Measurement and Modeling of Self-Heating in SOI NMOSFET's

Lisa T. Su, Student Member, IEEE, James E. Chung, Dimitri A. Antoniadis, Fellow, IEEE, Kenneth E. Goodson, and Markus I. Flik

Abstract—Self-heating in SOI nMOSFET's is measured and modeled. Temperature rises in excess of 100 K are observed for SOI devices under static operating conditions. The measured temperature rise agrees well with the predictions of an analytical model and is a function of the silicon thickness, buried oxide thickness, and channel-metal contact separation. Under dynamic circuit conditions, the channel temperatures are much lower than predicted from the static power dissipation. This work provides the foundation for the extraction of device modeling parameters for dynamic operation (at constant temperature) from static device characterization data (where temperature varies widely). Self-heating does not greatly reduce the electromigration reliability of SOI circuits, but might influence SOI device design, e.g., requiring a thinner buried oxide layer for particular applications and scaled geometries.

I. INTRODUCTION

Silicon-on-insulator (SOI) technology is a potential challenger to bulk silicon as the substrate material for future VLSI applications due to its potentially increased circuit speeds and simpler fabrication process [1]. These advantages arise from the presence of the buried insulating layer, most commonly silicon dioxide, which reduces the parasitic source/drain-to-substrate junction capacitance, limits the depth of the source/drain junction to form simple shallow junctions, and allows full dielectric isolation of the device to eliminate latchup. But, the low thermal conductivity of the underlying silicon dioxide layer, which is about one and a half orders of magnitude less than that of silicon, inhibits cooling in SOI devices and causes severe self-heating. This results in higher channel operating temperatures and is evidenced by the negative differential conductance at high gate biases that is characteristic of most SOI devices [2], [3]. The device mobility is reduced as a result of the elevated channel temperature rise and results in reduced maximum drain saturation current and more complicated device modeling. In addition, high channel temperatures lead to increased interconnect temperatures at the silicon-metal contact and make conduction cooling through the source, drain, and interconnects important [4]. These factors motivate a need for accurate measurement and modeling of channel and interconnect temperatures in SOI MOSFET's.

Self-heating effects in SOI MOSFET's have been measured previously using a liquid crystal technique [3], the temperature dependence of leakage currents [5], and noise thermometry [6]. Thermal models have also investigated the temperature rise [4], [7]–[9]. These initial data indicated a significant temperature rise. However, a systematic study of the temperature rise as a function of critical device dimensions, e.g., buried oxide thickness, and a comparison of direct temperature measurements to thermal modeling are not available. In the present work, a simple technique for temperature measurement in SOI MOSFET's is developed and channel and interconnect temperatures are measured over a wide range of device parameters. Measurements are compared with thermal model predictions and the impact on circuit design is assessed.

II. EXPERIMENTAL DEVICES AND PROCEDURE

The test structures for temperature measurement are MOSFET's with the gate configured for four-point resistance measurements as shown in Fig. 1. The polysilicon gate serves as a temperature-sensing resistor for the channel region, where the device power is dissipated. This technique was first proposed by Mautry et al. for temperature measurement in conventional bulk MOSFET's [10]. This work performs additional thermal analysis to make the technique applicable to SOI devices.
A. Device Fabrication

Fig. 2. Cross-sectional view of a completed SOI MOSFET. The device dimensions are given in Table I.

<table>
<thead>
<tr>
<th>THERMAL CONDUCTIVITIES AND DEVICE DIMENSIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>silicon dioxide conductivity, $k_{\text{ox}}$</td>
</tr>
<tr>
<td>aluminum conductivity, $k_m$</td>
</tr>
<tr>
<td>source/drain/gate (n + silicon)</td>
</tr>
<tr>
<td>conductivity $k_d$</td>
</tr>
<tr>
<td>channel and substrate conductivity, $k_c$</td>
</tr>
<tr>
<td>channel length, $L$</td>
</tr>
<tr>
<td>channel width, $W$</td>
</tr>
<tr>
<td>silicon thickness, $t_{si}$</td>
</tr>
<tr>
<td>gate oxide thickness, $t_{ox}$</td>
</tr>
<tr>
<td>SIMOX buried oxide thickness, $t_{box}$</td>
</tr>
<tr>
<td>channel-metal contact separation, $L_d$</td>
</tr>
<tr>
<td>channel-gate contact separation, $t_m$</td>
</tr>
<tr>
<td>interconnect width, $w_m$</td>
</tr>
<tr>
<td>gate polysilicon thickness, $t_g$</td>
</tr>
<tr>
<td>aluminum interconnect thickness, $t_m$</td>
</tr>
<tr>
<td>field oxide thickness, $t_{fox}$</td>
</tr>
<tr>
<td>gate-metal interlevel oxide thickness, $t_s$</td>
</tr>
</tbody>
</table>

B. Measurement Technique

The resistance of the polysilicon gate, $R_G(P)$, is measured as a function of device power, where $P = I_D V_{DS}$. The temperature dependence of $R_G$ is calibrated by varying the substrate temperature using a variable temperature chuck with the device turned off, i.e. an isothermal gate. From the resulting calibration function and $R_G(P)$, an average gate temperature, $T_{G(avg)}(P)$, is obtained.

The average channel temperature, $T_C$, is approximately equal to the local gate temperature directly above it because the thermal resistance of the gate oxide is small and heat transfer to the ambient air is negligible [11]. However, the value of $T_C$ is not equal to $T_{G(avg)}$ because the calibration technique assumes the entire gate is heated. During device operation, only the regions of the gate directly above the channel are heated and contribute to a change in $R_G$. The difference between $T_C$ and $T_{G(avg)}$ depends on the shape of the temperature profile in the gate region. In bulk devices, the high thermal conductivity silicon substrate tends to distribute the heat in the gate region and broaden the temperature profile; thus $T_C$ is well approximated by $T_{G(avg)}$ [10]. In contrast, in SOI devices, the length-scale for heat conduction is on the order of a few micrometers, as discussed in Section IV. The temperature distribution along the gate width is significant and must be considered.

The temperature profile can be estimated as a constant temperature in the heated portions of the gate and a linear decrease to the substrate temperature $T_0$ at the gate pads. The resultant temperature profile along the gate width is shown schematically in Fig. 3(a). The local gate electrode resistance per unit width in the $Z$ direction, shown schematically in Fig. 3(b), is obtained from the calibration function and the assumed temperature profile. The integration of the calculated gate electrode resistance per unit width in the $Z$ direction yields the total resistance $R_G$ for a gate electrode having a channel temperature. Using this technique, the channel temperature is found by iteratively calculating $R_G$ for increasing $T$ until the measured $R_G$ is reached.

This $T_C$ somewhat underestimates the actual channel temperature because the temperature profile assumes the channel is isothermal and neglects heat conduction through the buried oxide layer to the substrate in the region between the device edge and the gate contact pad. The error from these assumptions is less than 8% for our SOI test structures ($W = 10 \mu m$) and is included in the measurement error [11]. This error can be further reduced by using test structures with larger $W$.

III. EXPERIMENTAL RESULTS

Channel temperature measurements are performed on test structures with varying $t_{si}$, $t_{box}$, and $L_d$. Fig. 4 shows an example of the raw data from device and calibration measurements. Typical operating conditions for a 0.3 μm technology ($V_D = 2-3$ V, $V_G = 2-3$ V) are used for the device measurement. The sheet resistance of the arsenic-implanted gate resistor is typically 250–280 Ω/sq. The relatively high resistance of the arsenic-implanted gate resistor produces a large temperature coefficient consistent with the behavior of a semiconductor.
The technique would also work with a low resistance, silicided gate, although the temperature coefficient would likely be smaller. Polysilicon depletion is negligible for the temperature measurements because no change in the unheated $R_G$ is observed at different gate biases.

Fig. 5 shows the extracted channel temperature for SOI devices of varying silicon thicknesses and for a bulk device. The temperature rise, $T_C - T_0$, in each case is proportional to the power, and for a given power is much larger in the SOI device than in the bulk device, e.g., 70 K for $P = 8$ mW and $t_s = 78$ nm versus only 10 K in the bulk device. As $t_s$ is reduced, the channel temperature increases. The channel temperature versus power slope for each case can be interpreted as a thermal resistance from the channel at temperature $T_C$ to the chuck at temperature $T_0$, $(T_C - T_0)/P$.

Fig. 6 shows channel temperature data for varying buried oxide thicknesses. As $t_{box}$ is reduced, the temperature and its derivative with respect to power are reduced. Fig. 7 shows the channel temperature for varying channel-metal contact separation. Although the channel temperature decreases with decreasing $L_d$, there is little change for the range of $L_d$ investigated.

IV. STEADY-STATE THERMAL MODEL

A. Model Description

A steady-state thermal model [4] is used to understand the experimental data. The thermal model treats the source,
drain, gate, and interconnects as cooling fins for the power dissipated in the channel heater. Fig. 8 schematically shows the fin geometry. A plane of symmetry is drawn through half of the gate in the Y-Z direction and in the X-Y direction. The model assumes an isothermal channel and an isothermal heat equation is solved in each fin with the conditions of substrate at the temperatures of the gate in the metal interconnect, gate and drain respectively. The thermal conductivities and thickness of the fin. The parameter \( l/m \) is chosen at \( 349 \text{ K} \) which is the average of temperature from the gate resistance. The thermal conductivities and device dimensions used in the model. The thermal conductivities and thickness of the silicon film thickness. Each data point is the average of thermal resistance measurements at several powers on a single device. The error bars take into account the uncertainties in the device physical parameters (i.e., \( t_{si}, t_{box}, t_{ox}, k_{d}, k_{ox}, \text{etc.} \)) and the approximation employed for calculating the channel temperature from the gate resistance. As the silicon thickness is reduced, the temperature rise in the device increases, showing the same trend as the data. Physically, this can be understood through the relationship between \( l/m \) and \( t_{box} \). As \( t_{box} \) increases, the area of the source and drain that participates in appreciable cooling to the substrate is increased, and \( T_{c} \) and \( R_{c} \) are reduced. The agreement between the thermal model predictions and the data is very good considering the uncertainty in the thermal conductivities and device dimensions used in the model.

Fig. 10 shows the SOI device thermal resistance as a function of the channel-metal contact separation, for large \( L_{d} \), this parameter has little effect on the temperature rise because the temperature has decreased in the source and drain before the interconnect is reached. As \( L_{d} \) is reduced to less that 1 \( \mu \text{m} \), comparable to the thermal healing length in the source and drain, the interconnects become more important as cooling fins, and the device temperature is reduced. This causes the metal contact temperature to increase.

Fig. 11 illustrates the dependence of the channel-substrate thermal resistance on the buried oxide thickness; the device temperature decreases as \( t_{box} \) decreases because the thermal resistance of the oxide layer decreases. But the dependence is
on device dimensions. In the typical range of operation of modern MOSFET's, this static temperature rise is significant. However, these temperature rises are not necessarily problematic for device performance because under dynamic operating conditions typical of digital circuit operation, the temperature rise will not follow the instantaneous power dissipation. This occurs because the thermal time constants are much longer than typical electrical periods (e.g., clock); thus during dynamic operation, the temperature rise is significantly smaller than the static case [8], [15]. The main impact of self-heating lies in the a) extraction of model parameters for circuit simulation and b) potential for increased interconnect temperatures during circuit operation.

A. Extraction of Model Parameters

Because of the significant temperature differences between static device operation and dynamic circuit operation, model parameters for circuit design obtained from static device characterization can be in substantial error. Constant temperature parameters, i.e., the nonself-heated device characteristics, are needed for typical digital circuit applications. Although dynamic measurements of current-voltage data are possible, this would require high-speed measurement equipment and be a significant burden for device characterization. A simple technique for the extraction of the nonself-heated device characteristics can be implemented using the temperature measurements in conjunction with standard I-V data. Since most circuit models include temperature-dependent device parameters, once the $P$ versus $T$ slope is measured, the temperature is known at each point of the I-V data, and the nonself-heated device characteristics can be reconstructed from the static SOI I-V characteristics. This issue will be addressed in a future publication.

B. Interconnect Reliability

In a practical digital circuit, the maximum interconnect temperature is of concern because of electromigration considerations. In the worst case found here (static operation), for $t_o = 41$ nm, $L_d = 0.4$ μm, $t_{box} = 360$ nm, this temperature rise approaches 25 K for a device power of 1 mW/μm device width (from Fig. 10). However, as mentioned previously, devices are only operating a fraction of the time in a real circuit and the static power predicts too large a temperature rise in the interconnects.

The steady-state temperature rise during circuit operation in the metal contact can be estimated using the thermal model if the transient temperature fluctuations at the metal contact are small. The thermal diffusion length is a measure of the length over which the transient temperature fluctuations are significant and is approximately $(cYT)_{diff}$, where $Q = 0.33$ cm$^2$/s is the thermal diffusivity of heavily doped silicon, and $T$ is the clock period. Using $T = 5$ ns (200 MHz clock), the thermal diffusion length is 0.4 μm, and the time-averaged power can be used with the steady-state thermal model to estimate the temperature rise in the metal contact. The power dissipation of importance in this case is the average power...
dissipation per unit active area (not the power per total chip area) because cooling occurs primarily in the active area.

The average power dissipation per unit active area of a representative heavily loaded NMOS transistor (e.g., clock driver) in a 0.4 μm CMOS technology is estimated to be about 3000 W/cm². The length of the active area in that technology would be about 2.8 μm (Ld of 0.4 μm, contact hole width of 0.4 μm, contact spacing to edge of active area of 0.4 μm). Hence, the time-averaged power dissipation is 0.084 mW/μm device width. For this case, the model predicts a steady-state metal contact temperature rise on the order of 3 K.

The above analysis assumes a device layout of an isolated test device. If the device is very wide and is laid out in a serpentine fashion, as is common for such devices, the power dissipation per unit width almost doubles because neighboring devices share a source/drain. In this case, the source/drain regions and the metal interconnect become less important for heat removal. A worst case approximation of this case would return metal contact temperature rise on the order of 3 K.

VI. CONCLUSIONS

The steady-state temperature rise in SOI devices is measured and modeled as a function of various device parameters. The temperature rise is significant and dependent on the buried oxide thickness, silicon thickness, and channel-metal contact separation. The difference in temperatures between device characterization and actual circuit conditions requires accurate measurement and modeling of temperature rises. Channel-temperature measurements are needed in addition to typical current-voltage characterization to extract appropriate modeling parameters for dynamic circuit calculations. Steady-state analysis indicates that self-heating effects do not appear to limit the use of SOI technology, but optimization of device dimensions may be necessary for particular applications and scaled geometries.

ACKNOWLEDGMENT

The authors gratefully acknowledge IBIS Corp. for the SIMOX samples used in this study.

REFERENCES

James E. Chung received the B.S. degree in electrical engineering from the University of Illinois, Champaign-Urbana, in 1984 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1988 and 1990. He spent the fall of 1990 as a Visiting Researcher at the Motorola Advanced Products Research and Development Laboratory in Austin, Texas. Currently, he is the Analog Devices Career Development Assistant Professor in the Department of Electrical Engineering, Massachusetts Institute of Technology, and a Member of the Microsystems Technology Laboratories. He has research interests in the areas of sub-0.1 μm MOSFET device physics, VLSI technology, SOI materials and devices, and hot-electron and thin-dielectric reliability.

Dimitri A. Antoniadis received the B.S. degree in physics from the National University of Athens in 1970, the M.S.E.E. degree in 1973 and the Ph.D. degree in electrical engineering in 1976 from Stanford University. From 1969 to 1976 he conducted research in the area of measurements and modeling of the earth’s ionosphere and thermosphere. Starting with the development of the SUPREM process simulator in 1976, his technical activity has been in the area of semiconductor devices and integrated circuit technology. He has worked on the physics of diffusion in silicon, silicon-on-insulator technology and devices, and device measurements. At present his research focuses on a) extreme submicron and quantum-effect electronic devices; b) computer aids for the design and analysis of electronic devices and of the corresponding fabrication processes; c) novel concepts in semiconductor device fabrication. From 1970 to 1971 he was a Fellow of the National Research Institute, Athens. From 1976 to 1978 he was Research Associate and Instructor in the Department of Electrical Engineering at Stanford University. He joined the faculty at M.I.T. in 1978, where he is Professor of Electrical Engineering. From 1984 to 1990 he was Director of the MIT Microsystems Technology Laboratories which he has helped establish. He is author or coauthor of over 120 technical articles.

Kenneth E. Goodson received the B.S., M.S., and Ph.D. degrees from the Department of Mechanical Engineering at the Massachusetts Institute of Technology in 1989, 1991, and 1993, respectively. He is now with Daimler Benz Research and Technology in Ulm, Germany. His research interests include the submicrometer thermal engineering of electronic circuits and sensors, with a focus on thermal conduction processes in amorphous, high-Tc superconducting, and CVD diamond layers. At Daimler Benz, he is developing high-power transistor structures which use diamond layers to reduce operating temperatures.

Dr. Goodson held an Office of Naval Research Graduate Fellowship from 1989 to 1992. He is an M.I.T. Burchard Scholar and a member of Tau Beta Pi, Phi Beta Kappa, and Sigma Xi.

Markus I. Flik received the Dipl. Ing. degree from the Swiss Federal Institute of Technology, Zurich, in 1985 and the M.S. and Ph.D. degrees in mechanical engineering from the University of California, Berkeley, in 1987 and 1989, respectively. He took part in Post-Doctoral work at the Tokyo Institute of Technology. In 1989, he joined the Department of Mechanical Engineering, Massachusetts Institute of Technology, as an Assistant Professor, where he currently holds the Samuel C. Collins Career Chair. His research interests include radiation and conduction, heat transfer in microstructures, e.g., thin films and multiple quantum wells, and novel materials, e.g., high-Tc superconductors and diamond films.

Dr. Flik received the Silver Medal from the Swiss Federal Institute of Technology. He currently serves on the ASME Committee for Heat Transfer in Materials Processing.