simple equations, (2.1) and (2.2), which only consider the thermal activation and direct tunneling mechanisms, for the extraction of activation-energy levels. Fig. 4(f) shows the simulated results of lattice temperatures for nFinFET under V_{gs} = 0.75V and V_{ds} = 0.5V to 0.6V in consideration of the quasi-ballistic transport and hydrodynamic model. It was found that the thermal source was generated at the underlapped region between the gate and the drain. At V_{ds} = 0.5V, the temperature is slightly larger than the environmental temperature, 300K. With gradual increase of V_{ds}, the temperature around the thermal source raises correspondingly, up to ~330K, which is comparable to the results given in Fig. 4(b).

Since the increment of local temperature around RTN trap will assist the jump of trapped electron from the trap to the activation energy level, clearly, we can detect the change of channel local temperature by the inspection of τ_e from RTN measurement. However, the channel local temperature detected by the RTN trap is strongly dependent on the lateral distance of the trap to the thermal radiation. The closer the distance of RTN trap to the thermal source is, the higher the channel local temperature will be detected. Then, if we can identify each RTN trap in the lateral position of gate dielectrics, the lateral profile of channel local temperature caused by SHE can be delineated.

Fig. 5 shows the channel band-diagram in the direction of source-to-drain. It was found that the interaction between the RTN trap and electrons is most sensitive at the channel-barrier peak [21], [26]. If we apply a positive V_{ds}

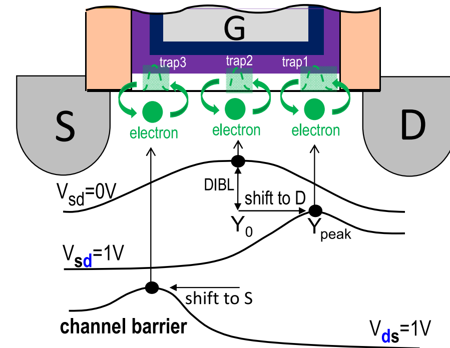

FIGURE 5. The RTN signals are most sensitive at the channel barrier peak. We can sweep the barrier peak to identify the lateral position of RTN traps by applying V_{ds} (shifting the peak to the source) or V_{sd} (shifting the peak to the drain).

on a FinFET, the barrier peak will be shifted to the source side owing to the drain-induce-barrier –lowering (DIBL). On the other hand, if a positive V_{ds} is applied at the source of a nFinFET, the barrier peak moves to the drain side. So, we can sense the RTN signals at each lateral position in the gate dielectric via the shift of the channel barrier peak by sweeping V_{ds} or V_{sd}. But it is to be noted that by using this approach, the V_{gs} is kept low around V_{th}, in our case: 0.25V~0.5V, to avoid the occurrence of SHE. After determination of the trap position, the V_{gs} can be ramped high to V_{dd} so that we can observe the SHE effect and detect the local temperature around the trap through RTN measurement.

Moreover, the relationship between the channel-middle position and the shifted channel barrier peak can be given by [26]:

$$\frac{Y_{peak}}{(L_{eff} - \Delta L)/2} = \frac{\left(V_{bi} - V_{c,\max}\right) - DIBL}{\left(V_{bi} - V_{c,\max}\right)}, \tag{3}$$

$$\frac{L_{eff} - \Delta L}{L_{eff}} = \frac{S.S. - 60mV}{S.S._0 - 60mV}, \tag{4}$$

where Y_{peak} is the position of shifted channel barrier peak from the channel middle, (L_{eff}-ΔL)/2. L_{eff} is the effective channel length, and ΔL is the depletion length of the channel-to-drain junction or the channel to source junction as V_{ds} or V_{sd} is applied respectively. S.S. is the subthreshold swing of I_d-V_{gs} at a certain V_{ds} bias, and $S.S._0$ is the subthreshold swing of I_d-V_{gs} at V_{ds} = 0.05V. V_{bi} is the channel-barrier height, and $V_{c.max}$ is the built-in potential between the source and channel. V_{bi} and $V_{c,max}$ can be extracted experimentally. However, L_{eff} should be further characterized.

As shown in Fig. 6(a), the effective channel length (L_{eff}) is defined as the distance in the channel between the source and drain(S/D) beneath the gate dielectric. Before 28nm node, the overlapped region of S/D under the gate dielectric (ΔL in Fig. 6(a)) have been adapted to aggressively reduce the conducting channel in terms of the decreasing S/D resistance. It is important to characterize the effective channel length for the determination of the RTN lateral position. We may

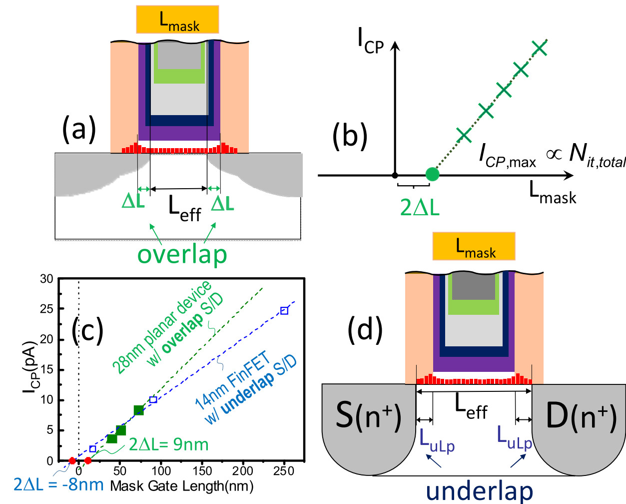

FIGURE 6. (a) The effective channel length is a subtraction of the S/D overlapped length from the gate mask length. (b) The charge pumping current can be used to characterize the effective channel length. (c) In experimental 14nm FinFET, there is a S/D underlapped length to the gate, as shown in the schematic of (d).

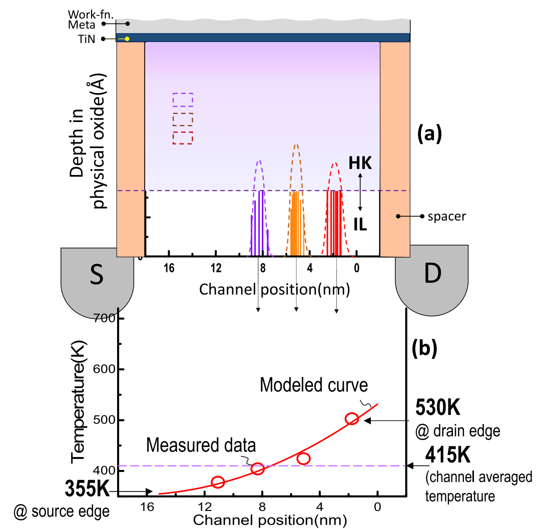

FIGURE 7. (a) Lateral profiles of the process induced RTN traps show Gaussian-like distribution in 14nm nFinFET. Peaks of Gaussian distribution are centers of RTN traps. (b) A lateral profile of the channel local temperature obeys the Fourier's law of thermal conduction, and the highest temperature happens near the drain around 530K, in which an averaged channel temperature is around 425K.

use the charge pumping technique to extract the effective channel length as follows.

If the distribution of interface traps between the gate dielectric and channel is uniform, the charge-pumping current is reasonably assumed to be proportional to the effective channel length. So, if we plot the charge-pumping current against the mask lengths of the same width devices, the intercept of this line extends to $I_{cp} = 0$ gives the value of the overlap between the gate and source/drain [27], i.e., 2 ΔL in Fig. 6(b) where the effective channel length is smaller than the mask gate length. The effective channel length is to subtract the overlapped length(L_{ov}) from the mask gate length (L_{mask}). Fig. 6(c) provided measured data of L_{eff} by using charge pumping technique in 28nm (the green curve) and 14nm (the blue curve) nMOSFET, respectively. For 28nm ones, the intercept of the straight line with x-axis gives 2 ΔL = 9nm, i.e., the effective channel length is 19nm. On the other hand, for 14nm ones, the intercept gives a value of $2\Delta L = -8$nm. This negative value has a different meaning from that of 28nm node which is positive. In other words, there are the so-called underlap regions underneath the spacer region in the channel, labeled as L_{uLp}, in which the source and drain do not overlap with the gate, as shown in Fig. 6(d) where the effective channel length is longer than the mask gate length. As a result, the effective channel length of 14nm FinFET is around 28nm.

In short, we have used the channel barrier peak, Fig. 5, to determine the lateral position of the trap from the source to the drain. Also, from the dependence of τ_e on the local temperature, Fig. 4, we may calculate the channel local temperature by using RTN measurement. Moreover, the effective channel length has been determined by the charge-pumping measurement, Fig. 6. To combine the above measured or calculated information, we may obtain the detailed lateral profile of the channel local temperatures.

IV. ANALYSIS OF THE CHANNEL TEMPERATURE

Fig. 7(a) shows the lateral profile of process-induced RTN traps from the source edge to drain edge of nFinFET. There are three traps characterized from three different nFinFETs with the same dimension. All traps are shown in a Gaussian-shape. As channel electrons are gradually approaching to the RTN trap, RTN trap will capture electrons more easily, in terms of the increment of τ_e. On the contrary, if electrons are away from the RTN trap, it is more difficult for RTN trap to capture electrons, leading to a decrease of τ_e. Therefore, the closer the electrons to the RTN trap is, the larger τ_e can be measured. Finally, a Gaussian distribution of RTN trap potential can be depicted. The peaks of these Gaussian distributions are the centers of RTN traps. After the determination of the trap positions, τ_e can also be used to correlate the local channel temperature and the V_{ds} bias based on the methodology mentioned in the last section, and the result are given in Fig. 7(b). The dotted circles of Fig. 7(b) are the measured raw data, the solid line is modeled by Fourier's law of thermal conduction, which is given by

$$Q = -kA\frac{dT}{dx}. \quad (5)$$

Q is the thermal conductance; k is the thermal conductivity; A is the cross-section area of the channel in FinFET, and x is the distance in the channel lateral direction. As far as the steady-state is concerned, and k is constant, if the heat generation rate is g in FinFET during SHE, the one-dimensional

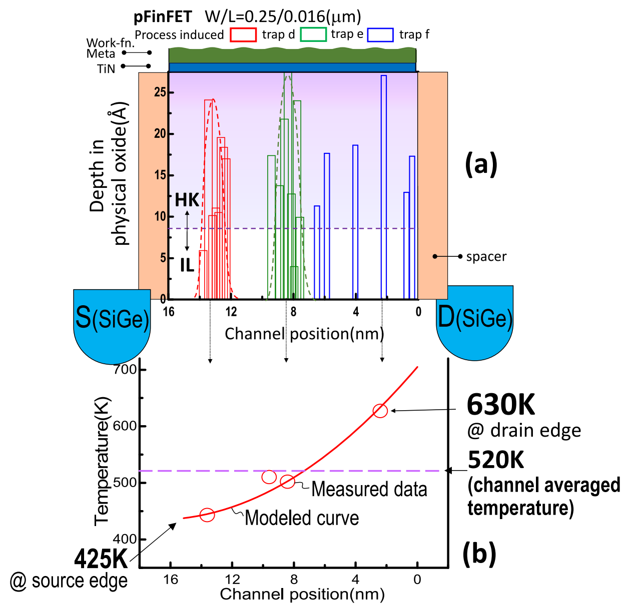

FIGURE 8. (a) In p-FinFET, lateral profiles of the process induced RTN traps show a Gaussian-like distribution. Peaks of Gaussian distribution are centers of RTN traps. (b) A lateral profile of the channel local temperature obeys the Fourier's law of thermal conduction, and the highest temperature happens near the drain around 700K, from which an averaged channel temperature is around 520K.

heat transfer equation can be derived from Eq. (5) as,

$$\frac{d^2T}{dx^2} + \frac{g}{k} = 0;\ \ T(x) = \frac{g}{k} \cdot x + c_1 \cdot x + c_2. \tag{6}$$

The unknown constant parameters, c_1 and c_2, can be extracted by the parabola regression algorithm, given that the values of temperature(T) for the channel lateral positions have been experimentally characterized. k is the inverse of thermal resistance, R_{th}, which can be extracted experimentally(see Fig. 9), and k used here is 0.85(nW/K) for nFinFET and 0.65(nW/K) for pFinFET with $L_{mask} = 20$nm. As far as the parameter, g, is concerned, it is used as a tunable parameter to fit the curve well. Hence, g is 0.98 nFinFET and 0.83 for pFinFET with unit of $(k/cm)^2(nW)^{-1}$. In Fig. 7, the dots are fitted on the modeled line very well, and the highest local temperature (>500K) has been detected near the drain. Moreover, the channel local temperature has been decayed sharply from the drain edge to the source edge. The latter is very close to the temperature around the environment. (350K) The result shows the efficiency of heat dissipation for the 14nm nFinFET is good enough because it still can be dissipated from the source-to-drain rapidly. On the other hand, Fig. 8(a) shows the experimental results of lateral profiles for process-induced RTN traps in pFinFET. Depths of RTN traps for pFinFET are deeper than those of nFinFET. The shape of RTN traps in pFinFET exhibits a Gaussian-like as well. The local temperature of each RTN traps has been plotted in Fig. 8(b), where the dotted circles are the measured raw data, and the curve is

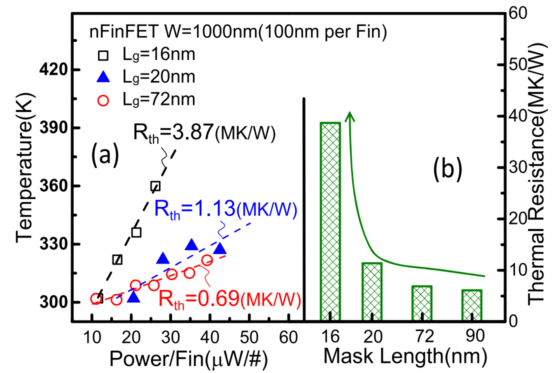

FIGURE 9. (a) The local temperature as a function of power per Fin, whose slopes are the thermal resistance, R_{th}. (b) R_{th} increases sharply as the mask gate length is scaled down to 20nm, raising a big concern for short channel devices.

a fitting line, predicted by Fourier's law of Thermal conductance. The trend of the fitting curve is similar to that of nFinFET in Fig. 7(b); however, the channel local temperature of pFinFET, 630K, near the drain, is higher than 500K of nFinFET, which shows that the efficiency of heat dissipation in pFinFET is much worse than that in nFinFET. This is because SiGe has been used as the stressor in S/D to enhance the mobility of pFinFET instead of Si in S/D of nFinFET. The thermal conductance of SiGe is much smaller than that of Si. Hence, the efficiency of heat dissipation is quite low in channel lateral direction for pFinFET, leading to much higher channel local temperature. As a result, pFinFET has suffered from serious SHE, which will deteriorate the device's electrical characteristics.

In the following, Fig. 9 will analyze the thermal resistance, which can be expressed as,

$$R_{th} \equiv \frac{T_{DUT} - T_{env.}}{p} = \frac{T_{DUT} - T_{env.}}{I_d V_{ds}/W_{Fin}}, \tag{7}$$

where R_{th} is the thermal resistance; T_{DUT} is average local temperature of DUT; T_{env} is environmental temperature. Fig. 9(a) plots averaged local temperature against power per fin, whose slope defines R_{th}. Fig. 9(b) compares R_{th} for the devices with different mask lengths. It was found that R_{th} increases as the mask length decreases. Moreover, R_{th} increases sharply as the mask length is smaller than 20nm, which shows that the short-channel effect aggravates the self-heating effect. The results show that R_{th} is 3.87(MK/W) for an 14nm nFinFET with the mask length equivalent to 16nm, which is close to the value of 3.82(MK/W) in the 16nm nFinFET, reported by IBM. [28] Fig. 10(a) shows the increment of channel total resistance, R_{tot} as function of the incremental channel local temperature. The result shows that R_{tot} is linearly proportional to the channel averaged temperature for n- and p-FinFET. Moreover, the slope of Fig. 10(a) can indicate how much FinFET has been suffered from the SHE. 0.88 of the slope in pFinFET is larger than 0.63 in nFinFET, which again confirms SHE in pFinFET is

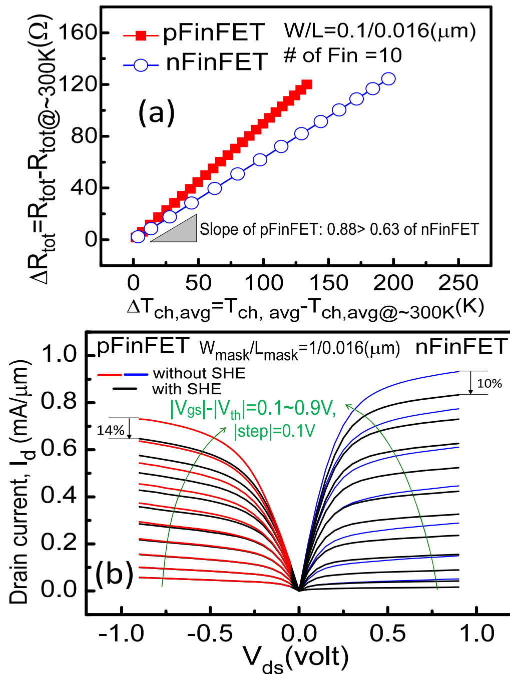

FIGURE 10. (a) The incremental channel total resistance is linearly proportional to the incremental channel local temperature, whose slope can be used to judge the degree of SHE on the devices. (b) The I_dV_{ds} characteristics of n- and pFinFET with (black curves) and without SHE(blue and red curves).

more serious than that in nFinFET. Since we have already known the amount of temperature generated in the channel of FinFET through R_{th} in Fig. 10(a), and R_{tot} was increased by this generated channel temperature, then we can obtain the characteristics of I_dV_{ds} without being suffered from SHE, as shown in Fig. 10(b). 10% decrease of I_dV_{ds} in nFinFET and 14% in pFinFET have been observed due to SHE. To benchmark our experimental results, the values of 11% and 10.8% for $I_{d,sat}$ degradation for 14nm n- and pFinFET have been reported in [29]. The authors used high speed(200ns) pulsed I-V measurement to exclude the self-heating effect so as to obtain SHE-free I-V characteristics. Since, the S/D of pFinFET in [29] is not fabricated by SiGe, the I_dV_{ds} degradation of pFinFET suffered from SHE is comparable to that of nFinFET, which is different from our results. But what is the exact physical mechanism behind this I_dV_{ds} degradation through SHE ? To answer this question, the drain current of FinFET can be governed by the quasi-ballistic transport, [30]

$$I_d = C_{ox} W V_{sat} B_{sat,q} \left(V_{gs} - V_{th}\right). \tag{8}$$

C_{ox} is the inversion capacitance of gate dielectrics; V_{sat} is the saturation velocity; $B_{sat,q}$ is the ballistic coefficient. Because the V_{gs} of I_dV_{ds} in Fig. 10(b) has been normalized by the V_{th}, the term, V_{gs}-V_{th}, in (8), can be ignored. Moreover, owing to the DIBL, the channel barrier has been shifted to the source side under the condition of the driving voltage, which is far from the thermal source of SHE. Thus, $B_{sat,q}$ can be also neglected. Finally, I_d is proportional to V_{sat}, which can be modeled by:

$$V_{sat}\left(T_{ch,avg}\right) = \frac{V_{sat}(300K)}{(1-A) + A\left(T_{ch,avg}/300K\right)} \propto I_d. \tag{9}$$

A is the temperature coefficient. By using Eq. (9), A can be extracted from Fig. 10(b), 0.27 for nFinFET; 0.22 for pFinFET. To benchmark the value of A, 0.26 has been reported in [31] and [32] for electron saturation velocity in intrinsic Si bulk, which is very close to our experimental data, 0.27. Therefore, it is confident for us to claim that the degradation of drain-current is caused by the decay of saturation velocity at a higher temperature due to SHE. On the other hand, in comparison of the A for pFinFET, there is a wide range of A form 0.06 [33], 0.1[32], to 0.37 [34] in previous experimental results, of which our result, 0.22, is in the range. The exact value of A in pFinFET is still an open question, which needs more experimental evidences.

V. SUMMARY AND CONCLUSION

In summary, a new methodology based on RTN measurement has been developed successfully to laterally profile the channel local temperature. By taking advantage of the proximity of traps to the channel from the RTN measurement, τ_e of RTN traps can be utilized to extract the local temperature along the channel. Results have shown that the highest local temperature is near the drain side, and the profile obeys Fourier' law of the thermal conduction very well. Furthermore, the local temperature of pFinFET is about 170K higher than that of nFinFET because a worse thermal conduction material, SiGe, in S/D of pFinFET resists heat dissipation from the lateral direction. Next, the thermal resistance, R_{th}, has been analyzed. The results showed that as the channel length is shrunk, R_{th} increases sharply, in terms of a noticeable short channel effect. To examine the impact of SHE on electrical characteristics of FinFET, it was found that the incremental channel resistance is proportional to the incremental channel averaged temperature, whose slope reflects the degree of SHE. The results double-confirmed that pFinFET exhibits a worse SHE than nFinFET. Finally, we have obtained the characteristics of I_D-V_{DS} with and without SHE experimentally. As far as SHE is concerned, the drain current decays 10% and 14% for n- and pFinFET respectively. And the degradation can be well explained by the degeneration of saturation velocity at high temperature, which can be represented by a temperature coefficient, A. The values of A in our work is 0.27 and 0.22 for n- and p-FinFET, and the former is very close to the previous published data [31], [32]. In addition, the self-heating effect degrades seriously the device mobility and saturation velocity. This further aggravates the noise margin of SRAM [35] which is crucial to the design of FinFET. This work provides valuable information on the understanding of SHE-resistant design of future generation FinFET and can be possibly extended to the study of future nanowires.

REFERENCES

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

[2] G. E. Moore, "Cramming more components onto integrated circuits," *Proc. IEEE*, vol. 86, no. 1, pp. 82–85, Jun. 1998.

[3] D. Hisamoto *et al.*, "FinFET-a self-aligned double-gate MOSFET scalable to 20 nm," *IEEE Trans. Electron Devices*, vol. 47, no. 12, pp. 2320–2325, Dec. 2000.

[4] C.-H. Jan *et al.*, "A 22nm SoC platform technology featuring 3-D tri-gate and high-k/metal gate, optimized for ultra low power, high performance and high density SoC applications," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, Dec. 2012, pp. 3.1.1–3.1.4.

[5] S. Natarajan *et al.*, "A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a $0.0588\mu m^2$SRAM cell size," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, Dec. 2014, pp. 3.7.1–3.7.3.

[6] C. Auth *et al.*, "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 *Proc. IEEE Int. Elecron Devices Meeting (IEDM)*, San Francisco, CA, USA, Dec. 2017, pp. 29.1.1–29.1.4.

[7] F. Menges *et al.*, "Temperature mapping of operating nanoscale devices by scanning probe thermometry," *Nat. Commun.*, vol. 7, Mar. 2016, Art. no. 10874.

[8] T. Takahashi, T. Matsuki, T. Shinada, Y. Inoue, and K. Uchida, "Direct evaluation of self-heating effects in bulk and ultra-thin BOX SOI MOSFETs using four-terminal gate resistance technique," *IEEE J. Electron Devices Soc.*, vol. 4, no. 5, pp. 365–373, Sep. 2016.

[9] J. Franco, B. Kaczer, and G. Groeseneken, "Poly-Si heaters for ultra-fast local temperature control of on-wafer test structures," *Microelectron. Eng.*, vol. 114, pp. 47–51, Feb. 2014.

[10] E. Bury *et al.*, "Experimental validation of self-heating simulations and projections for transistors in deeply scaled nodes," in *Proc. IEEE Int. Rel. Phys. Symp.*, Waikoloa Village, HI, USA, Mar. 2014, pp. XT.8.1–XT.8.6.

[11] M. Tledano-Luque *et al.*, "Correlation of single trapping and detrapping effects in drain and gate currents of nanoscaled nFETs and pFETs," in *Proc. IEEE Int. Rel. Phys. Symp.*, Anaheim, CA, USA, Apr. 2012, pp. XT.5.1–XT.5.6.

[12] P. Ren *et al.*, "New observations on complex RTN in scaled high-*k*/metal-gate MOSFETs—The role of defect coupling under DC/AC condition," in *Proc. IEEE Int. Elecron Devices Meeting (IEDM)*, Washington, DC, USA, Dec. 2013, pp. 31.4.1–31.4.4.

[13] F. M. Puglisi, A. Padovani, L. Larcher, and P. Pavan, "Random telegraph noise: Measurement, data analysis, and interpretation," in *Proc. IEEE Symp. Phys. Failure Anal. Integr. Circuits (IPFA)*, Chengdu, China, Jun. 2017, pp. 1–9.

[14] S. Guo, R. Wang, D. Mao, Y. Wang, and R. Huang, "Anomalous random telegraph noise in nanoscale transistors as direct evidence of two metastable states of oxide traps," *Nat. Sci. Rep.*, vol. 7, no. 1, p. 6239, 2017.

[15] T. Grasser *et al.*, "Switching oxide traps as the missing link between negative bias temperature instability and random telegraph noise," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Baltimore, MD, USA, Dec. 2009, pp. 1–4.

[16] C.-Y. Chen *et al.*, "Correlation of Id- and Ig-random telegraph noise to positive bias temperature instability in scaled high-k/metal gate n-type MOSFETs," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Monterey, CA, USA, Apr. 2011, pp. 190–195.

[17] T. Grasser *et al.*, "A unified perspective of RTN and BTI," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Waikoloa Village, HI, USA, Apr. 2014, pp. 4A.5.1–4A.5.7.

[18] G. Rzepa, M. Waltl, W. Goes, B. Kaczer, and T. Grasser, "Microscopic oxide defects causing BTI, RTN, and SILC on high-k FinFETs," in *Proc. Int. Conf. Simulat. Semicond. Processes Devices (SISPAD)*, Washington, DC, USA, Sep. 2015, pp. 144–147.

[19] F. M. Puglisi, F. Costantini, B. Kaczer, L. Larcher, and P. Pavan, "Monitoring stress-induced defects in HK/MG FinFETs using random telegraph noise," *IEEE Electron Device Lett.*, vol. 37, no. 9, pp. 1211–1214, Sep. 2016.

[20] W. Liu, A. Padovani, L. Larcher, N. Raghavan, and K. L. Pey, "Analysis of correlated gate and drain random telegraph noise in post-soft breakdown TiN/HfLaO/SiO_xnMOSFETs," *IEEE Electron Device Lett.*, vol. 35, no. 2, pp. 157–160, Feb. 2014.

[21] E. R. Hsieh *et al.*, "The experimental demonstration of the BTI-induced breakdown path in 28nm high-k metal gate technology CMOS devices," in *Proc. IEEE Symp. VLSI Tech. (VLSI)*, Honolulu, HI, USA, Jun. 2014, pp. 1–2.

[22] M. H. Liao, C.-P. Hsieh, and C.-C. Lee, " Systematic investigation of self-heating effect on CMOS logic transistors from 20 to 5 nm technology nodes by experimental thermoelectric measurements and finite element modeling," *IEEE Trans. Electron Devices*, vol. 64, no. 2, pp. 646–648, Feb. 2017.

[23] S. S. Chung and C. M. Chang, "The investigation of capture/emission mechanism in high-*k* gate dielectric soft breakdown by gate current random telegraph noise approach," *Appl. Phys. Lett.*, vol. 93, Nov. 2008, Art. no. 213502.

[24] M. Houssa *et al.*, "Trap-assisted tunneling in high permittivity gate dielectric stacks," *J. Appl. Phys.*, vol. 87, no. 12, pp. 8615–8620, 2000.

[25] S.-W. Yoo *et al.*, "Characterizing traps causing random telegraph noise during trap-assisted tunneling gate-induced drain leakage," *Solid State Elecron.*, vol. 109, pp. 42–46, Jul. 2015.

[26] E. R. Hsieh *et al.*, " The understanding of multi-level RTN in tri-gate MOSFETs through the 2D profiling of traps and its impact on SRAM performance: A new failure mechanism found," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, Dec. 2012, pp. 19.2.1–19.2.4.

[27] S. S. Chung *et al.*, "A novel and direct determination of the interface traps in sub-100 nm CMOS devices with direct tunneling regime(12/spl sim/16 A) gate oxide," in *Proc. IEEE Symp. VLSI Tech. (VLSI)*, Honolulu, HI, USA, Jun. 2002, pp. 74–75.

[28] S. Lee *et al.*, "Experimental analysis and modeling of self heating effect in dielectric isolated planar and fin devices," in *Proc. IEEE Symp. VLSI Tech. (VLSI)*, Kyoto, Japan, Jun. 2013, pp. 248–249.

[29] Y. Qu *et al.*, "Ultra fast($<$1ns) electrical characterization of self-heating effect and its impact on hot carrier injection in 14nm FinFETs," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, Dec. 2017, pp. 39.2.1–39.2.4.

[30] E. R. Hsieh *et al.*, "A new and simple experimental approach to characterizing the carrier transport and reliability of strained CMOS devices in the quasi-ballistic regime," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, Baltimore, MD, USA, Dec. 2009, pp. 779–782.

[31] R. Quay, C. Moglestue, V. Palankovski, and S. Selberherr, "A temperature dependent model for the saturation velocity in semiconductor materials," *Mater. Sci. Semicond. Process.*, vol. 3, nos. 1–2, pp. 149–155, 2000.

[32] C. Jacoboni, C. Canali, G. Ottaviani, and A. Quaranta, "A review of some charge transport properties of silicon," *Solid-State Electron.*, vol. 20, no. 2, pp. 77–89, 1977.

[33] J. Becker, E. Fretwurst, and R. Klanner, "Measurements of charge carrier mobilities and drift velocity saturation in bulk silicon of <111> and <100> crystal orientation at high electric fields," *Solid State Electron*, vol. 56, pp. 104–110, 2011.

[34] C. Canali, G. Ottaviani, and A. A. Quaranta, "Drift velocity of electrons and holes and associated anisotropic effects in silicon," *J. Phys. Chem. Solids*, vol. 32, no. 8, pp. 1707–1720, 1971.

[35] E. R. Hsieh *et al.*, "A novel approach to localize the channel temperature induced by the self-heating effect in 14nm high-k metal-gate FinFET," in *Proc. IEEE Electron Devices Technol. Manuf. (EDTM)*, Kobe, Japan, Mar. 2018, pp. 148–150.

E RAY HSIEH (M'10) received the Ph.D. degree in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2016, where he has been a Post-Doctoral Research Fellow with the Department of Electronics Engineering since 2016. He has authored over 40 technical papers and holds six U.S. patents. His current research interests include CMOS logic and nonvolatile memory devices and circuits.

MENG-RU JIANG received the B.Sc. degree in physics from National Chung Hsing University, Taichung, Taiwan, in 2015 and the master's degree from the Department of Electronic Engineering, National Chiao Tung University, Hsinchu, Taiwan, in 2017. His research interests during the course of Master study include advanced semiconductor device development and FinFET with emphasis on self-heating effect. He is currently with TSMC.

JIAN-LI LIN received the master's degree from the Department of Electronics Engineering, National Chiao Tung University, Hsinchu, Taiwan, in 2017. His research interests during the course of Master study include semiconductor manufacturing technology and the design of FinFET. He is currently with the Advanced Technology Development Division, United Microelectronic Corporation, Taiwan.

STEVE S. CHUNG (M'85–SM'95–F'06) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 1985. He is currently an NCTU and UMC Chair Professor with the National Chiao Tung University (NCTU). He has been the first Department Head of EECS Honors Program from 2004 to 2005, the Dean of International Affairs Office, and the Executive Director of school level research center from 2007 to 2008. He was a Visiting Professor with Stanford University in 2001 and 2009, the University of California, Merced, from 2009 to 2010, giving course lectures at both universities in 2009. He has been the Consultant of two world largest IC foundries, TSMC, and UMC. His recent current research areas include CMOS device physics and technology, nonvolatile memory technology, and device reliability/characterization. He was the first speaker (from Taiwan) to present the paper at VLSI Technology symposium in 1995. He presented over 27 times in IEDM/VLSI and is a holder of over 40 patents.

He was a recipient of three times outstanding research awards from the National Science Council and Pan Wen Yuan Award in recognizing his outstanding achievements in the semiconductor research, in 2013. He is a Distinguished PI and a Distinguished NSC Research Fellow from the National Science Council. He is an IEEE Distinguished Lecturer, an Editor of the IEEE JOURNAL OF THE ELECTRON DEVICES SOCIETY and *Applied Physics-A*, an EDS Taipei Chapter Chair, and with past involvements as EDS Board of Governor from 2012 to 2017, an AdCom Member from 2004 to 2009, an EDS Regions/Chapters Chair, and the Editor of EDL from 2002 to 2008.

TSE PU CHEN received the Ph.D. degree in microelectronics engineering from National Cheng Kung University, Tainan, Taiwan. He is currently with the Advanced Technology Development Division, United Microelectronic Corporation, Taiwan. His research is focused on the advanced device development, electrical analysis, and testkey layout design.

SHIH AN HUANG received the master's degree in electrophysics from National Chiao-Tung University, Hsinchu, Taiwan. He is currently with Advanced Technology Development Device Team, United Microelectronic Corporation, Hsinchu, as a Section Manager. His research interests include device development, electrical analysis, test measurement, and testkey layout design.

TAI-JU CHEN received the Ph.D. degree in electronic engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1998. He is focusing on the advanced device development in the Advanced Technology Development Division, United Microelectronic Corporation, Taiwan.

OSBERT CHENG (M'10) received the Ph.D. degree in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, and the master's degree from National Tsing Hua University, Taiwan. He is currently the Associate Vice President of the Advanced Technology Development Division, United Microelectronic Corporation, Taiwan. He has been involved in developing the foundry platform with more than eight technology nodes. He is currently leading the advanced device, SRAM, mixed signal, reliability, and TCAD research and development team.

He is currently a member of the Technical Program Committee of IEEE VLSI Technology Symposium.