Abstract—Thermal test structures and ring oscillators (ROs) are fabricated in 0.18-μm three-dimensional (3-D)–SOI technology. Measurements and electrothermal simulations show that thermal and parasitic effects due to 3-D packaging have a significant impact on circuit performance. A physical thermal model is parameterized to provide better prediction of circuit performance in 3-D technologies. Electrothermal simulations using the thermal model show good agreement with measurement data; the model is applicable for different circuits designed in the 3-D–SOI technology. By studying the phase noise of ROs, the device noise properties of 3-D–SOI technology are also characterized and compared with conventional bulk CMOS technology.

Index Terms—Device noise, electrothermal, thermal modeling, three-dimensional (3-D) integrated circuit, 3-D silicon-on-insulator.

I. INTRODUCTION

THREE-DIMENSIONAL (3-D) integrated-circuit technology offers the ability to stack die in the vertical dimension, and there is recent work using this technology to improve system performance [1]–[5]. By utilizing the vertical dimension, wiring lengths and chip sizes can be reduced. These reductions keep signal propagation delay and power consumption due to interconnections under control as the LSI capacity increases. Therefore, 3-D technology can be used to achieve large-scale integration with high performance and low power consumption. However, thermal and packaging effects need to be addressed before utilizing any 3-D technology, as these effects impact both system architecture and circuit design [6]. Furthermore, little is known about how thermal and packaging effects impact device noise characteristics in emerging 3-D technologies.

II. THERMAL AND PACKAGING EFFECTS IN 3-D–SOI TECHNOLOGY

There are many types of 3-D integrated-circuit technologies which are actively being studied, including monolithic [7], stacked-chip [8], and stacked-wafer [6] 3-D integrated circuits. All the circuits discussed in this paper are implemented with stacked-wafer 3-D technology, offered by the Lincoln Laboratory, Massachusetts Institute of Technology (MIT). In this 0.18-μm technology, three wafers using fully depleted SOI are bonded with interconnects using through-wafer vias [6]. The integrated-circuit layers are labeled Tier A (A), Tier B (B), and Tier C (C), with Tier A closest to the 675-μm-thick substrate, as shown in Fig. 1.

This paper presents our findings on the thermal and packaging effects, as well as device noise characterization, obtained from a 3-D multiproject run offered by the Lincoln Laboratory. The purpose of this paper is twofold. First, a physics-based thermal model will be presented. The parameters of this model were extracted from the thermal test structures and verified with a separate set of test structures from another research group [9]. It is also demonstrated that this model is applicable for the electrothermal simulations of different circuits. As an application example of the thermal model, the electrothermal simulation of ring oscillators (ROs) fabricated using 3-D–SOI technology is presented. Second, device noise characterization was done for 3-D–SOI devices with various geometric parameters. The device noise performance is then compared with a single-chip conventional 0.18-μm bulk CMOS process.
Fig. 2(a) shows the measured characteristics of a MOSFET implemented on all three tiers. Note that Tier A has the highest drive current and Tier C has the lowest drive current. The performance degradation of Tiers B and C is mainly due to 3-D packaging (through-wafer vias) and other parasitic resistive effects.

Fig. 2(b) shows the measured $I-V$ characteristics of diodes from the same 3-D circuit. Note that the Tier-C diode now has the highest current, and that of Tier A has the lowest. This trend is opposite to that shown in Fig. 2(a). The physical cause of this difference can be traced back to the thermal and packaging effects in the 3-D-SOI technology.

The MOSFET performance degradation shown in Fig. 2(a) could be mainly due to thermal-induced mobility reduction. Without losing its generality, the saturation current expression for a MOSFET can be approximated by

$$I_{DSat} = \frac{W}{2L} \cdot \left( V_{GS} - V_T \right)^2$$

where $R_{SP parasitic}$ is the parasitic resistance between the source pad to the source terminal of the transistor. For this test structure, $R_{SP parasitic}$ is a few ohms and $I_{DSat}$ is a few milliampere; therefore, the packaging effects are negligible. However, carrier mobility $\mu$ decreases considerably with increasing temperature $T$ and the effective mass of an electron $m^*$, following [10]:

$$\mu = \frac{\sqrt{2}\pi q^3 C_{11}}{3E_{dd}(m^*)^{3/2}(kT)^{3/2}} \propto (m^*)^{-5/2}(T)^{-3/2}.$$  \hspace{1cm} (2)

Although the threshold voltage decreases at higher temperatures, the effect of mobility lowering still dominates, resulting in an $I_{DSat}$ decrease with increasing temperature. It follows that the transistor on Tier A has the highest $I_{DSat}$, while Tier C has the lowest, as shown in Fig. 2(a).

Diode performance degradation due to 3-D packaging and parasitic resistive effects is shown in Fig. 2(b). The diode current can be expressed as

$$I_{Diode} = I_0 \left( e^{V_{Diode} / V_T} - 1 \right)$$

$$= I_0 \left( e^{(V_{Diode} - I_{Diode} \cdot R_{SP parasitic}) / V_T} - 1 \right) \hspace{1cm} (3)$$

For this test structure (with reference to Fig. 1), the product of $I_{Diode}$ and $R_{SP parasitic}$ is significant because $I_{Diode}$ in Fig. 2(b) is much larger than $I_{DS}$ in Fig. 2(a), and the $I_{Diode} \times R_{SP parasitic}$ product has an exponential effect on $I_{Diode}$. For this test structure, this effect dominates over the increase of the intrinsic carrier density as temperature increases. This is the reason for Tier A having the lowest diode current and Tier C the highest, as shown in Fig. 2(b). Parasitic resistance of each through-wafer via is independently measured to be approximately 3 $\Omega$. The current degradation of the diodes, as shown in Fig. 2(b), may be explained by parasitic resistances of this magnitude.

These effects on circuit performance show opposite trends. Therefore, both thermal modeling and parasitic extraction are required for an electrothermal simulator to help circuit designers better predict actual circuit performance.

III. THERMAL MODELING AND ELECTROTHERMAL SIMULATIONS

A. Experiments

An N-type MOSFET is surrounded by diode-connected MOSFETs, which act as on-chip heat sources to control the temperature of this device under test (DUT). The DUT has four connections to the gate, which can be used to accurately measure the gate resistance [11]. The test chip is first placed in a hot chamber, and the temperature dependence of the polysilicon gate resistance is calibrated in an enclosed temperature-controlled environment, as shown in Fig. 3. The test chip is then placed in ambient, and on-chip heat sources are used to control the DUT temperature. Changes in gate resistance can then be used to determine the temperature when the on-chip heat sources are turned on. The effects of process variation effects were taken into account by calibrating each gate resistor...
Fig. 3. Mapping of gate resistance as a function of temperature. With this characterization, the gate resistance can be used to determine temperature when the on-chip heat sources are turned on. The transistor shown has $W/L = 6 \, \mu m/0.2 \, \mu m$, and is on Tier C. Figure shown is a typical measurement. Each gate resistor is calibrated individually.

Fig. 4. (a) Layout of the DUT and the heat sources on each tier. Temperature of DUT can be controlled with surrounding heat sources. DUT gate resistance is used as a thermometer. (b) Three-dimensional view of test structures. Test structures are stacked on top of one another and can be turned on or off individually to simulate different operating conditions.

individually. The layout of the DUT and heat sources are shown in Fig. 4.

The thermal time constant for the self-heating of each tier can be determined with this test structure by using the ac output conductance technique [12]. The gate of the DUT is first biased to $V_{dd}$. An impedance analyzer is then used to apply both a dc bias and an ac signal with varying frequencies at the drain. The drain admittance can then be measured as a function of frequency. The drain admittance starts out negative at low frequencies and then eventually becomes positive as the frequency increases. This is due to the suppression of ac self-heating at higher frequencies. In this way, the thermal time constant may be experimentally determined. The results are shown in Fig. 6.

In this paper, thermal test structures described earlier were fabricated in order to develop a physics-based thermal model that can be experimentally verified.

A separate set of thermal test structures was used to validate the thermal model presented in this paper [9]. As shown in Fig. 5, this test structure is entirely different from the test structure used in this paper, including different layout areas [$\sim 522 \, \mu m^2$ for Fig. 4(a) versus $\sim 2400 \, \mu m^2$ for Fig. 5], heat sources (diode-connected transistors versus resistors), and thermal sensors (polysilicon gate resistors versus diodes). As shown in Fig. 5, in this test structure [9], four SOI active resistors are placed near a diode which is used as a thermometer. This test structure was replicated in each of the three tiers. Voltages were then applied to the resistors, and the diode current was used as a temperature sensor. Temperature measurements were made for all three tiers. The thermal model presented in this paper matches the temperature measurements made by the research group at Lincoln Laboratory, as will be presented in the next section.

This thermal model was then implemented in an electrothermal simulator, which was used to predict the actual circuit performance of ROs fabricated on the same chip (Section III-C). These same oscillators were also used to quantify device noise (Section IV).

B. Thermal Modeling

A physical lumped thermal model is shown in Fig. 6, and the model parameters can be extracted using measurement data collected from the thermal test structure shown in Fig. 4. This lumped model assumes an average tier temperature. This assumption holds because the test circuit is designed to be small compared with the geometrical scale for thermal conduction and also because the thermal conductivity of silicon is much larger than that of silicon dioxide. Therefore, the temperature profile across the tiers is more significant than the profile within a tier.

The thermal model is also verified, using thermal test structures independently designed by the Lincoln Laboratory, to be applicable for different circuits. A single known temperature
at a given power consumption level from the measured data is required to calibrate the thermal model (Fig. 7). Calibration is needed to reconcile measurement errors and experimental deviations for model assumptions made by both teams. The Lincoln Laboratory team designed a larger test structure, using resistors as heating elements and diodes as thermal sensors. A heated chuck was also used to control the external temperature. In contrast, the test structure described in this paper uses diode-connected transistors as heating elements and polysilicon gate resistors as thermal sensors. The gate resistances were measured using four-point measurements to ensure accuracy, and a hot chamber is used to control external conditions. Given the differences in the test structure designs and measurement setups of both teams, it is not surprising that a single calibration point is needed to reconcile both sets of measurement results.

Fig. 8 shows the measured temperature data, together with the temperature predicted by the thermal model presented in this paper. It is worthwhile to note here that more power is required to raise the temperature using the Lincoln Laboratory test structures than the test structures presented here since they are approximately four times larger than the test structures shown in Fig. 4(a). Nevertheless, the thermal model presented in this paper uses area as a parameter and is able to account for different layout areas. The agreement between the measured temperature data using two independently designed thermal test structures verifies that the thermal model is applicable for different circuits designed using the 3-D–SOI technology.

As an application example, the thermal model will be used to perform electrothermal simulations with asymmetric ROs designed for device noise measurements.
C. Electrothermal Simulations

Asymmetric ROs were fabricated on this test chip primarily to measure device noise [13]. However, since oscillation frequency is a function of oscillator temperature, these oscillators can also be used to verify electrothermal simulation results. The design of the ROs will be explained in Section IV, where the device noise is discussed.

Three different asymmetric ROs with minimum channel lengths of 0.18 \( \mu m \) (1\( \times \)), 0.54 \( \mu m \) (3\( \times \)), and 1.08 \( \mu m \) (6\( \times \)) are fabricated on each tier, resulting in a total of nine ROs.

Electrothermal simulations were performed to predict the oscillation frequencies of ROs implemented on different tiers. The temperature beneath the substrate and the temperature of the air far above the test chip are set to be ambient. These two boundary conditions mirror the experimental setup and are built into the thermal model. It is expected that both thermal and parasitic packaging effects will affect circuit performance. However, the exact operating conditions, under which each type of effect dominates, are unknown. Electrothermal simulations can help to gain design intuition.

For the first set of experiments, only one oscillator is turned on at any one time. Fig. 9 shows measured oscillation frequencies and simulation results.

For the second set of experiments, all oscillators are turned on at the same time. Measurements and simulation results are shown in Fig. 10. In Figs. 9 and 10, “No compensation” is the idealistic (and most unrealistic) case when neither thermal nor parasitic resistive effects are considered. “Parasitic Effect only” is when the temperature is set to 25 \(^\circ\)C for all tiers. “Thermal Effect only” is for parasitic resistances set to zero. “Electrothermal simulation” considers both thermal and parasitic resistive effects.

Fig. 9 shows the situation when only one tier is active at any one time. For this case, Tier C has the lowest oscillation frequency for the 1\( \times \) oscillators [Fig. 9(a)], and Tier A has the lowest oscillation frequency for the 3\( \times \) [Fig. 9(b)] and 6\( \times \) oscillators (not shown). This is due to thermal effects being dominant in 1\( \times \) oscillators, and 3-D packaging effects being dominant in 3\( \times \) and 6\( \times \) oscillators. Thermal effects are dominant in 1\( \times \) oscillators as they have a smaller area over which to dissipate heat through conduction. From Fig. 2(a), drive current is lowest on Tier C when thermal effects are dominant. Note that lower drive current translates to a lower oscillation frequency. It follows that Tier C has the lowest oscillation frequency for 1\( \times \) oscillators, as observed in Fig. 9(a).
For the larger 3x and 6x oscillators, heat dissipation is efficient enough that 3-D packaging effects become dominant. From Fig. 2(b), Tier A has the lowest current when 3-D packaging effects are dominant. Hence, Tier A has the lowest oscillation frequency for 3x and 6x oscillators [Fig. 9(b)].

For the situation when all three tiers are active at the same time, as shown in Fig. 10, Tier B has the lowest oscillation frequency for the 1x oscillators [Fig. 10(a)], and Tier A has the lowest oscillation frequency for the 3x [Fig. 10(b)] and 6x oscillators (not shown). It is shown in Fig. 1(a) that the dielectric layer thickness between Tiers A and B is twice of that between Tiers B and C. This difference is due to fundamental technology design considerations and can explain why the temperatures of Tiers B and C are very close for the 1x oscillators when all tiers are on [Fig. 7(b)]. The increase in temperature of Tier B, together with the ever-present parasitic effects, contributed to Tier B having the lowest oscillation frequency for the 1x oscillators when all tiers are on [Fig. 10(a)]. There is no straightforward way to predict this result a priori without electrothermal simulations, as shown earlier.

The observed different effects are validated through simulations (Figs. 9 and 10). Experimentally observed oscillation frequencies vary in the same direction, as predicted by the specific dominant effect (thermal and parasitic effects for the 1x and the 3x oscillators, respectively). Considering both effects as in “electrothermal simulation” gives the best match to actual circuit performance for all cases. Note that the transistor drive current degradation shown in Fig. 2(a) is due to mobility decreases. As transistors approach the ballistic transport regime, drive current increases with temperature. In this case, Tier A is expected to always have the worst performance, as both thermal and packaging effects work against it. More investigation is required to confirm this initial observation, as the impact of advanced packaging technologies with improved thermal dissipation performance was not considered in this paper.

### IV. DEVICE NOISE

#### A. Experiments

Nine asymmetric ROs were fabricated to measure device noise [13]. These ROs were designed so that accurate measurements of the device noise may be made on the circuit level.

This RO has seven alternating stages, where an inverter with a small width is followed by a large inverter with a width ten times as large (Fig. 11), except for the last stage. A large inverter is needed at the last stage to drive the output buffer. These inverters are capacitively loaded with large metal–insulator–metal (MIM) capacitors. These MIM capacitors are designed to be large enough to swamp the total capacitance of all internal nodes of the oscillator. Therefore, the total capacitance on all internal nodes is linear, has a high quality factor, and is fairly independent of temperature and device parasitic components.

Since all of the MIM capacitors have the same capacitance, Node A, which is driven by a 10x inverter, transits between high and low voltages much more quickly than Node B, which is driven by a 1x inverter. This behavior ensures that the noise of the oscillator is dominated by the small inverters, considering that the induced jitter at each stage is inversely proportional to the square of the voltage rate of change at its output [13]. Therefore, although the power spectral density of the current noise at the output of the large inverters is approximately ten times larger than that at the output of the small inverters, their jitter contribution is ten times smaller because of the faster voltage rate of change.

Furthermore, note that Node A, which is also the gate of the small inverter, is almost always high when the output of the small inverter, Node B, is transitioning between high and low (Fig. 11). This characteristic means that the bias conditions at which device noise is measured are known and can be used for reliable indirect characterization of device
noise through phase noise measurements for various bias voltages.

Three different asymmetric ROs with minimum channel lengths of 0.18 \( \mu \)m (1\times), 0.54 \( \mu \)m (3\times), and 1.08 \( \mu \)m (6\times) are fabricated on each of the three tiers, resulting in nine different oscillators.

Device noise is characterized by measuring the phase noise of the ROs in the 1/\( f^2 \) region, as shown in Fig. 12. In this region, the 1/\( f \) noise of the device does not affect the device noise. Furthermore, there is a one-to-one correspondence of the phase and device noises in the 1/\( f^2 \) region.

B. Discussion of Device Noise

Phase noise is a function of power and oscillation frequency. Therefore, comparing absolute phase noise across ROs without regard for power consumption and oscillation frequency is meaningless. This is because different ROs have different oscillation frequencies and consume different power, although they might be nominally designed to be the same. Process variations also cause performance differences between nominally similar oscillators.

Delta phase noise (DPN, Fig. 13) is the difference between the measured phase noise and the minimum achievable phase noise \( P_{N_{\text{min}}} (\Delta f) \) [14], [15]

\[
P_{N_{\text{min}}} (\Delta f) \approx \frac{7.33 f_0 k T}{N C \nu_{\text{dd}}^2 (\Delta f)^2} = \frac{7.33 k T}{P_{\text{min}}} \left( \frac{f_0}{\Delta f} \right)^2
\]

where \( P_{\text{min}} = f_0 N C \nu_{\text{dd}}^2, \) \( f_0 \) is the nominal oscillation frequency, \( \Delta f \) is the offset frequency, \( N \) is the number of stages in the RO, \( C \) is the capacitance of each stage, and \( \nu_{\text{dd}} \) is the supply voltage.

As shown in (4), the minimum achievable phase noise compensates for differences in power consumption and oscillation frequencies across ROs. Therefore, DPN, which is the difference between the measured phase noise for a particular oscillator and the minimum achievable phase noise of the same oscillator, compensates for differences in power consumption and oscillation frequencies of different oscillators. Thus, it is a figure of merit for measuring device noise [14], [15]. Lower DPN indicates better device noise performance.

Six different chips were measured, and the average DPN is shown in Fig. 13. The standard deviation of each point is represented by error bars. As shown in Fig. 13, device noise performance improved by 13.7 dB, going from 1\times to 3\times devices. This trend is consistent with a comparable 0.18-\( \mu \)m bulk CMOS process [13]. Further increase of the channel length does not improve the noise performance, which suggests that there is an optimal channel length when trading off between the drive current and the noise performance.

For the 3\times devices, device noise performance is best on Tier A and worst on Tier C. This is consistent with having a temperature gradient between tiers. From measurements, temperature is lowest on Tier A and highest on Tier C. For long-channel devices, this translates to the best device noise performance on Tier A and worst on Tier C, as shown in Fig. 13.

The reverse is true for 1\times devices. The device noise performance is best on Tier C and worst on Tier A, although Tier C has the highest operating temperature. This is because shot noise is dominant in short-channel devices and higher temperatures do not necessarily worsen the device noise performance in this regime [13].

V. CONCLUSION

The performance of 3-D ICs is influenced by thermal and 3-D parasitic effects. Actual circuit performance is difficult to predict as thermal and 3-D packaging effects act in opposite ways. To provide design insight, a stacked wafer 3-D–SOI technology was characterized, and a physics-based thermal model was developed and experimentally verified to be valid for various operating conditions. The resulting thermal model is also independently verified to be applicable for different circuits designed in the 3-D–SOI technology. An application example using the thermal model in electrothermal simulations was also demonstrated.

Electrothermal simulations of 3-D ICs were performed, and simulation results match measured data. If parasitic effects are more significant, the performance of the first (bottom) tier is expected to be the worst. In the cases where the thermal and parasitic effects are comparable, the proposed electrothermal simulations can be used to analyze the performance.
Device noise measured for this technology shows the same trends as a conventional 0.18-μm bulk CMOS technology. Measurement data are also consistent with having a temperature gradient between tiers for long-channel devices. Higher temperatures do not necessarily degrade the noise performance in short-channel devices, as shot noise is dominant.

ACKNOWLEDGMENT

The authors would like to thank DARPA for the support for 3-D IC projects at both Stanford University and Lincoln Laboratory, MIT. The authors would also like to thank key contributors to the 3-D fabrication process at Lincoln Laboratory, including J. A. Burns, C. K. Chen, J. M. Knecht, K. Warner, and D.-R. W. Yost.

REFERENCES


Yi-Chang Lu (S’00–M’05) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1993, the M.S. degree in electrical engineering, the M.S. degree in engineering–economic systems and operations research, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1997, 1999, and 2005, respectively.

Between 1993 and 1995, he was an Engineering Officer with the Naval Surveillance and Communication Command Department, Suao, Taiwan. In 2005, he was a Postdoctoral Scholar with Stanford University. Since 2006, he has been with the Graduate Institute of Electronics Engineering and the Department of Electrical Engineering, National Taiwan University, where he is currently an Assistant Professor. His current research interests include 3-D VLSI architecture and design, design for manufacturability, and technology computer-aided design.

Reza Navid (M’05) received the B.S. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 1996, the M.S. degree in electrical engineering from Sharif University of Technology, Tehran, in 1998, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 2005.

From 1998 to 2000, he was with ParsElectric MFG Corporation, Tehran, where he worked on RF circuits for TV tuners. From January 2000 to May 2001, he was with the University of Michigan, Ann Arbor, as a researcher working on MEMS for wireless communications. He spent the summer of 2001 with Maxim Integrated Products, Hillsboro, OR, where he worked on a 3.125-Gb/s high-speed communication receiver front end. From 2005 to 2008, he was with Rambus Inc., Los Altos, CA, working on high-speed integrated circuits for chip-to-chip and memory link applications. He is currently with True Circuits Inc., Los Altos. His current research interests include low-phase-noise timing circuits for various communication and data conversion systems.
Wei Wang (S’08) received the B.S. degree in electronics from Peking University, Beijing, China, in 2002 and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 2004, where she is currently working toward the Ph.D. degree in the Department of Electrical Engineering.

Her research interests include emerging nonvolatile memory device physics and fabrication, nonvolatile SRAM design, and nonvolatile memory application in 3-D FPGA. She is also involved in research activities focusing on oxide reliability and thermal modeling of 3-D integrated circuits.

Chang-Lee Chen (S’78–M’82–SM’02) received the B.S. degree in electronic engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1974, the M.S. degree in electrical engineering from the University of Cincinnati, Cincinnati, OH, in 1978, and the Ph.D. degree from the University of Michigan, Ann Arbor, in 1982.

He has been with the Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, since 1982 and was involved in the development of GaAs- and InP-based microwave devices, as well as in the design and fabrication of monolithic microwave integrated circuits. Currently, he is working on SOI-based devices, RF circuits, and application to 3-D integration.

Robert W. Dutton (S’67–M’70–SM’80–F’84) received the B.S., M.S., and Ph.D. degrees from the University of California, Berkeley, in 1966, 1967, and 1970, respectively.

He held summer staff positions at Fairchild, Bell Telephone Laboratories, Hewlett-Packard, IBM Research, and Matsushita during 1967, 1973, 1975, 1977, and 1988, respectively. He is currently a Professor of electrical engineering with Stanford University, Stanford, CA, and the Director of Integrated Circuits Laboratory. His research interests focus on integrated-circuit process, device, and circuit technologies—particularly the use of computer-aided design (CAD) in device scaling and for RF applications. He has published more than 200 journal articles and graduated more than four dozen doctorate students.

Dr. Dutton was the recipient of the 1987 IEEE J. J. Ebers and 1996 Jack Morton Awards and 1988 Guggenheim Fellowship to study in Japan. He was elected to the National Academy of Engineering in 1991, honored with the C&C Prize (Japan) in 2000, and was the recipient of the Career Achievement Award (2005) from the Semiconductor Industry Association (SIA) for sustained contributions in support of research that is critical to SIA needs. Most recently, he was the recipient of the Phil Kaufman Award. He was the Editor of the IEEE CAD Journal (1984–1986).