Received 1 March 2024; revised 17 April 2024; accepted 19 April 2024. Date of publication 24 April 2024; date of current version 13 May 2024.
The review of this article was arranged by Editor C. Bulucea.

Digital Object Identifier 10.1109/JEDS.2024.3393019

Heat Generation Mechanisms of Self-Heating Effects in SOI-MOS

ZHENG-LAI TANG AND BING-YANG CAO (Member, IEEE)

The Key Laboratory of Thermal Science and Power Engineering of Education of Ministry, Department of Engineering Mechanics, Tsinghua University, Beijing 100084, China

CORRESPONDING AUTHOR: B.-Y. CAO (e-mail: caoby@tsinghua.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant U20A20301, Grant 52327809, Grant 52250273, and Grant 82361138571; in part by the National Key Research and Development Program of China under Grant 2023YFB4404104; and in part by the Beijing Natural Science Foundation under Grant L233022.

ABSTRACT The development of microelectronic devices to the nanoscale intensifies self-heating challenges, affecting efficiency and durability. Understanding the mechanisms of heat generation at this scale is crucial, yet research extending beyond Joule heat remains limited. This paper simulates the self-heating effect of Silicon-On-Insulator Metal-Oxide-Semiconductor Field Effect Transistors (SOI-MOS) at the nanoscale and researches the characteristic and influence of different heat generation mechanisms, including the Joule heat, recombination heat and Peltier-Thomson heat. Our results provide a detailed two-dimensional distribution and intensity of various heat generation mechanisms within the silicon channel layer. It is found that Peltier-Thomson heat has the same magnitude as Joule heat at the nanoscale, and exhibits an alternating distribution pattern of hot and cold sources under the gate. But recombination heat is relatively negligible. The analysis of the influence of different heat mechanisms emphasizes the important role of Joule heat. While the offset effect limits the impact of Peltier-Thomson heat, its significance to device thermal performance should not be ignored. More importantly, this study investigates the impact of characteristic size on different heat generation mechanisms, revealing the size dependence of Peltier-Thomson heat.

INDEX TERMS Self-heating effect, heat generation mechanism, silicon-on-insulator metal-oxide-semiconductor field-effect transistors (SOI-MOS), size dependence.

I. INTRODUCTION

AS the characteristic size of transistors scales down to nanometer, self-heating effects (SHE) have become increasingly important. The temperature rise caused by the SHE enhances the scattering between the electrons and phonons, resulting in degraded electron mobility. Local hot spots can also increase thermal stress within semiconductor materials, potentially reducing device lifespan or causing failure [1]. It is predicted that the 10-15 °C changes of junction temperature will reduce devices lifetime to half [2]. The elevated temperature will also accelerate the process of oxide breakdown, leading to additional reliability issues [3]. High-accuracy electro-thermal simulation yields a complete prediction of device thermal and electrical performance. Exact hot spot temperature estimation is beneficial to design and optimization of devices, improving the thermal management and weakening the impact of SHE [4]. The most common electro-thermal simulation method is based on the drift-diffusion model, solving the Poisson equation and heat conduction equation to acquire the electric field and temperature distribution [5]. In the electro-thermal simulation, choosing an appropriate heat generation model is crucial, as it provides the mathematical expression representing the physical nature of the heat generation in the devices.

In early research, the heat generation in the device was usually attributed to Joule heating. Based on conductors' behavior, initial models proposed a heuristic approach using the product of current density and electric field to represent Joule heating [6]. Revised models were developed to include recombination heat as a contributing factor [7], [8]. The first rigorous heat generation model was provided by Wachutka using the laws of phenomenological irreversible thermodynamics [9]. Lindefelt also presented a model based on Boltzmann equation which considers the interaction between

the applied electric field and carriers [10]. These models illustrate the complexity of heat generation mechanisms in devices. Under steady state, there are two other important types of heat, one is the recombination heat caused by the recombination and generation process, another is the Peltier-Thomson heat resulted by the changes of the carrier's energy. Different heat generation mechanisms exhibit varied characteristics and make distinct impacts on the device. Many simulations often employ simplified expressions for heat generation, potentially overlooking the complexity of various mechanisms [11], [12]. The absence of a comprehensive comparative analysis of these mechanisms might introduce uncertainties in the simulations.

Tornblad et al. analyzed seven terms, including Joule heat and recombination heat, from Lindefelt's formula to assess heat generation in PIN diode under different operating conditions [13]. The results indicate that the peak values of other heat generation mechanisms are at least one order of magnitude smaller than Joule and recombination heat, with the peak widths not exceeding 8%. Gough et al. concluded from simulation of a multi-fingered bipolar power transistor that recombination and Peltier-Thomson heat constituted no more than 10% of total power consumption [14]. However, the studied device structure is too simple and cannot be generalized to other devices. Other studies have similarly explored comparisons among heat generation mechanisms. Leung et al. discovered that in LIGBT, the heat generation of recombination heat reached 10^7 W/cm^3, nearly matching that of Joule heat [15]. Qian et al. demonstrated that in high current and thermal breakdown stages, the effects of recombination and Thomson heat become more pronounced [16]. Muneer et al. presented the thermal behaviors of different types heat in the self-heated nano-crystalline silicon microwires under asymmetric melting, revealing that the maximum electronic-convective heat power increases to five times the Joule heat after melting [17]. Shang and Li stated that under optimal voltage in photovoltaic devices, the combined power consumption from Peltier heat and surface recombination heat equaled half of Joule heat [18]. Although their research considered various heat generation mechanisms in calculations, it didn't provide a detailed comparison of the characteristics and effects of these different types of heat. Whether to consider the complete heat generation mechanism in the simulation has become a subjective judgment of the researchers. More importantly, the devices examined in previously cited studies exhibit characteristic sizes from tens to hundreds of microns. With the urgent demand for integration and device performance, most devices have already entered the nanoscale. Comparative research and analysis of different heat generation mechanisms at the nanoscale are still lacking. Variations in device dimensions may alter internal temperature gradient, thereby affecting the generation of Peltier-Thomson heat and potentially leading to distinct heat generation characteristics. Therefore, it is necessary to analyze the characteristics and influence of different heat generation mechanisms at the nanoscale.

Despite the extensive application of SOI-MOS devices and numerous studies on the self-heating effects in advanced SOI-MOS devices [19], [20], [21], discussions on the underlying heat generation mechanisms are limited. This research provides an improved self-heating effect simulation for Silicon-on-insulator metal-oxide-semiconductor field-effect transistors (SOI-MOS). The main heat generation mechanisms under steady state, including Joule heat, recombination heat and Peltier-Thomson heat, are calculated and analyzed. The results illustrate the magnitude and distribution characteristics of different heat generation mechanisms, demonstrating that the impact of Peltier-Thomson heat on the device is limited by the offset mechanism of its heating and cooling effects. Additionally, the study examines the influence of characteristic size on different heat generation mechanisms, revealing that Peltier-Thomson heat is size-dependent, unlike Joule heat. This article aims to deepen the understanding of different heat generation mechanism of semiconductor devices.

II. METHOD

A. BASIC EQUATIONS

Common electro-thermal simulations depend on the solution of fundamental semiconductor equations. Derived from Maxwell's equations, the Poisson equation and the continuity equations for carriers can be obtained:

$$\begin{aligned} &\nabla \cdot (\varepsilon \cdot \nabla \phi) = -q \cdot (p - n + N_D - N_A) \\ &\frac{1}{q} \nabla \cdot \vec{J}_n - \frac{\partial n}{\partial t} = R \\ &\frac{1}{q} \nabla \cdot \vec{J}_p + \frac{\partial p}{\partial t} = -R \end{aligned} \tag{1}$$

where ε is the dielectric constant, ϕ is the potential, q is the elementary charge, p and n are the concentration of holes and electrons, N_D and N_A are the concentration of donors and acceptors, $\vec{J}_n$ and $\vec{J}_p$ are the current density of electrons and holes, t is the time, and R is the net recombination rate. There are five unknown variables while only three equations exist, so it is necessary to replenish the drift-diffusion model which give the transport equations of carriers [22]:

$$\begin{aligned} \vec{J}_n &= q \cdot n \cdot \mu_n \cdot \vec{E} + q \cdot D_n \cdot \nabla n \\ \vec{J}_p &= q \cdot p \cdot \mu_p \cdot \vec{E} - q \cdot D_p \cdot \nabla p \end{aligned} \tag{2}$$

where μ_n and μ_p are the mobility of electrons and holes, $\vec{E}$ is the electric field, coefficient D_n and D_p are determined by the Einstein relation [23]. If only the isothermal simulation is considered, the above equations are complete due to a constant lattice temperature. However, when including the self-heating effect, it is necessary to introduce the heat conduction equation:

$$\nabla \cdot (\kappa \nabla T) + H = 0 \tag{3}$$

where κ is the thermal conductivity, T is the lattice temperature, and H is the model of heat generation. It is noticed that, as the size of devices diminishes, approaching the mean

free path of phonons, the heat conduction process no longer conforms to Fourier's heat conduction law. This situation calls for alternative methods such as Monte Carlo simulations to accurately model phonon transport [24], [25]. This paper mainly concentrates on the comparison of different heat generation mechanism, instead of the heat conduction process. Additionally, some research has shown that in short-channel devices, the localized hot spot temperature rise due to phonon ballistic transport negligibly affects device performance [13]. Therefore, employing the heat conduction equation as an efficient method for solving lattice temperature is considered acceptable.

B. PHYSICAL MODELS

For simulation, most physical parameters need accurate models. In this paper, the mobility model employed is the Klaassen low field model, widely used in MOSFET simulation. It considers the impacts of lattice scattering, impurity scattering, carrier-carrier scattering, and impurity clustering effects at high concentration [26], [27]. The mathematical expression of the Klaassen mobility model is:

$$\frac{1}{\mu_b} = \frac{1}{\mu_L} + \frac{1}{\mu_I} \tag{4}$$

Among the (4), μ_b is the mobility of the bulk material, μ_L is the carrier mobility caused by lattice scattering, μ_I is the carrier mobility caused by impurities and carrier scattering. The specific calculation formula is given in [26], [27]. The Klaassen model, being a low field model, is not well-suited for high field conditions where drift velocity saturation occurs [28]. In such conditions, the increase in the electric field intensifies carrier scattering, consequently degrading mobility and device performance. This effect of velocity saturation is more pronounced in short-channel devices due to higher electric fields [29]. In this study, the velocity saturation is modeled using the Hansch equation [30]:

$$\mu = \frac{2\mu_0}{1 + \left(1 + 4\left(\frac{\mu_0 F}{v_{sat}}\right)^2\right)^{0.5}} \tag{5}$$

where μ_0 is the low field mobility depending on the Klaassen model, v_{sat} is the saturation velocity of carrier, F is the driving force. The chosen of driving force is not unique. Early work use the magnitude of electric field $|\vec{E}|$, but some research recommend using the gradient of quasi-Fermi potential $|\nabla\phi_{n(p)}|$ [31]. Although many parameters influence the saturation velocity, in this work, only the lattice temperature is considered [32].

The carrier concentration maintains dynamic equilibrium by the recombination and generation process, considering the SRH recombination only [33], [34]:

$$R^{SRH} = \frac{n \cdot p - n_i^2}{\tau_p \cdot (n + n_i) + \tau_n \cdot (p + n_i)} \tag{6}$$

where τ_n and τ_p is the carrier lifetime.

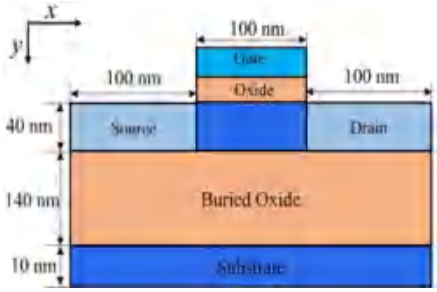

FIGURE 1. **The schematic of SOI-MOS structure.**

In the thermal simulation part, thermal conductivity is described using a model of the form of a power law:

$$\kappa(T) = \kappa_{ref} \cdot \left(\frac{T}{300}\right)^{\beta} \tag{7}$$

where κ_{ref} is the thermal conductivity at referred temperature, which is equal to 1.48 W/cm·K for silicon and 0.014 W/cm·K for SiO_2 at 300 K. The power exponent β is set as -1.65 [35]. Heat generation model determines the accuracy of the solution of thermal performance. The steady state model of Wachutka's derivation is used:

$$\begin{aligned} H = &\frac{\left|\vec{J}_n\right|^2}{q\mu_n n} + \frac{\left|\vec{J}_p\right|^2}{q\mu_p p} \\ &+ q\left(R' - G'\right)\left[\phi_p - \phi_n + T\left(P_n + P_p\right)\right] \\ &- qT\left(\vec{J}_n \cdot \overrightarrow{\nabla P_n} - \vec{J}_p \cdot \overrightarrow{\nabla P_p}\right) \end{aligned} \tag{8}$$

where R' and G' are the recombination and generation rate. P_n and P_p are the absolute thermoelectric powers for electrons and holes. The first term represents Joule heating, the second term accounts for recombination and generation heating or cooling, and the last term corresponds to the Peltier and Joule-Thomson effects. In this work, the last term is referred to as Peltier-Thomson heat, or simply PT heat.

C. SIMULATION SETTING

A layer of silicon dioxide is inserted between the substrate and the top silicon of the traditional MOS transistors, which is Silicon-On-Insulator Metal-Oxide-Semiconductor Field Effect Transistor (SOI-MOS). The simulated structure of SOI-MOS is as shown in Fig. 1. The gate length is 100 nm, and a 40 nm thick silicon channel layer is sandwiched between the gate oxide and a buried oxide layer 140 nm. The thickness of gate oxide is 2 nm. Let the lateral direction from source to drain be the x-axis, and the direction from the gate to the substrate be the y-axis. The donor doping concentration of the source and drain regions in the calculation is $N_D = 1\times10^{20}$ cm^{-3}, and the acceptor doping concentration of the bulk silicon layer is $N_A = 1\times10^{17}$ cm^{-3} [36]. Here, the lattice temperature of substrate is set as $T_0 = 300$ K. In actual transistors, the substrate thickness exceeds 10 nm. However, for simulation purposes, it only need solve the Poisson equation. Additionally, due to the high

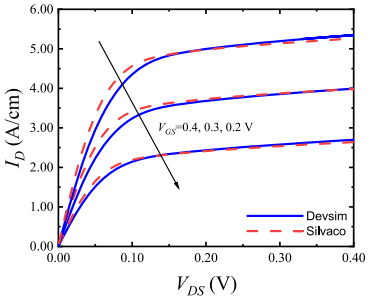

FIGURE 2. **Comparison of I-V output characteristic curves of SOI-MOS calculated by Devsim and Silvaco (V_{GS} = 0.2, 0.3, 0.4 V).**

thermal conductivity of silicon, heat is effectively transferred to the buried oxide layer through the isothermal boundary below. Therefore, the substrate's impact can be completely disregarded in the simulation. The simulation calculations in this article come from a program developed on the basis of Devsim, an open source TCAD software with higher flexibility and expandability [37], [38]. In addition, this paper only considers two-dimensional simulations, which may lead to an overestimation of the self-heating effect within the device. Nevertheless, since the main focus of this article is on different heat generation mechanisms, this omission does not affect the final conclusions.

III. RESULTS AND DISCUSSION

A. VERIFICATION FOR THE SIMULATION

The simulation program in this work is verified by commercial software Silvaco TCAD. The output characteristic curves for different V_{GS} are given in Fig. 2. The blue lines represent the results calculated by Devsim, while the red dash lines come from the Silvaco. In our simulation, only electron mobility and thermal conductivity were considered as temperature-dependent models, with others set as constant. Due to the convergence problem, the Caughey and Thomas velocity saturation model in Silvaco is not employed [39]. Although these simplifications led to discrepancies in the IV curves, the simulation results still demonstrate the program's ability to replicate the physical properties of SOI-MOS devices. Furthermore, the trend of the I-V curve and the hotspot location under the gate close to the drain side are consistent with experimental results [40], [41].

B. ANALYSIS OF HEAT GENERATION CHARACTERISTICS

As presented in the (8), the heat generation in the device includes the Joule heat, recombination heat and Peltier-Thomson heat. This part analysis primarily focuses on their distribution and magnitude within the channel layer, as other regions, such as the buried oxide, do not significantly contribute to heat generation. The results presented here are specifically calculated under the condition of $V_{GS} = 0.5$ V and $V_{DS} = 1.0$ V to ensure a pronounced self-heating effect. Importantly, while these results are derived under these specific condition, the heat generation characteristics identified

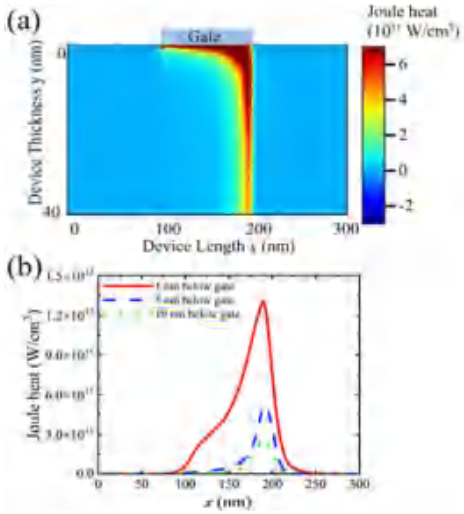

FIGURE 3. **The Joule heat distribution in (a) channel layer and (b) different thickness below the gate (V_{GS} = 0.5 V, V_{DS} = 1.0 V).**

are relevant to other reasonable operating conditions as well. The main variation under different conditions would likely be the scale of heat generation.

First, the 2D distribution of Joule heat in the channel layer is given in Fig. 3 (a). The result reveals that Joule heating is primarily concentrated in the conductive channel below the gate and at the interface of the right PN junction near the drain. Furthermore, these regions exhibit heat generation levels exceeding 10^{11} W/cm^3. The Joule heat is proportional to the current density, so the conductive channel is the main heat generation area in MOS. The distribution in different thickness below the gate is given in Fig. 3 (b). Due to its high carrier density, the thin channel layer formed by the gate voltage results in significantly higher Joule heat near the gate. The red line shows that the predominant heat generation distribution in the channel is significantly larger and more uniform, with the peak heat generation density reaching approximately 1.3×10^{12} W/cm^3. This peak is located at $x =$ 200 nm. The area beneath the gate and close to the drain is a common hot spot in the device [40]. This occurs because the high electric field present in this region increases the rate of carrier scattering, rather than further accelerating the drift speed, which is already at saturation. This heightened scattering under the strong electric field is the primary cause of increased heat generation. Additionally, the Joule heat at the interface of the right PN junction near the drain is also noteworthy. It is attributed to the high electric field in this area, as well as the significant diffusion current that should not be overlooked.

The heat generation due to electron-hole recombination and generation processes is depicted in Fig. 4. Unlike Joule heating, which exclusively results in positive heat generation, the recombination heat in the process of electron-hole recombination is more complex. Recombination process releases energy, leading to positive heat generation, while the generation of electrons and holes absorbs energy, thereby

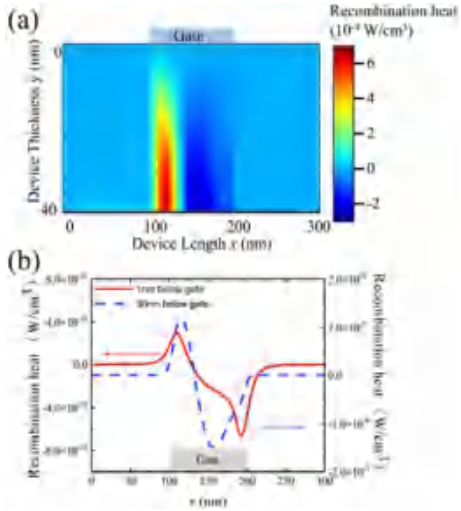

FIGURE 4. **The recombination heat distribution in (a) channel layer and (b) different thickness below the gate (V_{GS} = 0.5 V, V_{DS} = 1.0 V).**

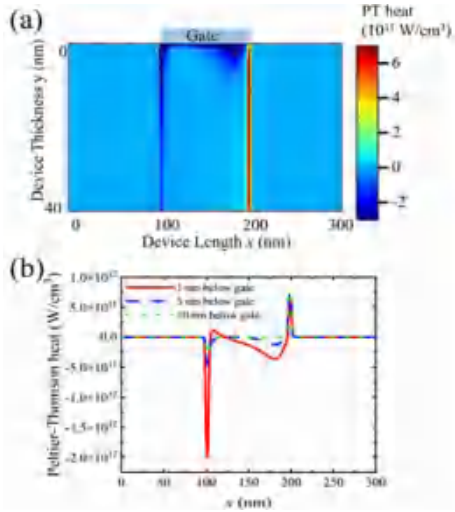

FIGURE 5. **The Peltier-Thomson heat distribution in (a) channel layer and (b) different thickness below the gate (V_{GS} = 0.5 V, V_{DS} = 1.0 V).**

causing negative heat generation. Fig. 4 (a) presents the two-dimensional distribution of recombination heat within the silicon channel layer of the MOSFET. As MOSFETs are majority carrier devices, the magnitude of recombination heat is minimal, approximately 10^{-9} W/cm^3, essentially negligible compared to Joule heat. Although the magnitude of recombination heat is relatively small, its characteristics of heat generation still deserve to further analysis. As (6) illustrates, recombination heat is directly proportional to the rate of recombination, which in turn depends on carrier concentration. The application of a positive gate voltage causes holes to move towards the bottom of the channel layer, intensifying the recombination and generation processes. This spatial variation in recombination heat, particularly its increase at the bottom of the channel layer, is detailed in Fig. 4 (b). It's important to note that in some minority carrier devices, recombination heat can be more pronounced and cannot be overlooked [15].

The last term in (8) represents Peltier-Thomson (PT) heat. Wachutka described this term as the energy exchange between the host lattice and carriers when particles flow through regions with spatially varying thermoelectric powers P_n and P_p [10]. In comparison with Joule heating, PT heat can be conceptualized as a modification in scattering events due to temperature variations occurring during the movement of carriers. The two-dimensional PT heat distribution in the channel layer, shown in Fig. 5 (a), exhibits both heating and cooling effects beneath the gate. Primarily generated in the conductive channel and at the PN junction interface, PT heat's sign is dictated by the alignment of current flow and temperature gradients. For instance, a negative PT heat is produced in the channel below the gate where the electron current flows from drain to source against a higher temperature at the drain side. Contrary to previous studies where PT heat is considered negligible [13], [14], our results reveal a PT heat magnitude of approximately 10^{11} W/cm^3, which is comparable to the magnitude of Joule heat. Detailed variations in PT heat across different depths below the gate are displayed in Fig. 5 (b). At 1 nm beneath the gate, there is an alternating distribution pattern of cooling and heating sources in the channel. The absolute peak of the cooling sources reaches about 2.0×10^{12} W/cm^3 at x = 100 nm, surpassing the heat source of 7.3×10^{11} W/cm^3 at x = 200 nm. As depth increases, cooling effects diminish, while heating effects remain consistent. At 10 nm below the gate, cooling sources drop to 2.0×10^{11} W/cm^3 at x = 100 nm. This variation in cooling and heating intensity is primarily due to changes in current density, given the minor temperature gradient changes in the device thickness (y) direction.

C. INFLUENCE ON THE DEVICE THERMAL AND ELECTRICAL PERFORMANCE

The different characteristic of heat generation determines their influence of the device thermal and electrical performance. However, recombination heat is excluded from this part of the analysis, as its heat generation is extremely low, almost negligible, contributing insignificantly to the overall performance of the device. The direct influence of heat generation is the rise in the device temperature. As shown in Fig. 6 (a), the two-dimensional temperature distribution within the silicon channel layer, when considering only Joule heat, indicates the highest hotspot temperature (T_{max}) of 437.5 K. This peak temperature is located right below the gate at x = 200 nm, exactly where the maximum Joule heat is generated. Due to the buried oxide's low thermal conductivity, the Joule heat generated is not efficiently dissipated, leading to its spread across the entire channel layer and resulting in an overall temperature increase. Including PT heat in the simulation, as illustrated in Fig. 6 (b), results in a further rise in temperature within the channel layer. PT heat does not significantly change the temperature

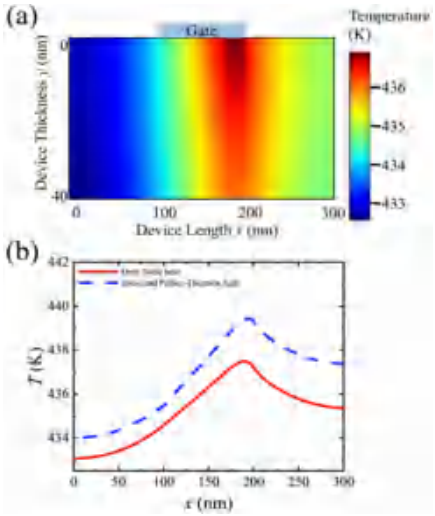

FIGURE 6. **(a) 2D lattice temperature considering Joule heat only; (b) Temperature comparison at 1 nm below the gate (V_{GS} = 0.5 V, V_{DS} = 1.0 V).**

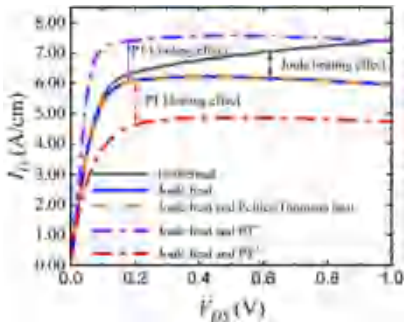

FIGURE 7. **The drain current versus the gate voltage curve considering different heat generation terms.**

field distribution pattern. Therefore, the two-dimensional temperature distribution under the combined effect of Joule and PT heat is not presented here. Although PT heat exhibits both cooling and heating effects, its net impact on temperature is positive. This is because its heating effect is stronger than cooling effect in total, which can be speculated in Fig. 5 (b). After considering the PT heat, the maximum temperature changes from 437.5 K to 439.5 K.

The self-heating effect in the device, causing a rise in temperature, enhances carrier-lattice scattering, thereby reducing electrical performance. Fig. 7 illustrates the impact of both Joule and PT heat on the device I-V output curves. To gain a deeper understanding of PT heat, its individual cooling (PT^-) and heating (PT^+) effects are analyzed. The definition of PT^- and PT^+ are:

$$\begin{cases} PT^+ = \frac{PT+\mathbf{abs}(PT)}{2} \\ PT^- = \frac{PT-\mathbf{abs}(PT)}{2} \end{cases} \tag{9}$$

In the simulation, setting the lattice temperature as a constant ($T = T_0 = 300$ K) represents the isothermal condition. When comparing these results with the non-isothermal simulation, which accounts for Joule heating, a significant drop in the drain current density I_D is observed. Besides, the current density degradation begins intensively with higher drain voltage V_{DS}, as a result of higher electric field and enhanced carrier collisions [42]. Despite the overall temperature increase in the channel due to PT heat, the current curve, as shown by the orange dash line in Fig. 7, remains largely unaffected. This is because the carrier mobility has already been significantly reduced by the high temperatures resulting from Joule heating, making the additional temperature rise from PT heat less impactful on the device performance. While the overall impact of PT heat is slightly, this does not imply that the heat generation of PT heat is insignificant. As demonstrated by the purple and red dash-dot lines in Fig. 7, the individual effects of PT heat on current density are pronounced. This suggests that the minor overall impact arises not from low levels of heat generation, but from the alternating cooling and heating effects of PT heat, which lead to an offsetting effect on channel temperature. This competitive interaction between heating and cooling sources limits the impact of PT heat on the device electrical performance.

While temperature is an indicator of self-heating effect strength within the device, it does not provide a complete picture of thermal performance. Hotspot temperatures is crucial for predicting device reliability and lifespan. The mean-time-to-failure (MTTF), a key metric for device lifespan, is calculated using Black's formula [43]:

$$\mathrm{MTTF} = AJ^{-n}\exp\left(\frac{E_a}{kT}\right) \tag{10}$$

where A is a constant, J is the current density, n is the current density acceleration factor, E_a is the activation energy, T is the temperature at working condition. To quantify the influence of PT heat, we utilize parameters from previous SOI-CMOS reliability studies, with an activation energy (E_a) of 1.386 eV. Since PT heat has minimal effect on current density, the primary variable altered by PT heat is the hotspot temperature. Using the T = 439.5 K and T = 437.5 K in equation, it reveals that the device lifespan can be misestimated by up to 15% if ignore the PT heat. However, it's important to note that this calculation is a simplification, as the parameters in Black's formula are highly dependent on real conditions. This underscores the importance of considering PT heat when evaluating device reliability under self-heating effects.

D. INFLUENCE OF CHARACTERISTIC SIZE TO HEAT GENERATION

Simulations for 100 nm gate length SOI-MOS devices indicate larger Peltier-Thomson heat generation, differing from previous research on micron-scale devices. Therefore, exploring the influence of characteristic size on various heat generation mechanisms is necessary. As SOI-MOS devices are a type of field-effect transistors, their scaling

TABLE 1. **Parameter settings under device scaling down.**

Parameter	Origin	Scaling (1/4)	Scaling (1/16)
Gate length (L_g)	400 nm	100 nm	25 nm
Device length (L)	1200 nm	300 nm	75 nm
Si thickness	160 nm	40 nm	10 nm
Oxide thickness	32 nm	8 nm	2 nm
BOX thickness	640 nm	160 nm	40 nm
N_A	2.5×10^{16} cm^{-3}	1×10^{17} cm^{-3}	4×10^{17} cm^{-3}
N_D	2.5×10^{19} cm^{-3}	1×10^{20} cm^{-3}	4×10^{20} cm^{-3}
V_{DS}	4 V	1 V	0.25 V
V_{GS}	4 V	1 V	0.25 V

down in characteristic size follows the constant electric field law, ensuring unchanged performance [44]. Three conditions with different characteristic sizes are compared, and the corresponding device structure parameters are listed in Table 1. Furthermore, the temperature boundary condition for this simulation is set to gate heat dissipation ($T_{gate} = T_0 = 300$ K). This approach is adopted due to the overheating issues caused by the thick buried oxide in devices with larger characteristic sizes and higher voltage bias. As the gate length approaches the order of nanometers, the ballistic transport of electrons becomes significant, rendering the drift-diffusion model less applicable. The influence of electron ballistic transport on heat generation is evident in the shift and reduction of the peak heat generation [12]. Although electron Monte Carlo simulations offer a more detailed physical description of electron movement and scattering processes, the calculations of heat generation primarily rely on the statistics of phonon emission and absorption, making the distinction between different heat generation mechanisms challenging. Therefore, employing the drift-diffusion model for a qualitative exploration of the influence of characteristic size on different heat generation mechanisms is a valid approach.

Fig. 8 illustrates the distribution of Joule and PT heat generation at 1 nm below the gate, under the constant electric field law. Theoretically, devices with different feature sizes should exhibit similar current densities. As depicted in Fig. 8 (a), the magnitude of Joule heat remains consistent across different characteristic sizes, which verified the similar device performance in different size. However, PT heat varies significantly with characteristic size, with smaller gate lengths resulting in higher heat generation in Fig. 8 (b). Although the current density also determines the PT heat, the significant magnitude changes cannot only come from the difference between the current density.

The primary difference between PT heat and Joule heat lies in PT heat strong dependence on the device internal temperature gradient. Fig. 9 shows the temperature gradient across various characteristic sizes. The $\alpha\Delta T$ is defined as $\alpha(T\text{-}T_{min})$, where α represents the scaling rate and is set to 1, 1/4, and 1/16 for different cases. T_{min} is the minimum

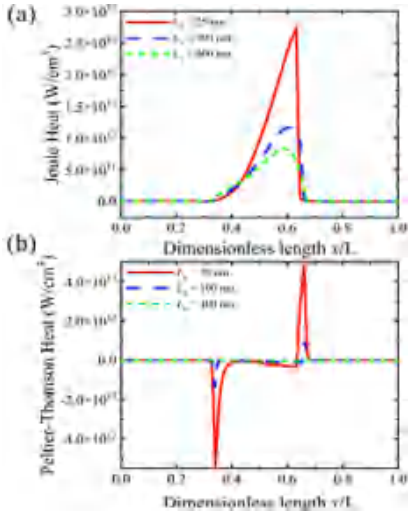

FIGURE 8. **The heat distribution of different characteristic size at 1 nm below the gate for (a) Joule heat and (b) PT heat.**

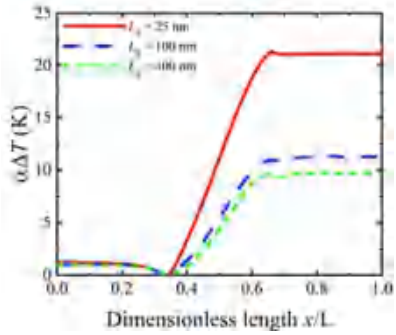

FIGURE 9. **The defined $\alpha\Delta T$ distribution of different characteristic size device at 1 nm below the gate.**

temperature, which located at $x = 100$ nm due to the cooling effect of PT heat. In non-dimensional coordinates, the slope in Fig. 9 directly represents the change of the internal temperature gradient. The steeper red line for $L_g = 25$ nm, compared to the blue dashed and green short-dashed lines, verifies the largest PT heat generation in devices with a 25 nm gate length. Analyzing the temperature gradient uncovers the intrinsic relationship between PT heat and device feature size: a shrinking device size results in a rapidly increasing temperature gradient, leading to stronger PT heat generation. This underscores the size-dependence of PT heat. However, this scaling approach results in smaller-scale SOI devices utilizing higher-doped silicon, which does not align with actual device configurations. Therefore, these findings are merely predictive.

IV. CONCLUSION

This study systematically investigates the self-heating effect in SOI-MOS devices using an electro-thermal simulation method based on the drift-diffusion model. We conduct a detailed comparison of the characteristics and effects of three key heat generation mechanisms: Joule heat, recombination

heat, and Peltier-Thomson heat. Our results indicate that Joule heat is the predominant source of power dissipation, significantly impacting the device temperature and performance. Additionally, the study finds that the impact of recombination heat in SOI-MOS devices is negligible. Peltier-Thomson heat, whose magnitude is similar to Joule heat, cannot be overlooked. While its alternating heating and cooling effects under the gate largely cancel each other out, minimally impacting device electrical performance, it significantly influences device reliability and lifespan.

Furthermore, our analysis of heat generation across different characteristic sizes uncover the size dependence of Peltier-Thomson heat. Our findings indicate that the impact of Peltier-Thomson heat becomes more pronounced in devices with gate lengths under 100 nm. This underscores the necessity of considering the full spectrum of heat generation mechanisms in simulations, particularly when focusing on thermal performance for devices with smaller characteristic sizes. This work provides further insights into the self-heating effects in SOI-MOS devices, potentially aiding in the optimization and design of more effective thermal management strategies.

REFERENCES

[1] K. Ranjan, S. Arulkumaran, G. I. Ng, and A. Sandupatla, "Investigation of self-heating effect on DC and RF performances in AlGaN/GaN HEMTs on CVD-diamond," *IEEE J. Electron Devices Soc.*, vol. 7, pp. 1264–1269, 2019, doi: 10.1109/JEDS.2019.2947564.

[2] R. Viswanath, V. Wakharkar, A. Watwe, and V. Lebonheur, "Thermal performance challenges from silicon to systems," *Intel Technol. J.*, vol. 4, pp. 1–16, Sep. 2001.

[3] B. Kaczer, R. Degraeve, N. Pangon, and G. Groeseneken, "The influence of elevated temperature on degradation and lifetime prediction of thin silicon-dioxide films," *IEEE Trans. Electron Devices*, vol. 47, no. 7, pp. 1514–1521, Jul. 2000, doi: 10.1109/16.848301.

[4] D.-S. Tang and B.-Y. Cao, "Phonon thermal transport and its tunability in GaN for near-junction thermal management of electronics: A review," *Int. J. Heat Mass Transf.*, vol. 200, Jan. 2023, Art. no. 123497, doi: 10.1016/j.ijheatmasstransfer.2022.123497.

[5] M. G. Ancona, "Modeling of thermal phenomena in GaN devices," in *Thermal Management of Gallium Nitride Electronics*. Amsterdam, The Netherlands: Elsevier, 2022, pp. 165–184, doi: 10.1016/B978-0-12-821084-0.00015-9.

[6] S. P. Gaur and D. H. Navon, "Two-dimensional carrier flow in a transistor structure under nonisothermal conditions," *IEEE Trans. Electron Devices*, vol. 23, no. 1, pp. 50–57, Jan. 1976, doi: 10.1109/T-ED.1976.18346.

[7] A. Chryssafis and W. Love, "A computer-aided analysis of one-dimensional thermal transients in n-p-n power transistors," *Solid-State Electron.*, vol. 22, no. 3, pp. 249–256, Mar. 1979, doi: 10.1016/0038-1101(79)90029-7.

[8] M. S. Adler, "Accurate calculations of the forward drop and power dissipation in thyristors," *IEEE Trans. Electron Devices*, vol. 25, no. 1, pp. 16–22, Jan. 1978, doi: 10.1109/T-ED.1978.19026.

[9] G. K. Wachutka, "Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling," *IEEE Trans. Comput. Aided Design Integr. Circuits Syst.*, vol. 9, no. 11, pp. 1141–1149, Nov. 1990, doi: 10.1109/43.62751.

[10] U. Lindefelt, "Heat generation in semiconductor devices," *J. Appl. Phys.*, vol. 75, no. 2, pp. 942–957, Jan. 1994, doi: 10.1063/1.356450.

[11] E. Pop, "Energy dissipation and transport in nanoscale devices," *Nano Res.*, vol. 3, pp. 147–169, Mar. 2010, doi: 10.1007/s12274-010-1019-z.

[12] X. Chen, S. Boumaiza, and L. Wei, "Self-heating and equivalent channel temperature in short gate length GaN HEMTs," *IEEE Trans. Electron Devices*, vol. 66, no. 9, pp. 3748–3755, Sep. 2019, doi: 10.1109/TED.2019.2926742.

[13] O. Tornblad, U. Lindefelt, and B. Breitholtz, "Heat generation in Si bipolar power devices: The relative importance of various contributions," *Solid-State Electron.*, vol. 39, no. 10, pp. 1463–1471, Oct. 1996, doi: 10.1016/0038-1101(96)00056-1.

[14] P. A. Gough, P. Walker, and K. R. Whight, "Electro-thermal simulation of power semiconductor devices," in *Proc. ISPSD*, 1991, pp. 89–94.

[15] Y.-K. Leung, A. K. Paul, K. E. Goodson, J. D. Plummer, and S. S. Wong, "Heating mechanisms of LDMOS and LIGBT in ultrathin SOI," *IEEE Electron Device Lett.*, vol. 18, no. 9, pp. 414–416, Oct. 1997, doi: 10.1109/55.622514.

[16] Q. Qian, W. Sun, J. Zhu, and L. Shi, "Investigation of the shift of hot spot in lateral diffused LDMOS under ESD conditions," *Microelectron. Reliab.*, vol. 50, no. 12, pp. 1935–1941, Dec. 2010, doi: 10.1016/j.microrel.2010.05.010.

[17] S. Muneer, G. Bakan, A. Gokirmak, and H. Silva, "Incorporation of GTR (generation–transport–recombination) in semiconductor simulations," *J. Appl. Phys.*, vol. 129, no. 5, Feb. 2021, Art. no. 055702, doi: 10.1063/5.0037411.

[18] A. Shang and X. Li, "Photovoltaic devices: Opto-electro-thermal physics and modeling," *Adv. Mater.*, vol. 29, Feb. 2017, Art. no. 1603492, doi: 10.1002/adma.201603492.

[19] R. Carter et al., "22nm FDSOI technology for emerging mobile Internet-of-Things and RF applications," in *Proc. IEEE Int. Electron Devices Meeting (IEDM)*, 2016, pp. 2.2.1–2.2.4, doi: 10.1109/IEDM.2016.7838029.

[20] T. Chohan et al., "Implication of self-heating effect on device reliability characterization of Multi-Finger n-MOSFETs on 22FDSOI," *IEEE Trans. Device. Mater. Rel.*, vol. 22, no. 3, pp. 387–395, Sep. 2022, doi: 10.1109/TDMR.2022.3183630.

[21] Q. Xing et al., "Investigation of self-heating effects in UTBBFD-SOI MOSFETs by a modified thermal conductivity model," *IEEE Trans. Electron Devices*, vol. 69, no. 8, pp. 4129–4137, Aug. 2022, doi: 10.1109/TED.2023.3297564.

[22] D. L. Scharfetter and H. K. Gummel, "Large-signal analysis of a silicon read diode oscillator," *IEEE Trans. Electron Devices*, vol. 16, no. 1, pp. 64–77, Jan. 1969, doi: 10.1109/T-ED.1969.16566.

[23] A. Bouchikhi, "Two-dimensional numerical simulation of the DC glow discharge in the normal mode and with Einstein's relation of electron diffusivity," *Plasma Sci. Technol.*, vol. 14, no. 11, pp. 965–973, Nov. 2012, doi: 10.1088/1009-0630/14/11/04.

[24] Y. Shen, H.-A. Yang, and B.-Y. Cao, "Near-junction phonon thermal spreading in GaN HEMTs: A comparative study of simulation techniques by full-band phonon monte carlo method," *Int. J. Heat Mass Transf.*, vol. 211, Sep. 2023, Art. no. 124284, doi: 10.1016/j.ijheatmasstransfer.2023.124284.

[25] H. Bao, J. Chen, X. Gu, and B.-Y. Cao, "A review of simulation methods in micro/nanoscale heat conduction," *ES Energy Environ.*, vol. 1, no. 39, pp. 16–55, Oct. 2018, doi: 10.30919/esee8c14.

[26] D. B. M. Klaassen, "A unified mobility model for device simulation—I. Model equations and concentration dependence," *Solid-State Electron.*, vol. 35, no. 7, pp. 953–959, Jul. 1992, doi: 10.1016/0038-1101(92)90325-7.

[27] D. B. M. Klaassen, "A unified mobility model for device simulation—II. Temperature dependence of carrier mobility and lifetime," *Solid-State Electron.*, vol. 35, no. 7, pp. 961–967, Jul. 1992, doi: 10.1016/0038-1101(92)90326-8.

[28] C. M. Liu, J. B. Kuo, and Y. P. Wu, "An analytical quasi-saturation model considering heat flow for a DMOS device," *IEEE Trans. Electron Devices*, vol. 41, no. 6, pp. 952–958, Jun. 1994, doi: 10.1109/16.293307.

[29] M. N. Darwish, "Study of the quasi-saturation effect in VDMOS transistors," *IEEE Trans. Electron Devices*, vol. 33, no. 11, pp. 1710–1716, Nov. 1986, doi: 10.1109/T-ED.1986.22732.

[30] W. Hansch, T. Vogelsang, R. Kirchner, and M. Orlowski, "Carrier transport near the Si/SiO2 interface of a MOSFET," *Solid State Electron.*, vol. 32, no. 10, pp. 839–849, Oct. 1989, doi: 10.1016/0038-1101(89)90060-9.

[31] A. Chaudhry and M. J. Kumar, "Controlling short-channel effects in deep-submicron SOI MOSFETs for improved reliability: A review," *IEEE Trans. Device Mater. Reliab.*, vol. 4, no. 1, pp. 99–109, Mar. 2004, doi: 10.1109/TDMR.2004.824359.

[32] S. A. Schwarz and S. E. Russek, "Semi-empirical equations for electron velocity in silicon: Part II—MOS inversion layer," *IEEE Trans. Electron Devices*, vol. 30, no. 12, pp. 1634–1639, Dec. 1983, doi: 10.1109/T-ED.1983.21424.
[33] R. N. Hall, "Electron-hole recombination in germanium," *Phys. Rev.*, vol. 87, no. 2, pp. 387–387, Jul. 1952, doi: 10.1103/PhysRev.87.387.
[34] W. Shockley and W. T. Read, "Statistics of the recombinations of holes and electrons," *Phys. Rev.*, vol. 87, no. 5, pp. 835–842, Sep. 1952, doi: 10.1103/PhysRev.87.835.
[35] V. Palankouski, R. Schultheis, and S. Selberherr, "Simulation of power heterojunction bipolar transistors on gallium arsenide," *IEEE Trans. Electron Devices*, vol. 48, no. 6, pp. 1264–1269, Jun. 2001, doi: 10.1109/16.925258.
[36] K. Raleva, D. Vasileska, A. Hossain, S.-K. Yoo, and S. M. Goodnick, "Study of self-heating effects in SOI and conventional MOSFETs with electro-thermal particle-based device simulator," *J. Comput. Electron.*, vol. 11, pp. 106–117, Jan. 2012, doi: 10.1007/s10825-012-0384-0.
[37] J. Sanchez, "DEVSIM: A TCAD semiconductor device simulator," *J. Open Source Softw.*, vol. 7, no. 70, p. 3898, Feb. 2022, doi: 10.21105/joss.03898.
[38] J. E. Sanchez and Q. Chen, "Element edge based discretization for TCAD device simulation," *IEEE Trans. Electron Devices*, vol. 68, no. 11, pp. 5414–5420, Nov. 2021, doi: 10.1109/TED.2021.3094776.
[39] D. M. Caughey and R. E. Thomas, "Carrier mobilities in silicon empirically related to doping and field," *Proc. IEEE*, vol. 55, no. 12, pp. 2192–2193, Dec. 1967, doi: 10.1109/PROC.1967.6123.
[40] 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, doi: 10.1109/JEDS.2016.2568261.
[41] 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.*, 2014, pp. XT.8.1–XT.8.6, doi: 10.1109/IRPS.2014.6861186.
[42] J. A. Spencer, A. L. Mock, and Y. Zhang, "Heating issues in wide-bandgap semiconductor devices," in *Thermal Management of Gallium Nitride Electronics*. Amsterdam, The Netherlands, Elsevier, 2022, pp. 1–19, doi: 10.1016/j.ijheatmasstransfer.2022.123497.
[43] K. O. Petrosyants, I. A. Kharitonov, S. V. Lebedev, L. M. Sambursky, S. O. Safonov, and V. G. Stakhin, "Electrical characterization and reliability of submicron SOI CMOS technology in the extended temperature range (to 300 °C)," *Microelectron. Rel.*, vol. 79, pp. 416–425, Dec. 2017, doi: 10.1016/j.microrel.2017.05.018.
[44] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideovt, E. Bassous, and A. R. Leblanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE J. Solid-State Circuits*, vol. 9, no. 5, pp. 256–268, Oct. 1974, doi: 10.1109/JSSC.1974.1050511.