Cryogenic CMOS for Quantum Processing: 5-nm FinFET-Based SRAM Arrays at 10 K

Shivendra Singh Parihar, *Member, IEEE*, Victor M. van Santen, *Member, IEEE*, Simon Thomann, *Member, IEEE*, Girish Pahwa, *Member, IEEE*, Yogesh Singh Chauhan, *Fellow, IEEE*, and Hussam Amrouch, *Member, IEEE*

***Abstract*— In this work, we are the first to investigate and model the characteristics of a commercial 5nm FinFET technology from room temperature (300K) all the way down to cryogenic temperature (10K). We focus on SRAM circuits demonstrating how cryogenic temperatures impact their power, delay, and reliability. SRAM memories are key components in quantum read-out and control circuits, and therefore characterizing their key figure of merits when building cryogenic-CMOS circuits is essential. To achieve that, we first measure the electrical characteristics of nFinFET and pFinFET devices from 300K down to 10K. Then, we carefully calibrate the cryogenic-aware BSIM-CMG, which is the first industry-standard compact model for FinFET technologies designed for cryogenic temperatures. This enables us to reproduce the experimental data in which SPICE simulations come with an excellent agreement with the measurements. Using our well-calibrated transistor models, we simulate a complete 32-bit SRAM memory array, including a write driver, sense amplifier, pre-charger, and output latch. Then, we investigate how cryogenic temperatures impact the SRAM read and write delays at several stages during the operation, as well as the power and energy. For a more comprehensive analysis, we perform our studies for different SRAM types covering high-density, high-performance, and low-voltage cells. All transistor and SRAM analyses are performed at both room temperature and cryogenic temperature to obtain detailed comparisons revealing the exact role that cryogenic temperature plays in SRAMs. All in all, we demonstrate that commercial 5nm FinFET is indeed suitable for cryogenic-CMOS circuits required in quantum processors, revealing that the performance of SRAMs at 10K does improve while power and energy consumption are reduced. Nevertheless, SRAM reliability is more challenging in which noise margins need to be carefully engineered to remain sufficient at 10K.**

***Index Terms*— Device characterization, device modeling, 5nm FinFET, SRAM, cryogenic CMOS, reliability.**

I. Introduction

CRYOGENIC CMOS circuits are crucial to enable efficient large-scale quantum computing. Quantum computers operating at near 0 K are considered to be the next leap in computation, allowing previously infeasible calculations to become feasible. However, a major concern in quantum computing is the operation of the supporting circuitry at room temperature (300 K). This results in limited scalability and integration of quantum computers, hindering the total number of Quantum Bits (Qubits) that can be implemented and reliably processed. Currently, the Qubits, which operate at close to absolute zero (e.g., 100 mK), are supported by circuitry at 300 K. This high thermal gradient between the classical and quantum layers and the limited available cooling power at cryogenic temperatures (in the order of 100 mW) enforce limited connectivity between the quantum and classical support layer to limit heat flux towards the cryogenic qubits. The wires, connecting the quantum bits with the processing circuits outside, are typically long to ensure that the heat dissipated from such circuits does not propagate and leak into the cryogenic refrigerator and jeopardize the quantum bits. However, a few long wires limit throughput and add considerable latencies via wire propagation delay and thus severely hinder the control of the quantum bits and their read-out. Control of the Qbits and quantum processing occurs in the range of ns and ms, respectively, which is barely met by room temperature CMOS circuitry. Hence, quantum computers with CMOS circuitry at 300 K are severely limited in scale (number of Qbits) and applications (since throughput or latency limitations limit possible quantum algorithms).

Throughout this paper, CMOS technologies represent different technology nodes (e.g., 5 nm, 14 nm, 28 nm, etc.) and the CMOS transistor refers to FinFETs, bulk- and SOI-MOSFETs. CMOS circuits are sub-circuits in control and readout circuitry (e.g., SRAM, ADC, Amplifier, etc.).

A. Cryogenic CMOS Circuits

Since the origin of the above-mentioned challenge is the high thermal gradient between supporting circuitry and quantum bits, cryogenic CMOS technology is proposed to alleviate this issue [1], [2]. Cryogenic CMOS limits the

Manuscript received 20 November 2022; revised 22 February 2023 and 5 April 2023; accepted 10 May 2023. Date of publication 1 June 2023; date of current version 28 July 2023. This work was supported in part by the Swarna Jayanti Fellowship under Grant DST/SJF/ETA02/2017-18 and in part by the Deutsche Forschungsgemeinschaft (DFG) AM 534/3-1 under Grant 428566201. This article was recommended by Associate Editor C. Wang. *(Corresponding author: Shivendra Singh Parihar.)*

Shivendra Singh Parihar is with the Semiconductor Test and Reliability (STAR) Research Group, University of Stuttgart, 70174 Stuttgart, Germany, and also with the NANOLAB, Department of Electrical Engineering, Indian Institute of Technology (IIT) Kanpur, Kanpur 208016, India (e-mail: parihasa@iti.uni-stuttgart.de).

Victor M. van Santen and Simon Thomann are with the Semiconductor Test and Reliability (STAR) Research Group, University of Stuttgart, 70174 Stuttgart, Germany (e-mail: van-santen@iti.uni-stuttgart.de; thomann@iti.uni-stuttgart.de).

Girish Pahwa is with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720 USA (e-mail: pahwa@berkeley.edu).

Yogesh Singh Chauhan is with the NANOLAB, Department of Electrical Engineering, Indian Institute of Technology (IIT) Kanpur, Kanpur 208016, India (e-mail: chauhan@iitk.ac.in).

Hussam Amrouch is with the AI Processor Design, Technical University of Munich (TUM), 80333 Munich, Germany, and also with the Munich Institute of Robotics and Machine Intelligence (MIRMI), 80992 Munich, Germany (e-mail: amrouch@tum.de).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCSI.2023.3278351.

Digital Object Identifier 10.1109/TCSI.2023.3278351

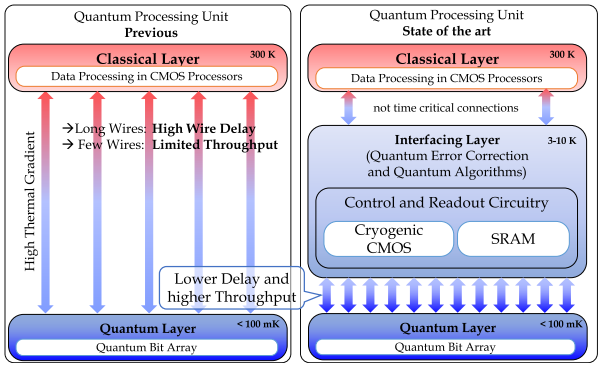


Fig. 1. Schematic view of different layers of quantum computers partitioned over the temperature (inspired by [1]).

thermal gradient by 30 times from a 300 degrees difference (100 mK to 300 K) to a 10 degrees difference (100 mK to 10 K) and thus allows integration within the same chip significantly increasing connectivity [3], [4] (Fig. 1). The main focus of cryogenic CMOS is to reduce the thermal gradient between qubit and control circuitry and not at all to improve the circuit or transistor performance. The operation of cryogenic CMOS at mK increases the cooling cost significantly [2]. The limited power handling capability of the state-of-the-art dilution refrigerators imposes a limit for operating the cryogenic CMOS circuitry at its higher temperature stages to reduce the thermal load [5]. Control circuitry operation at cryogenic temperature also lowers power dissipation. For example [6] shows that quantum control circuitry at 300 K consumes $>$ 1 W of the power while a cryogenic IC-based circuit operating at 3 K consumes $<$ 2 mW of power. Operating CMOS circuits at cryogenic temperatures (4.2 K to 33 K) enables better quantum control, lower latency, and higher throughput [2]. This, in turn, enables a wider range of quantum algorithms and applications since considerably more processing steps are possible within the quantum coherence of approximately 10 ms.

1) Challenges for Cryogenic CMOS: There are two key challenges with cryogenic CMOS circuits. The first challenge is the limited power budget. Cryogenic cooling is incapable of cooling the hundreds of Watts of traditional air- or water-cooled systems at 300 K. Instead, the power dissipation must be limited to $<$100 mW [1], [7]. Therefore, selecting the CMOS technology with the lowest possible power consumption is highly desirable. This power challenge is exacerbated by the second challenge, that quantum bits need to be controlled (apply quantum operations) in the nanosecond scale [8]. Hence, this ultra-low power operation requires very small delays (high performance). This again makes the cutting-edge CMOS technology desirable as it is fast enough at ultra-low power operation modes. Therefore, CMOS technology is *quantum suitable* when offering (i) High-performance circuits with small latency in which processing of quantum data can be done within milliseconds (e.g., processing within 10 ms [8], [9]). (ii) Short control loop of quantum bits (e.g., less than 10 ns [8]). (iii) Ultra-low power consumption (e.g., less than 100 mW [1], [7]).

2) Architecture for Quantum Processing With Cryogenic CMOS: The traditional quantum processor setup has to change to improve quantum processing with cryogenic CMOS circuits. The traditional quantum processor setup is shown on the left side of Fig. 1, and the new one is on the right. The state-of-the-art quantum processors consist of the quantum layer for processing near absolute zero, an interface layer for controlling and readout operations on the Qbits in cryogenic CMOS, and classical room temperature computers for correlating the quantum information into regular Boolean logic [1], [10]. The works in [3], [4], and [11] show potential implementations of such a quantum interface layer. However, to integrate a cryogenic CMOS circuit into such a quantum computer, the cryogenic CMOS technology needs to ensure the two quantum-compatible conditions (high performance, ultra-low power) to not disrupt quantum operations.

B. Cryogenic Transistors for Cryogenic-CMOS Circuits

This section presents how different CMOS technologies behave at cryogenic temperatures. The goal is to explore which technology is suitable for the cryogenic interface layer in quantum processors.

1) Challenges of Cryogenic Transistors Technology: Operating transistors at cryogenic temperatures provides great benefits, yet some challenges emerge at such low temperatures. In older technologies, enhanced threshold voltage (V_{TH}) variation and carrier freeze-out severely deteriorated transistor performance. Researchers also observed abnormal behavior (kink effect) in the drain-source current [12]. Additionally, hysteresis is observed in the output characteristics of the transistor when a transient gate pulse is applied [12]. From the circuit design perspective, transistor mismatch and increased V_{TH} are particularly challenging [13]. All these changes in transistor behavior are critical enough to affect circuit design.

2) Impact of Technology Scaling on Cryogenic Transistors: The transistor geometry scaling has reduced the feature size from 1 µm to 5 nm. The reduced supply voltage (V_{DD}) and thinner gate oxide with CMOS technology scaling led to an increased electric field inside the channel, improving the OFF and ON-state performance. Field-dependent mobility reduction and decreased V_{DD} lowered the impact-ionization and subsequent effects (kink effect, hysteresis). Thus, the kink effect is mainly observed in bulk MOSFET technologies [12], and hence this challenge became negligible in FinFET technologies. The other key changes are observable in all CMOS technologies, with weaker effects in advanced technology.

The transistors fabricated below 180 nm show similar sub-threshold and ON-state performance improvement when switching from 300 K to cryogenic temperatures [13], [14], [15]. Even though both 160 nm and 40 nm benefit from operating at cryogenic temperatures, the 40 nm is faster due to improved gate electrostatic and short-channel effects [14]. Improvement in the ON-state current (I_{ON}) with channel length reduction is accompanied by enhanced drain-induced barrier lowering (DIBL). Transistors operating at a low temperature suffer less from the DIBL, e.g., 28 nm CMOS technology n- and p-type transistors operating at 77 K show improvement in the DIBL by 15 mV and 8 mV, respectively [15].

In summary, with each newer technology, when switching from 300 K to cryogenic temperatures, the significant changes

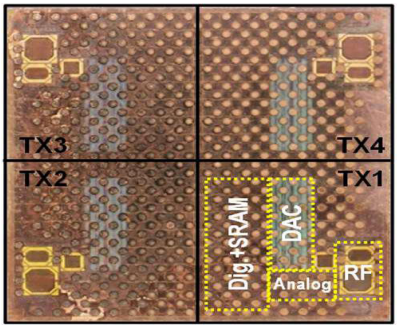


Fig. 2. First generation qubit drive controlling SoC (Horse Ridge 1) highlighting the use of SRAM as a memory to store the signal modulation information (taken from [23]).

are diminishing kink effects, increased mobility and V_{TH}, lower SS and I_{OFF}.

3) FinFETs at Cryogenic Temperatures: To address higher gate leakage and enhanced short-channel effects in bulk CMOS transistors, Fin-based multi-gate transistors (FinFETs) were adopted [16]. An undoped FinFET channel with better gate electrostatic and reduced interface trap density significantly lowers the I_{OFF} (leakage current). FinFET in the 22 nm technology shows 37 % improvement in performance with 50 % less power consumption than bulk MOSFET at 300 K [16], [17], [18]. Most works studied FinFETs in liquid nitrogen (77 K), while our work studies down to 10 K, yet most observations from 77 K hold at lower temperatures such as 10 K. At 77 K, 10 nm technology FinFETs show 48 % to 58 % mobility improvement, 0.01 × OFF-current, and ∼39 % speed improvement compared to 300 K [19].

Smaller technologies with lower V_{TH} and higher I_{ON} provide better Performance. Authors in [20] have reported that 5 nm FinFETs provides ∼18 % more drive current and 15 % more performance than the 7 nm FinFETs. Table I provides an overview of different technology's transistor performance improvement at cryogenic temperatures.

C. Dual Purpose of Evaluating SRAM Cells

Static Random-Access Memory (SRAM) is one of the most important circuits in CMOS technology. In this work, the majority of our evaluation is performed on SRAM cells. This is due to two key reasons. First and foremost, SRAM cells are excellent technology evaluation circuits and, as such, frequently employed to highlight the advantages and drawbacks of specific CMOS technologies. As our goal is to evaluate whether 5 nm FinFET technology is suitable for cryogenic operation in the interface layer of quantum processors, SRAM circuits are the obvious choice.

Additionally, SRAM cells are directly employed in these interface layers, as fast data storage is required [23], [24]. Hence, SRAM cells are not just a stand-in for cryogenic CMOS circuits to evaluate the 5 nm FinFET technology at 10 K; their direct use is also essential to store data fast and reliably within the interface layer of the quantum processor (compare interfacing layer in Fig. 1).

1) SRAM Cells to Evaluate CMOS Technologies: Evaluating SRAM area, power, and delay provide key insights into the capabilities of CMOS technology, and as such, it is used as a technology benchmarking tool [16], [17], [18], [20], [25], [26], [27], [28]. The high-density cell (HDC), low-voltage cell (LVC), and high-performance cell (HPC) are commonly used for technology benchmarking. For instance, HDCs are employed in processor caches, where memory should have the highest density possible to reduce the silicon area of large SRAM memory arrays. LVC cells are used in low-power applications (e.g., the internet of things) to reduce the considerable power consumption of SRAM cells. Lastly, HPC cells are employed where performance is of utmost importance, e.g., the register file of processors.

In summary, HDC encapsulates the area density of CMOS technology. LVC shows if the technology can control leakage and operate at sub-nominal V_{DD}, while HPC highlights the performance of the technology.

2) SRAM Cells as Cryogenic Memory in Quantum Computing: SRAM is one of the fastest memory used in CMOS technology and the most widely used for high-performance applications at 300 K, viz. buffers, cache, etc. This makes SRAM an inherent choice for memory by circuit designers for the quantum processor. Authors in [1], [9], and [23] show the use of SRAM as a memory element to store the intermediate signal's amplitude and phase modulation information generated by qubits (Fig. 2). Other memory technologies would not be suitable as performance in this data storage is of critical importance. The authors report that the amplitude and phase modulation information for pulse shaping is stored in on-chip SRAM, referenced by a look-up table (LUT) that defines eight instructions per qubit, each with a phase offset to reuse envelopes for different rotation axes efficiently. Employing SRAM as a memory technology allowed for envelopes as short as 41 µs [9], [23], highlighting the need for SRAM. The use of SRAM as an envelope memory in [9] suggests that for a successful control and read-out operation, a frequency of 1.5 GHz is necessary, enforcing the SRAM delay below 0.66 ns. Authors in [9], [24], and [29] have shown that the memory and digital logic block consumes 8 mW to 16 mW, 10 mW to 140 mW and 330 mW of power, respectively. Authors in [30] also reported the SRAM access delay, read and write power consumption of 0.95 ns, 1.87 mW, and 0.41 mW, respectively at 4.2 K.

Our novel contribution within this work is as follows:

(1) This is the first work to evaluate commercial mature 5 nm FinFET technology for cryogenic applications.
(2) We present a careful physical and electrical transistor model parameter calibration for commercial 5 nm FinFET at different cryogenic temperatures from 300 K to 10 K.
(3) An in-depth evaluation of SRAM arrays (instead of isolated memory cells) at high-density, high-performance, and low-voltage cells in commercial 5 nm FinFET at different cryogenic temperatures from 300 K to 10 K.
(4) We present circuit design insights for 5 nm FinFET-based cryogenic SRAM performance improvement by relating the transistor behavior with different SRAM cell configurations and highlighting the impact of different sub-circuits in an SRAM array.

All in all, this work focuses on the comprehensive characterization of state-of-the-art 5 nm FinFETs, as well as the

TABLE I
COMPARISON OF CRYOGENIC PERFORMANCE OF DIFFERENT CMOS TECHNOLOGIES

		5 nm FinFET (this work)		FinFET 10 nm [19]	FinFET 14 nm [21]	40 nm [22]	160 nm [22]
V_{DD}		0.75 V		0.75 V	0.8 V	1.1 V	1.8 V
Temperature		77 K	10 K	77 K	4.6 K	4.2 K	4.2 K
$\|V_{TH}\|$	n-type	+45.10 %	+47.28 %	–	+20.00 %	+18.18 %	+27.27 %
	p-type	+37.71 %	+38.98 %		+12.03 %	+27.27 %	+40.00 %
$\|SS\|$	n-type	-77.37 %	-82.99 %	-68.16 %	-69.23 %	-68.52 %	-73.79 %
	p-type	-71.04 %	-77.82 %	-65.86 %	-76.92 %	-69.66 %	-58.17 %
I_{ON}	n-type	+7.13 %	+3.09 %	–	-3.8 %	1.5 %	5 %
	p-type	+1 %	-2.7 %		-4.1 %	0 %	0 %
I_{OFF}	n-type	$10^{-5}\times$	$10^{-5}\times$	–	–	$10^{-2}\times$	$10^{-1}\times$
	p-type	$10^{-5}\times$	$10^{-5}\times$			$10^{-2}\times$	$10^{-2}\times$

*All results at cryogenic temperatures are relative to 300 K of respective technologies.

performance evaluation of 5 nm FinFET-based SRAM cells for quantum computing applications. Analyzing the impact of leakage current and RC extraction, designing and exploring the impact of different memory designs and array sizes are out of the scope of this work.

II. RELATED WORK

Transistors characteristics are affected by temperature, which further impacts the circuit design [31]. The advancement in quantum computing has further fueled cryogenic CMOS research. In the following, we briefly summarize the recent developments in the field of cryogenic CMOS.

A. Older CMOS Technologies

1) Transistor's Cryogenic Performance: The work in [15] reports the impact of transistor width in 28 nm CMOS technology down to 77 K. The n-type MOSFETs exhibit a reverse narrow-width effect, while p-type MOSFETs show anomalous trends with width reduction. References [1] and [23] present the performance and integration capability of 22 nm FinFET technology operating at cryogenic temperatures for quantum computing application. A functional Silicon-on-Chip for quantum computing application presented in [23] demonstrates the capability and benefits of FinFET technologies. Transistor characterization at 77 K of the 14 nm FinFET technology in [21] shows that a customized cryogenic CMOS technology (accounting for increased V_{TH}) can benefit from a considerable power reduction by operating at a lower V_{DD}. The cryogenic performance of the 10 nm FinFET technology has been characterized in [19]. The improved SS at 77 K offers a similar speed as at 300 K with 0.27 × reduced power. Most of these studies were either based on an older technology [1], [15], [23] or have performed transistor characterization as low as 77 K [19], [21].

2) SRAM at Cryogenic Temperatures: Quantum driver circuits in [1], [9], and [23] have used SRAM as a memory element to store the instruction for read-out operation. However, they have not shown a detailed performance analysis (impact of different cell types, V_{DD}, and temperatures) of SRAM and used much older 22 nm FinFET technology.

A regular V_{TH} transistor-based 6T high-density SRAM cell fabricated in 10 nm FinFET technology shows an increment in the Read Noise Margin (RNM) while operating at 77 K [32]. A low V_{TH} transistor-based HDC offers a reduction in RNM. Apart from noise margin, a 128 × 128 SRAM array of both regular and low-V_{TH} HDC show significant improvement in the energy-delay product compared to 300 K. Although a detailed analysis of the SRAM cell is presented in [32], the operating temperature was limited only up to 77 K, which is not suitable for interfacing layers operating at ∼3 K to 10 K.

B. The 5 nm Technology

1) Transistor's Cryogenic Performance: As per the authors' knowledge, cryogenic performance evaluation of the commercial 5 nm technology remains untouched. Due to the remarkable improvement in electrical performance of CMOS technology at cryogenic temperature and their possible use in quantum computers, in this work, we include the cryogenic characterization of 5 nm technology FinFETs.

2) SRAM at Cryogenic Temperatures: There are a few studies on the 5 nm SRAM; however, they are either TCAD based [33] or were presented as a technology benchmark only [20], [34]. Reduction in V_{DD} and the increase in temperature results in the degradation of the noise margin. Previous studies on 5 nm SRAM lacked a detailed investigation of each variable (V_{DD}, temperature) impacting the SRAM performance. As per the authors' knowledge, this is the first work demonstrating the cryogenic performance of 6T-SRAM. We present a comprehensive simulation analysis of the SRAM array at cryogenic temperatures to assess the usability of mature commercial 5 nm technology for quantum computing applications without cryogenic-specific alterations.

III. MEASUREMENT SETUP

To evaluate the commercial 5 nm technology, we need to characterize (measure) it first. The Fin-based transistors used in this study are n- and p-FinFETs fabricated in 5 nm commercial bulk FinFET technology. Static electrical characterization of these FinFETs is being performed over a wide temperature range, starting from 10 K to 300 K.

Static electrostatic measurements of the minimum channel length n- and p-type multi-fin, multi-finger FinFETs have been performed at 300 K and in the cryogenic temperature regime (starting from 10 K). Lakeshore CRX-VF, a cryogenic probe station, has been used for on-wafer DC measurements (Fig. 3). The probe station consists of a 51 mm diameter sample stage,

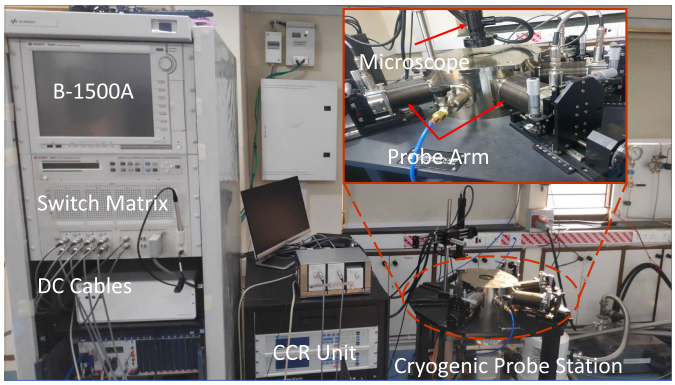


Fig. 3. On-wafer lakeshore's cryogenic probe station.

a vacuum pump, a two-stage closed cycle refrigerator unit (CCR), and probe handlers/positioners. During the cooling process (and only here, not during measurement), the wafer glued on the sample stage by vacuum grease was kept at a slightly higher temperature to avoid condensation. The thermal load imposed by the number of probes (probe holding stage temperature is 15 K) limits the substrate stage temperature between 3.5 K to 8.5 K. Hence, a minimum temperature of 10 K is chosen to avoid temperature variations throughout the measurement. Once the vacuum level is below 5 μbar, the CCR unit starts the cooling procedure toward the 10 K temperature. DC voltages and currents at all terminals have been measured using Keysight's B1500A controlled via PC using the General Purpose Interface Bus.

IV. Compact Model Calibration

To extract an optimal model parameter set, which closely captures the key electrostatic phenomena of the advanced FinFETs, we use the cryogenic-aware industry-standard BSIM-CMG compact model [35]. The device modeling tool ICCAP and HSPICE simulator are used for parameter extraction. A comprehensive characterization approach is followed before parameter extraction to ensure the accuracy of the measurements.

A. Parameter Extraction at 300 K

The most important step in model parameter extraction is setting the correct environment. We start with setting the ambient temperature for simulations in each setup. The model extraction begins with the process-related parameters, viz. flat-band voltage, channel doping, and metal gate work function, followed by the model parameters having the strongest influence in the sub-threshold and linear regime of the transistor operation. The work function, interface trap, and source/drain coupling capacitance model parameters (PHIG, CIT, and CDSC, respectively) from BSIM-CMG are used to accurately reproduce the transistor's sub-threshold behavior. We use mobility model parameters (U0: low field mobility, UA and EU: phonon/surface roughness scattering dependent mobility, UD: Coulomb scattering dependent mobility, and ETAMOB: effective field parameter) to extract the low-field mobility and vertical field-induced mobility degradation from the transfer characteristics at lower drain biases (V_{DS}) in moderate inversion. The strong inversion regime is used for series resistance extraction while the transistor operates in the linear region. At V_{DS} = 0.75 V, sub-threshold swing and Drain Induced Barrier Lowering (DIBL) are captured using drain bias sensitivity of CDSC via CDSCD model parameter and DIBL model coefficient ETA0, respectively, followed by the extraction of parameters for the velocity saturation effect. The impact of DIBL and channel length modulation is incorporated by carefully selecting the model parameters to minimize the error between measured data and model in the transfer and output characteristics simultaneously.

State-of-the-art FinFETs exhibit significant self-heating (SH), which leads to severe performance degradation [36]. As transistors transfer and output characteristics are susceptible to temperature, during the model calibration, one must extract temperature dependence of V_{TH}, mobility, series resistance, DIBL, velocity saturation, channel length modulation, etc. To accurately estimate thermal resistance, R_{TH}, and thermal capacitance C_{TH} values for self-heating, the temperature dependence of the aforementioned physical parameters and effects need to be carefully calibrated based on the transistor measurements in the range of 10 K to 300 K. The R_{TH} and C_{TH} parameters were extracted following the similar approach we have used in [37] (fitting the g_{DS} vs. frequency, C_{DD} vs. frequency, C_{DG} vs. frequency simultaneously by carefully selecting the parameters R_{TH} and C_{TH}).

B. Parameter Extraction at Cryogenic Temperatures

Silicon-based field effect transistors show different characteristics at cryogenic temperatures from nominal temperature, viz. nonlinear sub-threshold slope due to the band-tail effect, enhanced Coulomb and Surface roughness scattering, and nonlinear saturation velocity [35]. We use the model equations developed by [35] to incorporate these effects in the simulation using the BSIM-CMG compact model. The impact of the band-tail and SS saturation is captured by first extracting the effective temperature and then carefully choosing the additional V_{TH} model parameters (KT_{11}, KT_{12}, T_{VTH}) to minimize the error between measured data and simulated model data. Reducing the temperature to cryogenic temperatures leads to lower thermal-induced lattice vibration and enhancement in peak mobility. However, this reduction in temperature is also responsible for the decrease in the thermal velocity of the charge carriers. The mobile charge carriers with lower thermal velocity suffer from increased surface roughness scattering and may lead to a reduction in effective mobility at an elevated vertical field. These effects have been captured in the simulations by optimizing the values of mobility temperature coefficients (model parameters EU1, UA2). Cryogenic temperatures dependent Coulomb scattering is captured using UD1 and UD2. The impact of cryogenic temperature on velocity saturation has been incorporated with the help of modified expressions of the existing velocity saturation model, i.e., modifying the expression of parameters related to effective drain-source voltage (MEXP, KSATIV), and velocity saturation (VSAT) [35]. Note that the minor discrepancy between the model and measurement in the low V_{GS} regime is not due to inaccurate modeling. The model

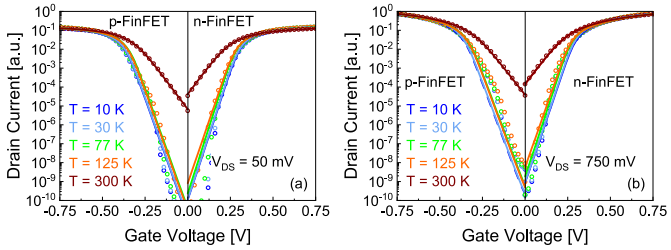


Fig. 4. Temperature dependent transfer characteristics of p- and n-FinFET at: (a) $V_{DS} = 50\,mV$ and (b) $V_{DS} = 750\,mV$.
Symbols: measured data, Line: Our calibrated model.

deviates from the data due to inherent randomness in the measurement (measurement results in Fig. 4 show multiple slopes with increasing V_{GS}), and discrepancies between measurement and model results can be attributed to noisy measurements at a lower current. The extracted model results show a good agreement with measured data, and simulation results from the extracted model for cryogenic temperatures, along with the measured data in Fig. 4, demonstrate the quality of the extracted model.

V. SRAM Memory: Noise Evaluation Methods, Array Structure, and Cell Types

SRAM is a crucial and frequently used circuit for semiconductor technology evaluation and as a volatile high-performance memory component (compared to DRAM, Flash, etc.). A typical six-transistor (6T) SRAM bit-cell consists of a cross-coupled inverter pair and two access transistors to isolate the stored data from the bit lines. The access transistor also provides access to the bit stored inside the cell during read and write operations (Fig. 5). Therefore, the SRAM has three operations: 1) Hold the stored data. 2) Read the stored data via the bit lines to a sense amplifier and a latch. 3) Use strong write drivers to force an external state into the bit cell as a write operation.

While the hold operation is simply doing nothing to maintain the status, the read and write operation involves peripheral circuitry on the column of the SRAM arrays. During the read operation, the bit line is pre-charged to V_{DD}, followed by activating the word line (WL) to access the stored data. If the data stored at the node (Q or QB) is '0', then it pulls the bit line connected at that node to ground (Fig. 5). For a write operation, data transfer over the bit line (BL) and complement bit line (BLB) is followed by the WL activation to access the storage cell.

A. Noise Margins

The maximum allowed noise voltage without disturbing the stored bit state is quantified as the Static Noise Margin (SNM). During the hold operation, the SRAM cell can resist a specific voltage noise before its internal state flips. This is called the "hold noise margin" (HNM). During the read operation, the SRAM cell must maintain its state (i.e., maintain voltage at Q) while current is withdrawn to discharge the bit lines. Hence the cell is most vulnerable during the read operation. Therefore, the "read noise margin" (RNM) is typically smaller than the write and hold noise margins. The minimum BL

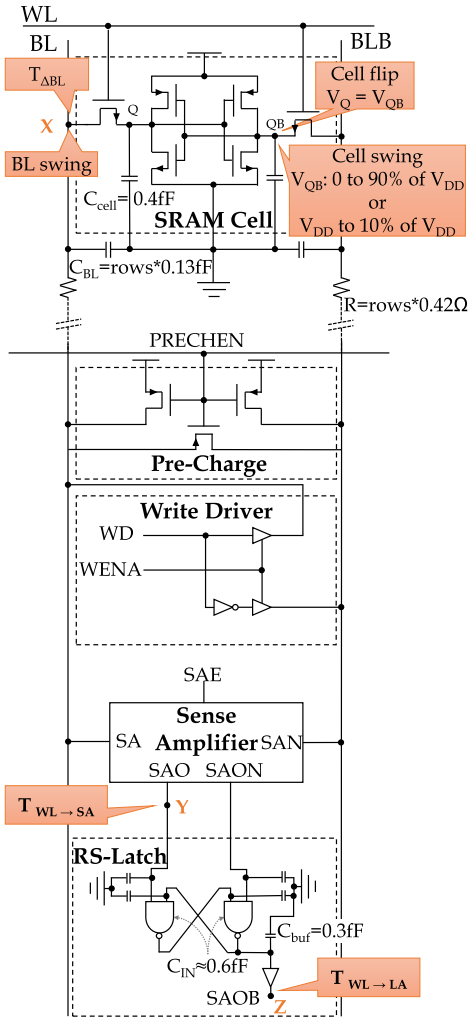


Fig. 5. Schematic view of an SRAM array.

voltage required to destroy the internal cell state is the "write noise margin" (WNM).

The hold, read and write noise margins are extracted using mathematical operations on the measured butterfly curve (BFC) (Fig. 6(a)) and cannot be measured directly. The N-curve of the cell offers an alternative to measure the read stability by the inline testers [38]. The N-curve is the metric which shows the relation between the current and voltage of the noise source injected at the Q or QB node. It also provides additional information about the resiliency against the current noise, which can not be obtained from BFC measuring purely voltage noise. The cell's N-curve is extracted by keeping it in the read mode; a noise voltage is then injected at the '0' storage node, simultaneously measuring the respective current flowing into that node.

B. SRAM Arrays: Memory, Pre-Charging, Sense Amplifiers, Latches and Write Drivers

A classical SRAM memory array contains millions of these bit-cells along with peripheral circuitry to enable the read-write operation on the memory array. In an SRAM array

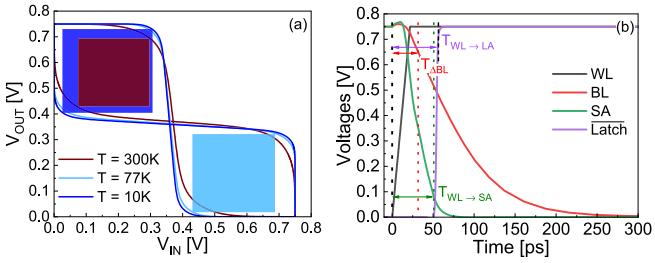


Fig. 6. (a) Impact of ambient temperature on the butterfly curve of High-Density Cell (HDC) in the hold state. (b) Simulated output waveform during read '0' case in HDC at 300 K.

(Fig. 5), each bit-cell is connected to a shared word line (WL) in each row and a complementary bit-line pair (BL, BLB) in each column. The number of bit cells in a memory array is enormous (64 Mbit SRAM cache consists of 67 108 864 bit-cells). As the bit-cells dominate overall area and power, a slight improvement in bit-cell results in an almost equal improvement in the memory array.

Typically, each column features peripheral circuits next to the SRAM cells (Fig. 5). In our SRAM array implementation, we employ 32 cells (rows) in a column with a single set of peripheral circuits per column. Since we do not have the exact numbers for parasitics of 5 nm technology and neglecting them would have an impact on power and energy results, we have approximated them as lumped resistance and capacitance for schematic-level simulations, as shown in Fig. 5 [39]. First is the pre-charge circuit (PC), which charges both bit lines (BL and BLB) to V_{DD} to support the read operation of an SRAM cell, as it can easily create a difference when both bit lines are on the same potential. The sense amplifier (SA) then amplifies this minuscule difference between BL and BLB as a differential amplifier. This amplified output is then stabilized in an output latch (LA). During writing, the data that needs to be written is applied on bit lines in the SRAM bit-cell. Strong write buffers are necessary to overpower the SRAM cell and potentially flip its internal state (if the opposite value to the stored value must be written). This is performed by a strong writer driver (WD).

C. SRAM Cell Types: HDC, LVC, and HPC

As explained before, any new CMOS technology capabilities (area, power, performance) are evaluated with SRAM cells. Three key SRAM cell types are used: HDC, LVC, and HPC. These cells are characterized by their fin numbers for pull-up (PU), pass-gate (PG), and pull-down (PD). With the introduction of FinFET technology, transistor sizing is not just a continuous width (W) and length (L) but instead quantized to a discrete number of fins (one fin, two fins (double the strength), three fins (triple driving current), etc.). The following paragraphs present the typical SRAM bit cell types used in the industry to support the trade-offs between cell area, performance, and power consumption.

The HDC of any technology must have minimum width per transistor, i.e., a single fin for PU, PG, and PD transistors. Although employing single fin FinFETs for PU, PG, and PD transistors (1:1:1) in an HDC gives minimum area and lowest leakage, employing solely weak transistors leads to performance issues [17], [27], [28].

A typical LVC in FinFET technologies has one fin for PU and PG to keep the area minimum and two fins in the PD transistor to improve read performance, enabling operation at low voltages (and thus reducing dynamic and static power) [17]. The increased strength of the PD transistor improves the read stability, especially at low voltages with marginal increment in leakage current and area. This makes the LVC a benchmark for the low-power capabilities of CMOS technology.

Similar strength of PU and PG transistors in HDC and LVC may result in poor read performance. SRAM bit-cells with a stronger PG help achieve very high performance (small read and write delay) at the cost of increased area and leakage. These HPC are a benchmark for technology speed [17], [28].

VI. SPICE Framework for SRAM Characterization

The SRAM array from the previous section is evaluated to assess the CMOS technology with SRAM benchmarking and as a memory element for quantum computers (as shown in Fig. 2). For this purpose, we employ circuit simulations using the SPICE circuit simulator family. We developed a SPICE framework to evaluate the delay, power, and noise margin of the different cell types (HDC, LVC, HPC) from room to cryogenic temperatures. Depending on the user inputs, the framework generates the power consumption, delay results, butterfly curves (hold, read, and write noise margins), and N-curves. The framework is designed to perform all the analyses in parallel to minimize the run time. The hold, read, and write margins are determined by the size of the largest possible square that can be accommodated between the two eyes of the butterfly curve. N-curve read/write margins are extracted from the peak and valley of the n-curves. Extraction of these noise margins from a graph (BFC, N-curve) by visualization is complex and error-prone. At the end of simulations, the framework generates the butterfly and n-curves and extracts the noise margin values by auto-tracing those curves.

For a detailed picture of the read and write operations, we report their propagation delays in Table II at several intermediate stages along the signal propagation path within the circuit (Fig. 6(b)). The time measurement for all read delays starts at the beginning of WL activation, and the SA enable signal triggers simultaneously with the WL signal. The $T_{\Delta BL}$ in Fig. 5 is the time SRAM cell takes to produce a 100 mV potential difference between BL and BLB during the read operation [40]. Further, along the path, delay $T_{WL \rightarrow SA}$ at the SA output (point 'Y' in the SRAM array, see Fig. 5) is measured between WL activation and when the respective output (SAO/SAON) crosses 10 % of V_{DD}. Finally, the delay $T_{WL \rightarrow LA}$ at the latch output terminal (point 'Z' in the SRAM array, see Fig. 5) is the time between WL activation to when LA output reaches 90 % of its target potential.

For the write delays, we differentiate between the cell flip, cell swing, and bit line swing delay. We consider the cell to be flipped when Q equals QB (Table II). The cell swing delay is measured until Q has reached 90 % of the target potential and is shown in line five of Table II. Similar to the read delays, the time measurement starts at the beginning of the WL activation.

TABLE II
DELAY COMPARISON OF SRAM CELL OPERATING AT 300 K, 77 K AND 10 K

Operation	T = 300 K Delay [ps]			T = 77 K Delay [ps]			T = 10 K Delay [ps]		
	HDC	LVC	HPC	HDC	LVC	HPC	HDC	LVC	HPC
Read ($T_{\Delta BL}$)	31.72	28.72	24.24	30.87	28.12	23.74	31.22	28.44	23.96
Read ($T_{WL \rightarrow SA}$)	51.50	51.27	50.21	46.47	46.30	45.42	46.46	46.25	45.23
Read ($T_{WL \rightarrow LA}$)	58.84	58.66	57.67	55.33	55.20	54.45	56.01	55.84	54.98
Write (Cell flip)	26.98	31.99	21.76	26.55	30.82	21.52	26.66	30.84	21.66
Write (Cell swing)	31.93	36.54	25.02	30.86	34.94	24.33	30.88	34.99	24.41
Write (BL swing)	31.88	31.88	33.56	30.91	30.91	32.36	31.25	31.25	32.69

Fig. 5 highlights the locations where various measurements are performed.

TABLE III
POWER COMPARISON BETWEEN SRAM CELLS OPERATING AT 300 K, 77 K AND 10 K IN 5 nm FINFET TECHNOLOGY

Average Power during Operation	T = 300 K Power [μW]			T = 77 K Power [μW]			T = 10 K Power [μW]		
	HDC	LVC	HPC	HDC	LVC	HPC	HDC	LVC	HPC
Read until SRAM charged BL	72.04	74.55	81.15	69.72	72.60	78.66	68.36	71.30	76.94
Read until signal passed SA	87.27	92.75	109.11	78.88	84.63	102.50	75.75	81.40	99.39
Read until signal passed LA	93.71	99.12	115.56	88.23	94.32	112.18	85.66	91.85	109.90
Read (just SRAM)	20.60	23.10	31.80	20.60	23.10	31.40	20.50	22.90	31.00
Read (just PC)	7.89	7.89	9.27	7.08	7.07	8.03	7.09	7.08	8.04
Read (just SA)	46.60	46.40	46.00	46.70	46.50	46.20	45.70	45.50	45.30
Read (just LA)	25.70	25.80	25.80	20.80	20.80	21.10	19.60	19.70	20.00
Write until Cell flip	103.43	101.78	145.88	102.68	101.67	145.92	106.74	104.98	150.30
Write until Cell swing	95.77	96.92	138.13	95.21	96.10	138.13	98.61	98.40	142.08
Write until BL swing	188.58	188.62	192.74	184.33	184.34	189.18	180.14	180.15	185.66
Write (just SRAM)	10.90	16.90	15.30	8.22	13.30	11.50	7.39	12.10	10.30
Write (just PC)	9.42	9.41	10.70	8.69	8.69	9.64	8.66	8.65	9.60
Write (just WD)	187.00	187.00	191.00	183.00	183.00	188.00	180.00	180.00	185.00
	Power [nW]			Power [nW]			Power [nW]		
Hold (All Cells)	1 220.00	1 480.00	1 740.00	0.14	0.38	0.51	0.24	0.23	0.38
Hold (Single Cell)	39.90	46.50	55.20	0.65	0.15	0.01	0.19	0.06	0.02
Hold (just PC)	8.64	8.65	8.15	18.50	18.50	18.00	18.30	18.30	17.80
Hold (just SA)	31.00	31.00	31.00	0.31	0.31	0.31	0.33	0.33	0.32
Hold (just LA)	323.00	323.00	323.00	0.10	0.08	0.08	0.00	0.01	0.01
Hold (just WD)	297.00	297.00	297.00	0.03	0.09	0.08	0.05	0.04	0.04

SA = Sense Amplifier, LA = Output Latch, PC = Pre-Charging Circuit, WD = Write Driver. If not clarified by bracket, entire array is measured.

The bit line swing measurement features a different start time, as the WD is enabled before the WL activation to prepare the bit lines in advance. Therefore the bit line swing delay starts at the beginning of the WENA activation and stops at the 90 % voltage target of the respective bit line (Table II).

VII. 300 K EVALUATION OF 5 nm FINFET SRAM

In this section, we evaluate the performance of the commercial 5 nm FinFET technology with SRAM cells at 300 K. This is mainly to benchmark the CMOS technology and as a baseline to compare our cryogenic results with others. Our focus on quantum reappears in the cryogenic evaluation later.

A. Power Consumption of SRAM Cells

The highest power consumption (138.132 μW in HPC) occurs in the write operation of the SRAM cell. It is considerably higher than hold and read for all studied cell types at 300 K, as shown in Table III. Comparing cell types, HDC types consume minimum power during the hold (0.039 μW), read (93.710 μW), and write (95.772 μW) operations.

We also present the overall energy consumption of SRAM in Table V. The HDC consumes the lowest energies at 300 K for the hold, read, and write operations.

B. Performance (Read, Write Delay) of SRAM Cells

Table II contains the delay values from a single cell to an array, as mentioned in Section I-C and Section V (positions shown in Fig. 5). The HPC results in the lowest delay (at 300 K) for bit-cell during a read operation. On the other hand, the latch masks some of these performance advantages, evening out the delay across the cell types. The HPC in Table II shows the lowest write delay for cell flip and cell swing compared to other cells. All studied SRAM cells built with 5 nm FinFETs perform read and write operations within a few ps and thus support multi-GHz operations.

C. Static Noise Margin Evaluation

SRAM possesses a certain noise margin to avoid data loss (to tolerate voltage noise, e.g., V_{DD} variations). Hence, we present a detailed analysis of the impact of V_{DD} on the SNM of different cell types.

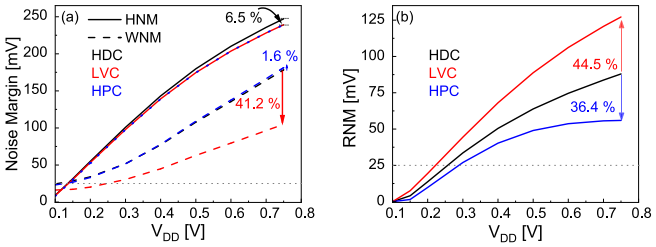


Fig. 7. Impact of different fin configurations and V_{DD} on noise margins at 300 K. (a) Hold and Write Noise margins variation with increasing V_{DD}. (b) Read Noise margin. The black dotted line in both figures shows the thermal noise present at 300 K.

1) Impact of Supply Voltage: For the HDC, we observe that noise margin degrades with V_{DD} reduction with different rates as we move towards lower V_{DD}. While HNM and RNM dropped by 59.2 % and 61.81 %, respectively, WNM dropped considerably more (79.97 %) in the same range (Fig. 7). The SNM of HDC being ≥50 mV for V_{DD} of 0.4 V shows that the studied SRAM cell can operate at such low V_{DD}.

2) Impact of Fin Configuration: Fin configurations significantly impact the SNM (see Section V). We observe two slopes in the HNM for all cell types with increasing V_{DD} and a sharp decrease in static performance at lower voltages compared to nominal V_{DD}, i.e., 0.75 V (Fig. 7(a)). The WNM of HDC and HPC degrade at a higher rate than LVC and shows a linear relationship with V_{DD} for ≥0.4 V. Due to the '0' storage node rising to a higher potential during the read operation, the cell becomes most vulnerable while reading. The LVC results in a ∼45 % improvement in the read noise margin but also degrades the write noise margin by 41 % at nominal V_{DD} (Fig. 7(a) and 7(b)). Simulation results also indicate that the HDC cell will be more resilient to the voltage noise during read if one operates at V_{DD} = 0.75 V.

D. Current Noise Resilience

SRAM cells which exhibit the same SNM value are solely equally resilient against voltage noise. Another source of noise is current noise; unfortunately, current and voltage noise do not correlate. The SNM extracted from the butterfly curve does not provide any information on noise margins against the current noise. Hence, in this section, we have analyzed the static voltage noise margin (SVNM) and current noise margin (SINM) of the 6T SRAM (extracted from N-Curve) for all the cell types and operating conditions discussed in Section V. The voltage difference between points A and B (point B and C) is defined as the read (write) SVNM. The SINM for read (write) is extracted from the positive (negative) peak current of the simulated n-curve.

1) Cell Type Impact on Noise Margin: Fig. 8(a) presents the N-curve of HDC, LVC, and HPC at 300 K. The difference among two zero crossings (points A and B) in Fig. 8(a) represents the maximum allowable voltage noise level (allowable voltage until bit flips) for cell stability in a read operation and between points B and C for a write operation. The SINM for a read operation is defined as the peak current between points A and B (Fig. 8(a)). Similarly, the peak value of the current between points B and C represents the write SINM. As mentioned before, only N-curve captures current noise resiliency (compared to voltage only in BFC). A comprehensive comparison of noise margins extracted from the N-curve of different cells in Table IV shows that LVC (1:1:2) is more resilient to noise than others.

TABLE IV
N-CURVE NOISE MARGINS AT 300 K

Cell	Read		Write	
	SVNM [mV]	SINM [μA]	SVNM [mV]	SINM [μA]
HDC	232.10	29.30	331.70	22.70
LVC	264.80	70.30	379.20	39.90
HPC	212.60	38.90	339.80	19.60

2) N-Curve Noise Resiliency Against Supply Voltage: Fig. 8(b) and 8(c) present the N-curve's voltage and current noise margins for a wide V_{DD} range. The HDC has the lowest read, write current and write voltage noise margins for V_{DD}≤500 mV. The HPC shows a marginal reduction in both read voltage and write current noise margins compared to the HDC and, at the same time ∼2.35 × improvement in the read current noise margins for V_{DD} ranging from 0.4 V to 0.75 V. The LVC has higher noise margins than the other two cells, and the read and write current noise margins of this cell decrease by ∼97 % of their maximum values at V_{DD}≤300 mV.

VIII. CRYOGENIC EVALUATION OF 5 nm FINFET SRAM

After baseline results at 300 K, we now explore 77 K (liquid nitrogen) and 10 K (near liquid helium). Results at 77 K are common and allow comparisons to other works, while 10 K is our target for a quantum computer interface. We explore how an unaltered commercial FinFET technology (optimized to operate at 300 K instead of 77 K to 10 K) would be suitable for a quantum interface. To work as a part of a quantum interface, SRAM read/write delay should be 0.66 ns or lower [9] to control qbits before their quantum state collapses. Additionally, SRAM array power must be a few mW/qubit to allow operation on a large number of qubits inside the dilution refrigerator with its limited cooling (see Section I-C). Lastly, noise margins should be comparable to 300 K, to prevent data corruption in the SRAM array at 77 K to 10 K.

A. SRAM's Power Consumption

The read-out and control circuitry of the quantum processors has a very tight power budget (a few 100 mW). SRAM is an integral part (as a memory element) of quantum circuits (see Fig. 2). However, SRAM might become too power-hungry for quantum circuits due to quantum computers' fast data transfer requirements during the short qubit control loops.

Table III compares the power consumed by an SRAM cell at 300 K, 77 K, and 10 K. The static/leakage power is the typical drawback of SRAM compared to other memory technologies. We have observed that hold power consumption (i.e., leakage power) significantly reduces (≥99 %) as we move from 300 K to 10 K. However, read-and-write power consumption must be reduced to operate under the power budgets of cryogenic chambers. Indeed, cryogenic temperatures reduce average read power consumption by 4.89 % to 8.59 %, at 10 K for different cell types. During the write operation, the write driver

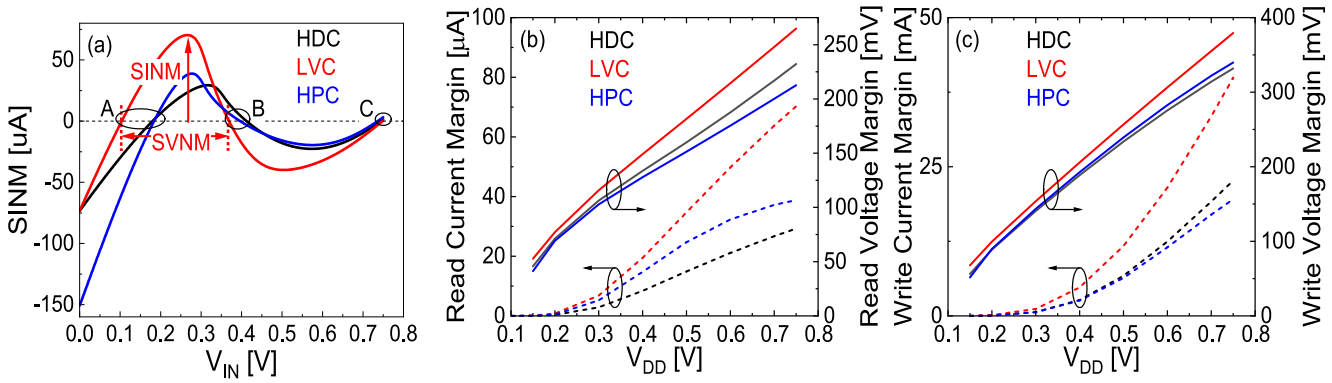


Fig. 8. Room temperature (a) N-curve, and (b) Read, (c) Write, Current and Voltage noise margins extracted from N-curve for HDC, LVC, and HPC cell at nominal V_{DD}.

TABLE V

ENERGY CONSUMPTION DURING READ AND WRITE OF SRAM CELL OPERATING AT 300 K, 77 K, AND 10 K

Energy Consumption for Operation	T = 300 K Energy Consumption [fJ]			T = 77 K Energy Consumption [fJ]			T = 10 K Energy Consumption [fJ]		
	HDC	LVC	HPC	HDC	LVC	HPC	HDC	LVC	HPC
Read until SRAM charged BL	2.284	2.140	1.964	2.154	2.040	1.864	2.133	2.025	1.847
Read until signal passed SA	4.495	4.758	5.477	3.668	3.918	4.654	3.523	3.769	4.492
Read until signal passed LA	5.510	5.818	6.668	4.879	5.207	6.114	4.797	5.125	6.044
Read (just SRAM)	0.654	0.664	0.771	0.636	0.650	0.745	0.638	0.651	0.743
Read (just SA)	2.402	2.381	2.310	2.169	2.152	2.100	2.122	2.105	2.051
Read (just LA)	1.514	1.511	1.491	1.149	1.150	1.147	1.100	1.102	1.098
Read (just PC)	3.946	3.947	4.634	3.540	3.536	4.017	3.546	3.542	4.018
Write until Cell flip	2.793	3.257	3.180	2.731	3.132	3.137	2.850	3.233	3.262
Write until Cell Swing	3.055	3.538	3.453	2.942	3.354	3.357	3.047	3.444	3.467
Write until BL swing	6.016	6.017	6.476	5.696	5.696	6.129	5.638	5.639	6.071
Write (just SRAM)	0.349	0.617	0.382	0.254	0.464	0.279	0.228	0.425	0.250
Write (just PC)	4.708	4.707	5.347	4.347	4.343	4.820	4.328	4.323	4.798
Write (just WD)	5.973	5.973	6.399	5.671	5.671	6.090	5.612	5.612	6.031
	Energy Consumption [aJ]			Energy Consumption [aJ]			Energy Consumption [aJ]		
Hold (All Cells)	291.921	354.215	416.179	0.033	0.091	0.122	0.058	0.054	0.090
Hold (Single Cell)	9.542	11.115	13.189	0.155	0.035	0.002	0.046	0.014	0.006
Hold (just PC)	4.338	4.342	4.092	9.288	9.287	9.027	9.199	9.199	8.944
Hold (just SA)	7.415	7.413	7.415	0.073	0.075	0.073	0.079	0.079	0.077
Hold (just LA)	77.310	77.308	77.308	0.024	0.019	0.020	0.001	0.002	0.002
Hold (Just WD)	71.095	71.104	71.111	0.006	0.022	0.020	0.013	0.009	0.011

SA = Sense Amplifier, LA = Output Latch, PC = Pre-Charging Circuit, WD = Write Driver. If not clarified by bracket, entire array is measured.

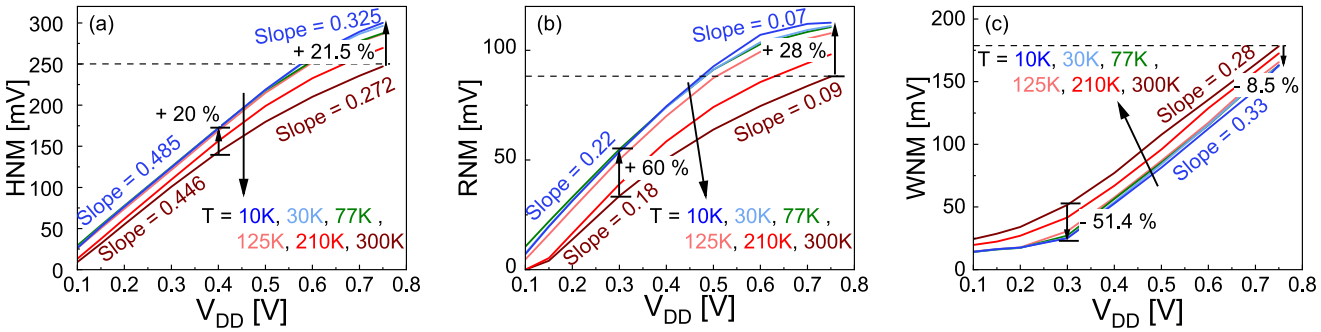


Fig. 9. Impact of ambient temperature on (a) Hold, (b) Read and (c) Write noise margins of HDC are shown for multiple V_{DD} for varying temperatures. Hold, read, and write noise margin show saturation below 77 K for all V_{DD}. The horizontal dashed line shows the maximum noise margin of HDC at 300 K.

masks the benefits from the cell and results in similar power consumption across temperatures. As leakage is negligible, additional dynamic power reduction of 4.8 % to 8.5 % during read makes 5 nm technology a suitable candidate for SRAM memory design for quantum computing applications.

Table V presents the total energy consumption of the SRAM array for different cells at room and cryogenic temperatures. At 10 K we observe 99.99 % less energy consumption for holding data. For dynamic operations such as read, energy consumption decreases by ∼9 % to ∼13 % at 10 K.

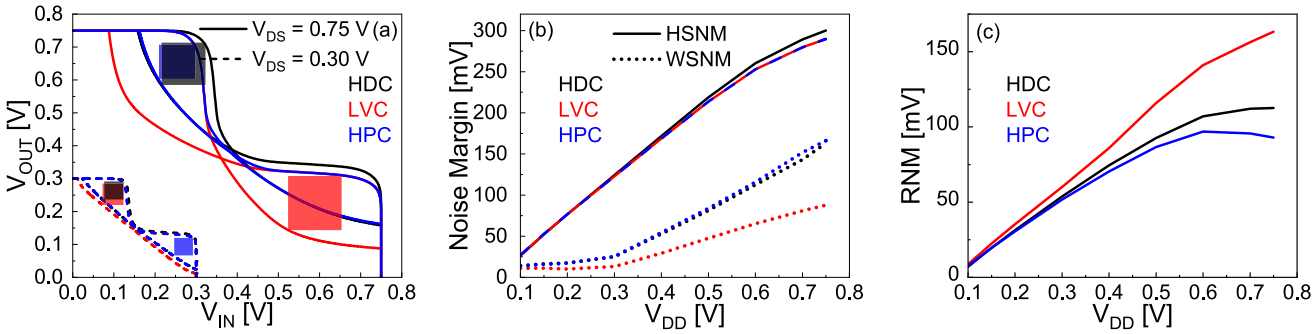


Fig. 10. Impact of different fin configurations and supply voltage on noise margins at cryogenic temperature (T = 10 K): (a) Read butterfly curve at nominal and low V_{DD} with varying cell types. (b) Hold and Write Noise margin variation with increasing V_{DD}. (c) Read Noise margin.

B. Cryogenic Performance (Delay)

To assess the speed of SRAM at cryogenic temperatures, Table II presents the effect of temperatures on SRAM cell delay for different operations and resulting performance enhancements of HDC, LVC, and HPC. At 77 K, SRAM bit-cell shows only ∼2 % to ∼3 % and ∼2 % to ∼5 % performance improvement for read and write operations, respectively, at nominal V_{DD}. Unfortunately, further lowering of temperature provides no further benefit. Note that this is in stark contrast to previously reported results [32], where a 30 % and 28 % reduction in read and write time is observed at 77 K compared to 300 K. Exploring cell types, the HPC shows the lowest latency compared to HDC and LVC. While power improved drastically with ≥99 % in static power and ∼5 % to 9 % in dynamic power, delay barely improved. However, as 5 nm SRAM is already very fast at 300 K (ps-regime) and maintains this high performance at 10 K, this 5 nm FinFET SRAM is still suited for quantum circuits.

C. Static Noise Margin Evaluation

Now let's explore the resiliency of 5 nm FinFET SRAM against noise in cryogenic conditions.

1) Impact of Cryogenic Temperatures: An enhanced transistor performance (lower leakage, higher mobility, and stronger drive current) at cryogenic temperatures should result in stronger noise margins. Fig. 9 shows that cryogenic temperatures affect the resiliency of the SRAM in two primary ways: 1) It improves the hold and read noise margins by ≥21 % and 28 %, respectively, with only an 8.5 % decrease in write noise margin at nominal V_{DD}. 2) Cryogenic temperatures (≤77 K) enable SRAM to operate at ∼250 mV lower V_{DD} while maintaining hold and read noise margins compared to 300 K. Comparing noise margins at 10 K to 77 K across V_{DD} shows that this temperature range exhibits near identical margins across a wide range of V_{DD}. In summary, higher noise margins observed at cryogenic temperatures highlight the SRAM noise tolerance and enable lowering the V_{DD} to achieve comparable noise margins as regular operation (300 K) at nominal V_{DD}. Therefore, solely performance limits how much below nominal V_{DD} the SRAM can be operated, as noise margins are not a concern. This potentially further improves the power efficiency of 5 nm SRAM in quantum computers.

2) Impact of Supply Voltage: As the previous section motivates lower V_{DD}, this section explores that in more detail. Fig. 9 to 11, present how reducing V_{DD} expectedly degrades the SNM and thus increases the chance of data corruption. Similar to operating at 300 K, at colder temperatures, too V_{DD} governs SNM. Lowering V_{DD} from 0.75 V to 0.5 V or further to 0.3 V results in 27 % and ∼59 % degradation in HNM at 10 K, respectively. Fig. 9(a) and 9(b) show that SNM decreases at two different rates with V_{DD} reduction. Fig. 9(a) shows that the RNM degrades by only 5 % for V_{DD} = 0.75 V to 0.6 V and 50 % for V_{DD} = 0.6 V to 0.3 V. On the other hand, the WNM in Fig. 9(c) depicts a constant slope with V_{DD}, degrading the WNM by ∼87 % for a V_{DD} reduction from 0.75 V to 0.3 V. At 10 K, SNM always remains >3 × of ambient thermal noise for V_{DD} >0.3 V. This highlights the 5 nm SRAM as a suitable candidate for quantum circuits.

3) Impact of Cell Types (Transistor Fins): Circuit designers use different 6T SRAM cell types (HDC, LVC, HPC) in memory design based on the application (see Section I-C and Section V). To evaluate their cryogenic performance, Fig. 10 presents the impact of different cell types on noise margins of SRAM at 10 K. Fig. 10(b) shows that the hold noise margin remains almost invariant to fin configurations. The HDC and HPC cells show near identical read and write noise margins at lower V_{DD} (Fig. 10(b) and 10(c)). HPC offers a nearly constant read noise margin for V_{DD} ≥525 mV, which can help reduce power consumption by operating at a lower V_{DD} for a similar read noise margin. The LVC shows the highest RNM (163 mV) and lowest WNM (87 mV) among all the studied cells at nominal V_{DD}. Noise margins of all cells remain higher than the ambient thermal noise for all V_{DD}. The HDC and HPC seem to be optimal solutions for read and write noise margins among all studied cells.

D. Impact of Cryogenic Temperatures on N-Curve

We observe that read and write SVNM of HDC remain almost invariant (≤2 % variation) in the studied temperature range (Fig. 11(a) and 11(d)); however, read SINM shows ∼1.5 × increment (at 10 K) for HDC at V_{DD} = 0.75 V (Fig. 11(b)). The write SINM shows an increase at nominal V_{DD} up to 77 K only and deteriorates with further reduction in temperature (Fig. 11(c)). The p-FinFETs' threshold voltage below 77 K is so high that the amount of I_{ON} decreases with further reduction in temperature. This leads to the degradation of the write SINM while moving from 77 K to 10 K.

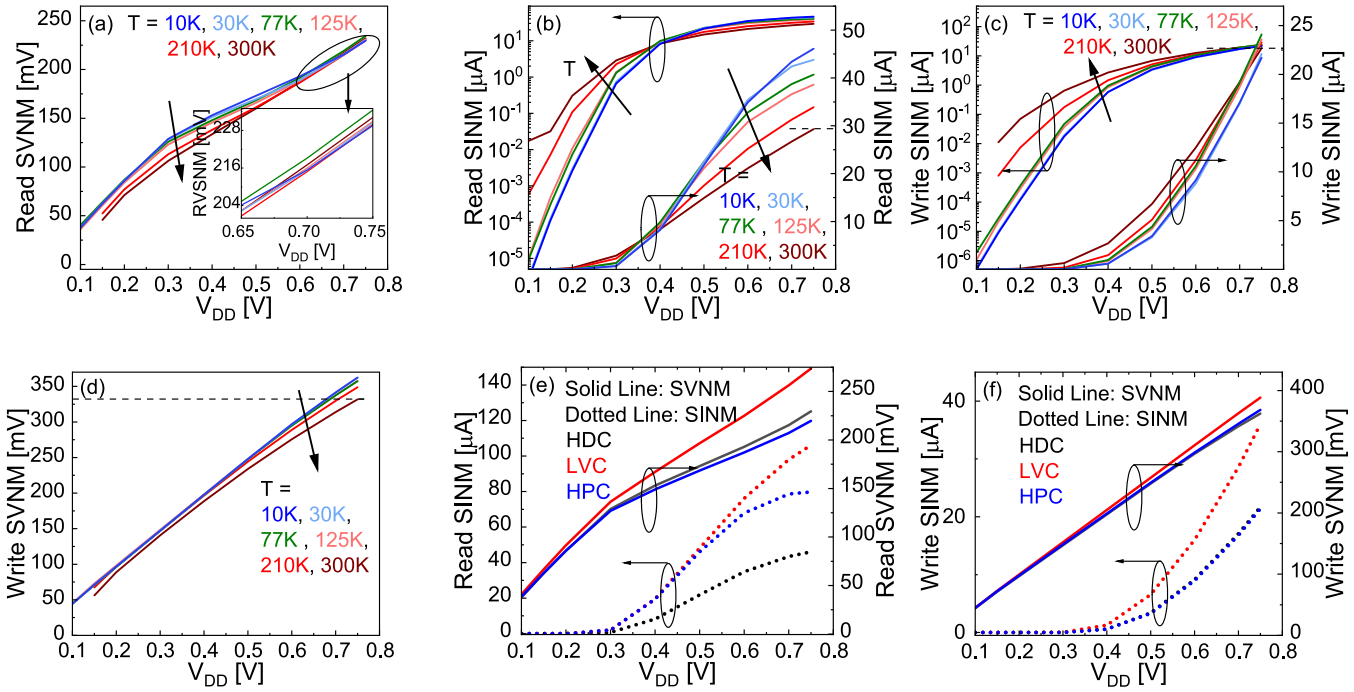


Fig. 11. (a)-(d) Cryogenic temperature influence on HDC read, write voltage and current noise margins extracted from N-curve. (e)-(f) N-curve's Read, Write voltage and current noise margin at $T = 10$ K for read and write operations of different cell types. Horizontal dashed line in (b), (c) and (d) shows the maximum noise margin of HDC at 300 K.

A similar SVNM and decreased SINM show that the SRAM cell will be less resilient at 10 K as the total noise charge needed to flip its state will be lower. Therefore, studying solely at 77 K as other works have done [19], [41], does not capture the full picture for CMOS in quantum processors.

1) Impact of Cell Types: Fig. 11(e) and 11(f) present the impact of cryogenic temperature on different cell types. The LVC in Fig. 11(e) and 11(f) shows 2.3 × and 1.6 × increments in the read and write SINM compared to the HDC. Write SINM, read and write SVNM for HDC and HPC remain almost the same, while HPC shows a significant improvement (1.73 ×) in the read SINM. The LVC offers enhanced SVNM and SINM, followed by HPC at 10 K.

2) Impact of Supply Voltage: Read SVNM and SINM in Fig. 11 show two different trends at lower and higher V_{DD} with a decrease in temperature. Fig. 11(a) shows the highest read SVNM for HDC at 10 K in low V_{DD} regime; however, at a higher V_{DD}, this trend reverses itself around 77 K and starts decreasing with further reduction in temperature. At 10 K, read SVNM reduces with a slope of 0.21 and 0.24, respectively, between 0.3 V to 0.6 V and 0.6 V to 0.75 V. Read and write SINM of HDC degrades by 98 % and 99 % while operating at 0.3 V compared to the nominal V_{DD}.

To summarise, cryogenic temperatures also improve voltage and current noise margins for ≥0.4 V. An unmodified commercial 5 nm technology analysis shows that it is well suited for low-power quantum processing applications.

IX. Insights of Cryogenic SRAM Array Design

A. Cryogenic Circuit Design

Our exploration of the SRAM array in cryogenic conditions highlights the following circuit design challenges and options. For the write operation, performance is mainly governed by the write driver and not by the cell types. This indicates that the write operation should be optimized to reduce its power consumption, e.g., by employing write assist [18], [34].

For the hold operation, there is little to optimize, as operation in cryogenic temperature reduces the transistor I_{OFF} significantly. Our observed ∼1000 × decrease in the power dissipation of all the cells during the hold operation (i.e., leakage) means the typical drawback of SRAM cells is eliminated. As cells are not the main concern at 10 K, contrary to operation at room temperature, a strong focus should be given to the periphery optimization.

B. Optimized Transistor Sizing in SRAM Cells

Once the periphery is optimized, the impact of the cells on power and delay becomes visible again and should be optimized via transistor sizing (i.e., determining fin count). The hold operation enforces no change in transistor sizing, as leakage currents are too low at 10 K. For the write operation, the power dissipation for the cell swing operation will be higher in a cell with a higher number of fins in PG and PD transistors (i.e., higher power dissipation in HPC than LVC and HDC). In the HPC, since the current driving capability of PG transistors are higher than LVC and HDC, they will help in faster switching (less delay). They may also result in a lower power dissipation due to the shorter delay while looking at the cells' only power dissipation. The read operation can also be optimized by transistor sizing. Read speed depends on the amount of current flowing from PG and PD. Hence, the cell with a higher number of fins in the PG and PD transistor (i.e., LVC and HPC) shows a smaller read delay (Table II).

A higher slope of the V_{out} vs. V_{in} characteristics in the butterfly curve increases the SNM. Both temperature reduction and increase in PD to PG "width/length" ratio results in a larger slope raising the SNM for hold and read cases, as shown in Fig. 6, Fig. 9, and Fig. 10. Also, as in the existing 5 nm technology, transistors do not have a significant increase in the current at cryogenic temperatures. Both the read and write delay of a cell at all temperatures show only a minor variation.

Since lower latency is one of the primary concerns during quantum operations, our overall suggestion for optimal transistor sizing is to choose HPC. Its power consumption is low enough to satisfy the cooling power budget constraints. Hence, it is the best trade-off between power and the delay at 10 K.

X. Conclusion

In this work, we have evaluated a commercial 5 nm FinFET technology as a candidate for cryogenic CMOS circuits in the interface layer of quantum processors. For this purpose, we have studied (via SRAM cells as our benchmark circuit) its delay, power, and energy at 300 K and two key cryogenic temperatures. First at 77 K (liquid nitrogen) and then at 10 K for temperatures suitable for a quantum interface layer. Our results highlight that leakage is essentially negligible at 77 K and 10 K, while delay actually shortens (improves) and power consumption is also reduced. Therefore studied commercial 5 nm FinFET technology is suitable for operation in the interface layer of quantum processors, as it offers ultra-low power capabilities at high performance.

Acknowledgment

The authors would like to thank Central Research Facility IIT Delhi for facilitating the cryogenic characterization of FinFETs.

References

[1] F. Sebastiano et al., "Cryo-CMOS interfaces for large-scale quantum computers," in *IEDM Tech. Dig.*, Dec. 2020, p. 25.

[2] F. Sebastiano et al., "Cryogenic CMOS interfaces for quantum devices," in *Proc. 7th IEEE Int. Workshop Adv. Sensors Interfaces (IWASI)*, Jun. 2017, pp. 59–62.

[3] S. J. Pauka et al., "A cryogenic CMOS chip for generating control signals for multiple qubits," *Nature Electron.*, vol. 4, no. 1, pp. 64–70, Jan. 2021, doi: 10.1038/s41928-020-00528-y.

[4] R. Maurand, "A CMOS silicon spin qubit," *Nature Commun.*, vol. 7, p. 13575, Nov. 2016.

[5] B. Patra et al., "Cryo-CMOS circuits and systems for quantum computing applications," *IEEE J. Solid-State Circuits*, vol. 53, no. 1, pp. 309–321, Jan. 2018.

[6] J. C. Bardin et al., "A 28 nm bulk-CMOS 4-to-8 GHz 2 mW cryogenic pulse modulator for scalable quantum computing," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2019, pp. 456–458.

[7] L. Geck, A. Kruth, H. Bluhm, S. V. Waasen, and S. Heinen, "Control electronics for semiconductor spin qubits," *Quantum Sci. Technol.*, vol. 5, no. 1, Dec. 2019, Art. no. 015004, doi: 10.1088/2058-9565/ab5e07.

[8] L. M. K. Vandersypen et al., "Interfacing spin qubits in quantum dots and donors—Hot, dense, and coherent," *npj Quantum Inf.*, vol. 3, no. 1, p. 34, Sep. 2017, doi: 10.1038/s41534-017-0038-y.

[9] J. P. G. Van Dijk et al., "A scalable cryo-CMOS controller for the wideband frequency-multiplexed control of spin qubits and transmons," *IEEE J. Solid-State Circuits*, vol. 55, no. 11, pp. 2930–2946, Nov. 2020.

[10] N. C. Jones et al., "Layered architecture for quantum computing," *Phys. Rev. X*, vol. 2, no. 3, Jul. 2012, doi: 10.1103/PhysRevX.2.031007.

[11] L. L. Guevel et al., "A 110 mK 295 μW 28 nm FDSOI CMOS quantum integrated circuit with a 2.8 GHz excitation and nA current sensing of an on-chip double quantum dot," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2020, pp. 306–308.

[12] B. Dierickx, L. Warmerdam, E. R. Simoen, J. Vermeiren, and C. Claeys, "Model for hysteresis and kink behavior of MOS transistors operating at 4.2 K," *IEEE Trans. Electron Devices*, vol. 35, no. 7, pp. 1120–1125, Jul. 1988.

[13] P. A. T. Hart, M. Babaie, E. Charbon, A. Vladimirescu, and F. Sebastiano, "Subthreshold mismatch in nanometer CMOS at cryogenic temperatures," in *Proc. 49th Eur. Solid-State Device Res. Conf. (ESSDERC)*, Sep. 2019, pp. 98–101.

[14] H. Homulle, "Cryogenic electronics for the read-out of quantum processors," Ph.D. dissertation, Dept. Quantum Comput. Eng., Delft Univ. Technol., Delft, The Netherlands, 2019.

[15] T. Tsai, H. Lin, and P. Li, "Temperature-dependent narrow width effects of 28-nm CMOS transistors for cold electronics," *IEEE J. Electron Devices Soc.*, vol. 10, pp. 289–296, 2022.

[16] C. Auth et al., "A 22 nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors," in *Proc. Symp. VLSI Technol. (VLSIT)*, Jun. 2012, pp. 131–132.

[17] C.-H. Jan et al., "A 22 nm SoC platform technology featuring 3-D tri-gate and high-K/metal gate, optimized for ultra low power, high performance and high density SoC applications," in *IEDM Tech. Dig.*, Dec. 2012, pp. 1–4.

[18] E. Karl et al., "A 4.6 GHz 162 mb SRAM design in 22 nm tri-gate CMOS technology with integrated read and write assist circuitry," *IEEE J. Solid-State Circuits*, vol. 48, no. 1, pp. 150–158, Jan. 2013.

[19] H. L. Chiang et al., "Cold CMOS as a power-performance-reliability booster for advanced FinFETs," in *Proc. IEEE Symp. VLSI Technol.*, Jun. 2020, pp. 1–2.

[20] G. Yeap et al., "5 nm CMOS production technology platform featuring full-fledged EUV, and high mobility channel FinFETs with densest 0.021 μm^2 SRAM cells for mobile SoC and high performance computing applications," in *IEDM Tech. Dig.*, Dec. 2019, p. 36.

[21] A. Chabane et al., "Cryogenic characterization and modeling of 14 nm bulk FinFET technology," in *Proc. IEEE 51st Eur. Solid-State Device Res. Conf. (ESSDERC)*, Sep. 2021, pp. 67–70.

[22] E. Charbon et al., "Cryo-CMOS for quantum computing," in *IEDM Tech. Dig.*, Dec. 2016, p. 13.

[23] S. Pellerano et al., "Cryogenic CMOS for qubit control and readout," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Apr. 2022, pp. 1–8.

[24] M. Prathapan et al., "A cryogenic SRAM based arbitrary waveform generator in 14 nm for spin qubit control," in *Proc. IEEE 48th Eur. Solid State Circuits Conf. (ESSCIRC)*, Sep. 2022, pp. 57–60.

[25] S. Wu et al., "An enhanced 16 nm CMOS technology featuring 2nd generation FinFET transistors and advanced Cu/low-k interconnect for low power and high performance applications," in *IEDM Tech. Dig.*, Dec. 2014, pp. 1–4.

[26] S. Natarajan et al., "A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 μm^2 SRAM cell size," in *IEDM Tech. Dig.*, Dec. 2014, pp. 1–3.

[27] E. Karl et al., "A 0.6 V, 1.5 GHz 84 mb SRAM in 14 nm FinFET CMOS technology with capacitive charge-sharing write assist circuitry," *IEEE J. Solid-State Circuits*, vol. 51, no. 1, pp. 222–229, Jan. 2016.

[28] T. Song et al., "13.2 A 14 nm FinFET 128 Mb 6T SRAM with VMIN-enhancement techniques for low-power applications," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2014, pp. 232–233.

[29] J. Park et al., "A fully integrated cryo-CMOS SoC for state manipulation, readout, and high-speed gate pulsing of spin qubits," *IEEE J. Solid-State Circuits*, vol. 56, no. 11, pp. 3289–3306, Nov. 2021.

[30] K. Kuwabara, H. Jin, Y. Yamanashi, and N. Yoshikawa, "Design and implementation of 64-kb CMOS static RAMs for Josephson-CMOS hybrid memories," *IEEE Trans. Appl. Supercond.*, vol. 23, no. 3, Jun. 2013, Art. no. 1700704.

[31] J. van Staveren et al., "Voltage references for the ultra-wide temperature range from 4.2 K to 300 K in 40-nm CMOS," in *Proc. IEEE 45th Eur. Solid State Circuits Conf. (ESSCIRC)*, Sep. 2019, pp. 37–40.

[32] V. P. Hu et al., "High-density and high-speed 4T FinFET SRAM for cryogenic computing," in *IEDM Tech. Dig.*, Dec. 2021, p. 8.

[33] T. Huynh-Bao, S. Sakhare, J. Ryckaert, A. Spessot, D. Verkest, and A. Mocuta, "SRAM designs for 5nm node and beyond: Opportunities and challenges," in *Proc. IEEE Int. Conf. IC Design Technol. (ICICDT)*, May 2017, pp. 1–4.

[34] J. Chang et al., "A 5 nm 135 Mb SRAM in EUV and high-mobility-channel FinFET technology with metal coupling and charge-sharing write-assist circuitry schemes for high-density and low-VMIN applications," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2020, pp. 238–240.

[35] G. Pahwa, P. Kushwaha, A. Dasgupta, S. Salahuddin, and C. Hu, "Compact modeling of temperature effects in FDSOI and FinFET devices down to cryogenic temperatures," *IEEE Trans. Electron Devices*, vol. 68, no. 9, pp. 4223–4230, Sep. 2021.

[36] C. Prasad, "A review of self-heating effects in advanced CMOS technologies," *IEEE Trans. Electron Devices*, vol. 66, no. 11, pp. 4546–4555, Nov. 2019.

[37] S. S. Parihar, J. Z. Huang, W. Wang, K. Imura, and Y. S. Chauhan, "Self-heating characterization and modeling of 5 nm technology node FinFETs," in *Proc. Device Res. Conf. (DRC)*, Jun. 2022, pp. 1–2.

[38] C. Wann et al., "SRAM cell design for stability methodology," in *Proc. IEEE VLSI-TSA Int. Symp. VLSI Technol.*, Apr. 2005, pp. 21–22.

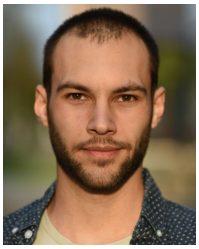
[39] H. Amrouch et al., "Reliability challenges with self-heating and aging in FinFET technology," in *Proc. IEEE 25th Int. Symp. On-Line Test. Robust Syst. Design (IOLTS)*, Jul. 2019, pp. 68–71.

[40] K. Chul Chun, P. Jain, T.-H. Kim, and C. H. Kim, "A 667 MHz logic-compatible embedded DRAM featuring an asymmetric 2T gain cell for high speed on-die caches," *IEEE J. Solid-State Circuits*, vol. 47, no. 2, pp. 547–559, Feb. 2012.

[41] E. R. Hsieh, Z. Y. Wang, Y. S. Huang, T. C. Hung, R. Y. Lyu, and K. Y. Lee, "Cryogenic quasi-ballistic transport enhanced by strained silicon technologies in 14-nm complementary fin field effect transistors through virtual source model," *IEEE Trans. Electron Devices*, vol. 69, no. 7, pp. 3575–3580, Jul. 2022.

Shivendra Singh Parihar (Member, IEEE) is currently pursuing the Ph.D. degree with the Indian Institute of Technology Kanpur, Uttar Pradesh, India. He is associated with the Semiconductor Test and Reliability (STAR) research group, University of Stuttgart, Germany, as a Research Scholar. His research interests include the characterization and compact modeling of advanced CMOS technologies for circuit designs.

Victor M. van Santen (Member, IEEE) received the M.Sc. degree in computer science and the Ph.D. degree from the Karlsruhe Institute of Technology (KIT) in 2014 and 2022, respectively. He is currently a Research Group Leader with the Chair for Semiconductor Test and Reliability (STAR), University of Stuttgart. His research interests include microprocessor reliability estimation, circuit and transistor simulation, failure analysis, and reliability estimation, with a special interest in aging phenomena.



Simon Thomann (Member, IEEE) received the bachelor's and master's degrees in computer science from the Karlsruhe Institute of Technology (KIT), Germany, in 2019 and 2022, respectively. He is currently pursuing the Ph.D. degree with the Chair of Semiconductor Test and Reliability (STAR), University of Stuttgart. His research interests include device to system level, circuit designs, emerging technologies, and cross-layer reliability modeling from device to circuit level.

Girish Pahwa (Member, IEEE) received the M.Tech. and Ph.D. degrees in electrical engineering from the Indian Institute of Technology (IIT) Kanpur in 2020. He is currently an Assistant Professional Researcher with the Department of Electrical Engineering and Computer Sciences (EECS), University of California (UC) at Berkeley, where he is also the Executive Director of the Berkeley Device Modeling Center (BDMC). Prior to this, he worked as a Post-Doctoral Researcher with the EECS, UC Berkeley, from 2020 to 2021. He is the co-developer of industry-standard BSIM-CMG, BSIM-IMG, BSIM-BULK, BSIM-SOI, and BSIM4 models. He has also developed the first industry-standard models for cryogenic FinFET and FDSOI FETs for quantum computing and cold electronics applications. He has more than 50 technical publications in prominent journals and conferences in the field of device modeling and simulation. His research interests include the modeling and simulation of nanoscale devices and device circuit co-design and optimization of emerging transistor technologies with a special emphasis on ferroelectric devices. He was a recipient of the IEEE EDS Early Career Award in 2022, the Outstanding Ph.D. Thesis Award from IIT Kanpur in 2020, and the Best Paper Award at the IEEE International Conference on Emerging Electronics (ICEE), Mumbai, India, in 2016. He serves as the reviewer of several reputed journals.

Yogesh Singh Chauhan (Fellow, IEEE) was with the Semiconductor Research and Development Center, IBM Bengaluru, from 2007 to 2010; the Tokyo Institute of Technology in 2010; the University of California at Berkeley from 2010 to 2012; and ST Microelectronics from 2003 to 2004. He is currently a Class of 1984 Chair Professor with the Indian Institute of Technology Kanpur, India. He is the developer of several industry standard models, such as ASM-GaN-HEMT model, BSIM-BULK (formerly BSIM6), BSIM-CMG, BSIM-IMG, BSIM4, and BSIM-SOI models. His research group is involved in developing compact models for GaN transistors, FinFET, nanosheet/gate-all-around FETs, FDSOI transistors, negative capacitance FETs, and 2D FETs. He has published more than 300 papers in international journals and conferences. His research interests include characterization, modeling, and the simulation of semiconductor devices. He is a fellow of the Indian National Academy of Engineering. He received Ramanujan Fellowship in 2012, IBM Faculty Award in 2013, P. K. Kelkar Fellowship in 2015, CNR Rao Faculty Award, Humboldt Fellowship, and Swarnajayanti Fellowship in 2018. He is the Chair of the IEEE-EDS Compact Modeling Committee. He is the Founding Chairperson of IEEE Electron Devices Society U.P. Chapter and the Elected Chairperson of IEEE U.P. Section. He has served in the technical program committees of IEEE IEDM, IEEE SISPAD, IEEE ESSDERC, IEEE EDTM, and IEEE VLSID. He is an Editor of IEEE TRANSACTIONS ON ELECTRON DEVICES and a Distinguished Lecturer of the IEEE Electron Devices Society.

Hussam Amrouch (Member, IEEE) received the Ph.D. degree (summa cum laude) from the Karlsruhe Institute of Technology (KIT) in 2015. He is currently a Professor heading the Chair of AI Processor Design, Technical University of Munich (TUM), and the Munich Institute of Robotics and Machine Intelligence (MIRMI), Germany. He is also the Head of the Semiconductor Test and Reliability, University of Stuttgart, Germany. Prior to that, he was a Research Group Leader with the KIT, where he was leading the research efforts in building dependable embedded systems. He has more than 210 publications in multidisciplinary research areas (including 85 journals) across the entire computing stack, starting from semiconductor physics to circuit design all the way up to computer-aided design, and computer architecture. His research in HW security and reliability have been funded by the German Research Foundation (DFG), Advantest Corporation, and the U.S. Office of Naval Research (ONR). His research interests include design for reliability and testing from device physics to systems, machine learning for CAD, HW security, approximate computing, and emerging technologies with a special focus on ferroelectric devices. He holds eight HiPEAC paper awards and three best paper nominations at top EDA conferences: DAC'16, DAC'17, and DATE'17, for his work on reliability. He has served in the technical program committees of many major EDA conferences, such as DAC, ASP-DAC, and ICCAD. He serves as a Reviewer for many top journals, like *Nature Electronics*, IEEE TRANSACTIONS ON ELECTRON DEVICES, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS, IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS, and IEEE TRANSACTIONS ON COMPUTERS. He currently serves as an Editor for the *Nature Scientific Reports* journal.