

Maximize CMOS Data Processing Speed Through Low Temperature Operation

Mingchen Zhuang^{1†}, Yangziqian Guo^{2†}, Zixian Liu^{3†}, Yichen Cai^{4†}, Rui Ma^{5*†}

¹*Department of Engineering, University of California, Davis, USA*

²*Department of Engineering, The Ohio State University, Columbus, USA*

³*RCF Experimental School, Beijing, China*

⁴*University College London, London, United Kingdom*

⁵*Jinan New Channel-JUTES High School, Jinan, China*

**Corresponding Author. Email: 2623526194@qq.com*

†These authors contributed equally to this work and should be considered as co-first authors.

Abstract. Low temperature operation Complementary Metal Oxide Semiconductor (CMOS) is a promising method to improve high performance computing after traditional CMOS fails to break through the “power wall” composed of and heat loss and Moore's Law is pushed to the limits of physical scale and technological capability. This study discusses the increasing of data processing speed of CMOS under low temperature operation. The focus is on carrier concentration and mobility, in addition, two experimental methods and theoretical modeling methods for measuring carrier concentration and mobility are presented. and finally discuss the challenges and opportunities of low temperature CMOS.

Keywords: low temperature, CMOS, Data processing, Maximize

1. Introduction

Complementary Metal Oxide Semiconductor (CMOS) technology has become integral to modern electronics, playing a vital role in a wide range of applications. Its widespread use is attributed to several key advantages, including low static power consumption, high noise immunity, and cost efficiency. However, as the demand for higher performance and computational power continues to grow, traditional CMOS technology is encountering significant limitations. In particular, the reduction of chip size to the nanometer scale has led to challenges such as increased power consumption, leakage currents, and reduced device performance—a phenomenon often referred to as the "Power Wall." These issues arise because further scaling of CMOS technology is reaching physical and material boundaries, making it increasingly difficult to improve performance through traditional methods [1].

One promising approach to overcoming these challenges is operating CMOS devices at low temperatures. Lowering the operating temperature can enhance carrier mobility, reduce leakage currents, and improve overall device performance. By minimizing thermal noise and improving electrical characteristics, low-temperature CMOS technology offers the potential to significantly boost performance metrics such as data transmission speed and energy efficiency. In this study, we

summarized previous research data and identified two methods for creating low-temperature environments: cryogenic cooling using liquid nitrogen and thermoelectric cooler refrigeration. Additionally, we compiled data on the changes in carrier mobility, power consumption, threshold voltage, and switching speed over time, finding a 50%-60% improvement across the board, with carrier mobility showing exponential growth. Furthermore, we make a summary that the maximum data processing speed in these scenarios, which indicated that the maximum data processing speed improved at different low temperatures compared to room temperature, with the most significant increase of 50% observed at 77K. Finally, we reached our conclusions and discussed the future directions of low-temperature CMOS for high-performance computing.

2. Research methods

2.1. Review of existing literature

A comprehensive review of existing literature was conducted to gather insights into the effects of low temperatures on CMOS technology. The review covered key topics such as noise reduction, carrier mobility, power consumption, and switching speed in CMOS circuits.

2.2. Comparative analysis

In this article, much of the discussions were constrained to 70 K to 100 K range since it is believed to provide the best device performance at reasonable cooling cost for the majority of CMOS technology applications.

2.2.1. Power consumption

In recent years, a research group investigated that the benefits of operating CMOS technology at low temperature. At 77 Kelvin, there is a 7 times power advantage at iso-frequency and a 1.4 times performance advantage at iso-power compared to room temperature [2]. Figure 1 illustrates the normalized performance across different temperatures for various supply voltages, highlighting the corresponding improvements relative to the room temperature (RT) design. Additionally, Figure 2 presents a plot of performance per watt versus overall performance across temperatures and supply voltages. This figure demonstrates the iso-power performance enhancement and indicates over a 4x power improvement at iso-performance by leveraging low supply voltage (low-VDD) [3,4].

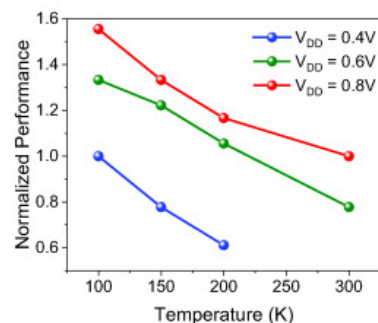


Figure 1. Normalized performance of 64-bit Arm Cortex-A53 across multiple supply voltages at different temperatures. At 300 K, devices cannot operate at 0.4 V and hence no data point [5]

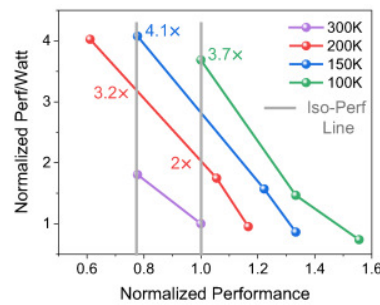


Figure 2. Performance per watt versus performance of 64 bit Arm Cortex-A53 indicating up to 4x improvement at iso-frequency by scaling down the temperature from 300 K to 150 K and corresponding supply voltages from 0.6 V to 0.4 V and up to 3.7x by going from 300 K to 100 K and reducing supply voltage from 0.8 V to 0.4 V. Notice the 300 K 0.6 V frequency is lower than 0.8 V value [5]

2.2.2. Carrier mobility

As the temperature decreases, mobility increases significantly—by a factor of four to six—due to a reduction in atomic lattice vibrations, as shown in Fig. 3 [4].

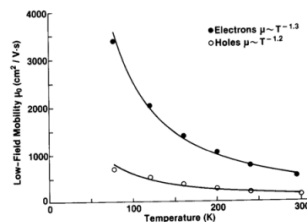


Figure 3. Carrier mobility plotted versus temperature [6]

3. Effects of low temperature on CMOS

3.1. Thermal noise reduction

The reduction of thermal noise in CMOS devices through low-temperature operation is a highly effective method for enhancing the performance of electronic circuits, particularly in high-speed and low-power applications. Thermal noise, also known as Johnson-Nyquist noise, is directly proportional to the operating temperature and the resistive elements of a circuit. As the operating temperature decreases, the thermal agitation of charge carriers lessens, reducing the overall noise levels in CMOS transistors [2]. However, the behavior of noise in CMOS is complex, especially at lower frequencies.

3.2. Flicker noise in n-MOS and p-MOS devices

Flicker noise, the dominant source for low-frequency noise in CMOS transistors, exhibits distinct behaviors in n-MOS and p-MOS devices. The paper written by Chang et al. (1994) analyzes flicker noise in both types of transistors under various conditions [2].

As shown in fig. 4, the results indicate that flicker noise in n-MOS devices is primarily due to carrier-density fluctuations, which arise from the random capture and emission of charge carriers by interface traps in the gate oxide. Additionally, the study found that noise activity in n-MOS devices

is not relatively sensitive to temperature, which means lowering the temperature does not significantly reduce flicker noise in n-MOS devices. The noise spectrum remains relatively stable from room temperature down to 5K. In contrast, p-MOS devices are more sensitive to temperature changes. It can be told from fig.5, as the temperature decreases, the mobility of holes in p-MOS devices increases, leading to a reduction in flicker noise. This is especially evident when the temperature drops to around 150K [2]. However, the flicker noise begins to rise again at temperatures lower than 150K. This increase is caused by the generation-recombination (g-r) noise, resulting in more trapped carriers and noise. This suggests the existence of both opportunities and challenges for noise reduction in CMOS circuits.

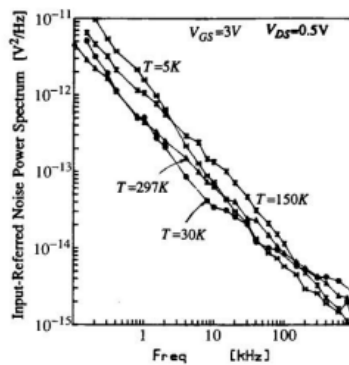


Figure 4. Noise spectra for the $80 \times 6 \mu\text{m}^2$ n n-channel transistor from Fabricator H from room temperature down to 5K

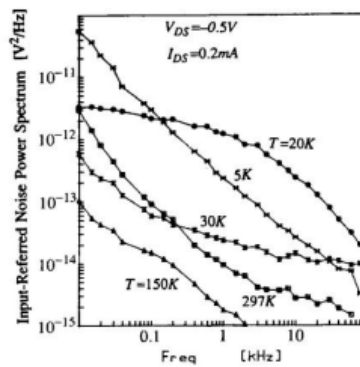


Figure 5. Noise spectra for the $80 \times 6 \mu\text{m}^2$ P-channel transistor from Fabricator H from room temperature down to 5K at a fixed drain voltage of -0.5V and a drain current of 0.2mA

3.3. Enhanced carrier mobility

Carrier mobility in CMOS devices is a critical parameter that influences the overall performance of the transistor, especially under different temperature conditions. As temperature decreases, carrier mobility typically increases due to the reduction in scattering mechanisms like phonon scattering.

3.3.1. Phonon scattering reduction

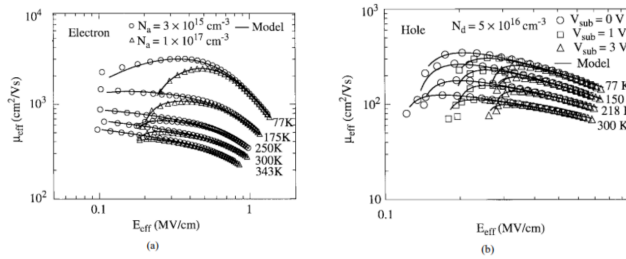


Figure 6. μ_{eff} versus effective electric field (E_{eff}) dependence of the inversion layer mobility for (a) n-channel devices at temperatures of 77 K, 175 K, 250 K, 300 K, and 343 K. (b) p-channel devices at temperatures of 77 K, 150 K, 218 K, and 300 K. Symbols correspond to experimental mobility data, while solid lines represent the simulated mobility data given by

At room temperature, phonon scattering is the dominant mechanism that limits carrier mobility. Phonon scattering arises due to the interaction of electrons and holes with the vibrating atoms in the semiconductor lattice.

According to A Temperature-Dependent MOSFET Inversion Layer Carrier Mobility Model for Device and Circuit Simulation by Cheng and Woo (1997), the reduction in phonon scattering is a major factor causing the increase in electron and hole mobilities at low temperatures [3]. As shown in figures 1a for n-channel devices and 1b for p-channel devices, electron mobility increases significantly as the temperature drops from 300 K to 77 K due to reduced phonon scattering.

3.3.2. Temperature dependence of electron and hole mobilities

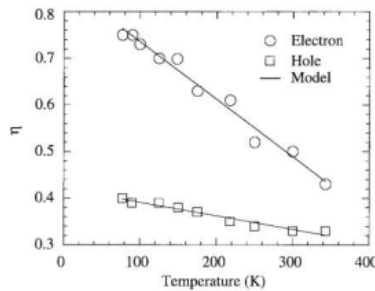


Figure 7. Temperature dependence of the electric field parameter for electrons and holes at temperature range from 77K to 343K

As shown in Figure 7 from Cheng and Woo's study [3], changing the temperatures from 343K to 77K, electron mobility shows a sharp increase in the electrical field parameter, which means electron mobility is significantly dependent on temperature changes. On the other hand, hole mobility improves after lowering the temperature, but with a relatively slower increase in the electric field parameter compared to electron mobility. This indicates that electron mobility shows a stronger dependence on the electric field at lower temperatures compared to hole mobility. The increased mobility at low temperatures for electrons and holes improves the performance of CMOS devices in cryogenic applications [7].

3.4. Power consumption minimization

In a low temperature environment, reducing the power consumption of CMOS circuits can be achieved mainly by reducing leakage current and dynamic power consumption. For CMOS circuits, low temperature not only improves the energy efficiency of the system, but also provides favorable technical support for high performance and low power consumption applications.

3.4.1. Power consumption in CMOS circuits

In CMOS technology, power consumption is mainly composed of two parts. They are dynamic power consumption and static power consumption. Dynamic power consumption mainly comes from the switching of the circuit, that is, the charging and discharging behavior of the capacitor in the circuit, while static power consumption is related to leakage current.

Dynamic power consumption:

$$P_{dynamic} = \alpha cv^2f$$

α is the activity factor

c is the capacitance

v is the operating voltage f is frequency

Actually, at low temperature situations, capacitance does not change significantly, but for reducing the operating voltage and minimizing switch events can reduce dynamic power consumption effectively.

Static power consumption:

$$p_{static} = I_{leak}v$$

A low temperature environment can reduce the leakage current of transistors, thereby effectively reducing static power consumption.

3.4.2. Effect of low temperature on power consumption

Leakage current in a transistor is significantly reduced at lower temperatures because lower temperatures reduce the thermal energy of carriers, minimizing unnecessary current leakage.

3.4.3. Strategies for power consumption minimization

Lowering the operating voltage is one of the most effective strategies to reduce power usage.

By using lower voltages and maintaining performance without degradation, operating at low temperatures greatly lowers dynamic consumption. According to calculations, the minimum power's temperature behavior is highly dependent on the assumptions made about circuit speed. High-performance logic circuits require a current density J_{ON} that is rather high (around the industry standard of 600A/m). In this instance, as shown by the upper solid line in Figure 8, the minimal power and the optimal VDD both gradually decrease with temperature and reach saturation near 100K.

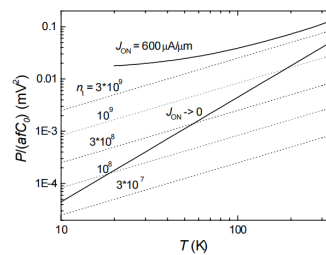


Figure 8. The relation of between temperature and the power consumption [8]

3.5. Changes in threshold voltage and switching speed

“The temperature has an impact on transistor mobility characteristics and threshold voltage V_{th} . The performance of the device increases as the temperature rises, and it decreases as the temperature falls. Off-state power consumption and junction leakage current decrease as V_{th} rises and temperature decreases mobility. A PD transistor can transition to FD mode at low temperatures because the depletion region enlarges.”

As was already mentioned, increased mobility is a result of low temperatures. Nonetheless, increased mobility also implies that the circuit can maintain an effective switching speed at lower voltage, which is highly beneficial for lowering dynamic power consumption and enhancing system performance in general. "At low temperatures, the behavior of V_{th} varies depending on the length of the channel. The transition from PD to FD mode is shown in a change in the slope dV_{th}/dT at 300 K for a long device ($L_g = 1\mu m$). The transition area moves to a lower temperature range (140 K) for very small channels in the nanometer length range ($L_g = 40\text{ nm}$) [5].

The second impact of the pockets is observed in the PD linear region ($140\text{ K} < T < 300\text{ K}$), where the shorter channel ($L_g = 40\text{ nm}$) has a larger slope dV_{th}/dT . In order to explain this, Eq. (2) incorporates the fluctuation of N_{eff} versus gate length, with short channels having a greater effective doping level. Figure 9: Characteristics of the threshold voltage versus temperature for a range of gate lengths and temperatures, the slope of $V_{th}(T)$ was examined.

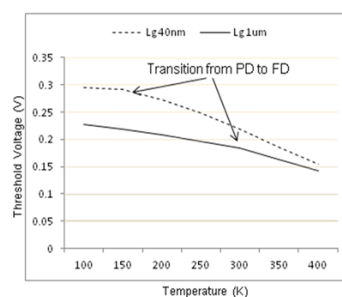


Figure 9. the relation of between temperature of Threshold the Voltage

4. Conclusion

In summary, the low temperature CMOS technology provides a large number of advantages, including reducing thermal noise, enhancing the carrier mobility. This is quite critical for improving signal clarity and reliability in high-performance electronic systems.

Moreover, decreasing the power is another key improvement. Lower temperatures reduce leakage currents, which directly influences energy efficiency. This is a necessary in modern electronics,

where energy efficient design is a priority. The research also observes that the changes in threshold voltage and the switching speed, which provides the possibility of optimizing CMOS applications in several devices.

Finally, the researches highlight the potential of the CMOS technology, which could be applied to the next generation devices. As the demand for faster, smaller, and more efficient devices, the low-temperature operation would play an important role in the application of the next generation devices.

5. Future potential and applications

In summary, the low temperature CMOS technology provides a large number of advantages, including reducing thermal noise, enhancing the carrier mobility. This is quite critical for improving signal clarity and reliability in high-performance electronic systems.

Moreover, decreasing the power is another key improvement. Lower temperatures reduce leakage currents, which directly influences energy efficiency. This is a necessary in modern electronics, where energy efficient design is a priority. The research also observes that the changes in threshold voltage and the switching speed, which provides the possibility of optimizing CMOS applications in several devices.

Finally, the researches highlight the potential of the CMOS technology, which could be applied to the next generation devices. As the demand for faster, smaller, and more efficient devices, the low-temperature operation would play an important rule in the application of the next generation devices.

Author contributions

Mingchen Zhuang, Yangziqian Guo, Zixian Liu, Yichen Cai and Rui Ma contributed equally to this work and should be considered co-first authors.

References

- [1] Cheng R, Li B, Wang Z W, et al. Low-temperature CMOS technology for high-performance computing: development and challenges (in Chinese). *Sci Sin Inform*, 2024, 54: 88–101, doi: 10.1360/SSI-2023-0347
- [2] Chang, J., Abidi, A. A., & Viswanathan, C. R. (1994). Flicker Noise in CMOS Transistors from Subthreshold to Strong Inversion at Various Temperatures. *IEEE Transactions on Electron Devices*, 41(11), 1965–1971.
- [3] Cheng, B., & Woo, J. (1997). A Temperature-Dependent MOSFET Inversion Layer Carrier Mobility Model for Device and Circuit Simulation. *IEEE Transactions on Electron Devices*, 44(2), 343-345. <https://doi.org/10.1109/TED.1997.560090>
- [4] Zhang, Y., Xu, J., Lu, T.-T., Zhang, Y., Luo, C., & Guo, G. (2021). Hot Carrier Degradation in MOSFETs at Cryogenic Temperatures Down to 4.2 K. *IEEE Transactions on Device and Materials Reliability*, 21(4), 620-625. <https://doi.org/10.1109/TDMR.2021.3124417>
- [5] Saligram, R., Raychowdhury, A., & Datta, S. (2024). The future is frozen: cryogenic CMOS for high-performance computing. *Chip*, 3(1), 100082.
- [6] Hu, J. (2006). Low temperature effects on cmos circuits. Dept. Elec. Eng., Northwestern Univ, Evanston, IL.
- [7] NehaGoel, Ankit Tripathi, "Temperature effects on Threshold Voltage and Mobility for Partially Depleted SOI MOSFET" *International Journal of Computer Applications*, vol.42, no. 21, pp. 56, March 2012.
- [8] Victor Sverdlov, Yehuda Naveh, Konstantin Likharev, "Temperature scaling of CMOS circuit power consumption" *Physica E: Low-dimensional Systems and Nanostructures*, vol. 18, no. 1-3, pp. 2, August, 2002.