71st International Electron Devices Meeting

2025 IEEE International Electron Devices Meeting

100 YEARS of FETs:
SHAPING the FUTURE of DEVICE INNOVATIONS

2025 Technical Program



December 6 - 10, 2025
Hilton San Francisco Union Square
San Francisco, California
www.ieee-iedm.org

IEEE
Electron Devices Society

As of November 26, 2025

A DISTINGUISHED WELCOME TO IEDM 2025

On behalf of the IEEE International Electron Devices Meeting (IEDM) Committee, it is my distinct honor to welcome you to IEDM 2025, taking place in the city of San Francisco, December 6 - 10, 2025. For over seven decades, IEDM has stood as the world's foremost stage for unveiling transformative advances in semiconductor technology and device innovation. This year, we proudly continue that legacy with a program that reflects the pinnacle of scientific achievement and visionary thought leadership as well as celebrating the 100 years of FET. Our technical sessions will feature groundbreaking contributions from global industry pioneers, leading academic institutions, and the brightest emerging minds. From paradigm-shifting research to forward-looking perspectives, IEDM 2025 promises to be an unparalleled forum for knowledge exchange and strategic insight. I invite you to preview the program and paper summaries at www.ieee-iedm.org and access the electronic digest post-conference via IEEE Xplore. For those unable to attend in person, we encourage registration for IEDM OnDemand, which will include access to Tutorials and Short Courses. This option will be available on the website prior to the conference.

The conference starts with our renowned tutorial program, now in its 15th year of excellence. These sessions are designed for students, industry professionals, and anyone eager to bridge the gap between foundational knowledge and the forefront of semiconductor research. Led by distinguished experts, the tutorials offer a rare opportunity to gain deep insights into emerging technologies and their real-world applications – making them an essential start to your IEDM experience.

On Sunday, IEDM 2025 will feature two comprehensive short courses designed to deliver unparalleled insights into the technologies shaping the AI and High-Performance Computing era. These courses are curated and presented by internationally recognized experts from leading academic institutions and global technology companies. Each session is crafted to offer broad appeal, with material tailored for both emerging professionals and seasoned experts. Whether you seek to expand your foundational knowledge or gain strategic insights into next-generation architectures, these short courses are an essential part of your IEDM experience.

The main conference opens with three keynote presentation exploring the challenges and opportunities of enabling AI at the edge, provide insights into enterprise mission-critical computing, and highlight the impact of semiconductor innovation on global industries. In addition to the excellent contributed paper sessions, special Focus Sessions will feature talks from leading experts in exciting new and rapidly advancing areas.

On Tuesday evening, please join us for an interactive panel session: "FETs at 100: Yesterday's Sparks, Tomorrow's Arcs – Human Wisdom Meets AI Insight.". This unique discussion will bring together industry luminaries and AI-driven perspectives to explore how artificial intelligence is reshaping the semiconductor landscape – and how leaders are leveraging AI to transform their strategies and innovations.

Furthermore, I invite you to join us Tuesday for the IEDM Luncheon, a signature event combining celebration and strategic insight. We honor a century of the Field-Effect Transistor with reflections on its technological milestones and a forward-looking market outlook. As AI-driven demand accelerates, the semiconductor industry faces both unprecedented opportunities and significant challenges. This session will explore the forces shaping the future and strategies to navigate a rapidly evolving landscape.

In addition, this year's program, IEDM 2025 proudly presents an Exhibits Section – a curated showcase of the latest products, technologies, and publications shaping the future of electronics. Please explore the exhibition during opening hours, engage with leading vendors, and discover innovations that define our industry.

I extend my gratitude to the authors and speakers who contributed to the technical program. Your work continues to propel IEDM as the premier global conference in electron devices and related technologies. My sincere appreciation also goes to the members of the IEDM Executive and Technical Subcommittees, whose dedication and expertise were instrumental in crafting this year's exceptional program.

IEDM is proudly sponsored by the IEEE Electron Devices Society (EDS). If you are not yet a member of IEEE and EDS, I encourage you to join this distinguished institution – a cornerstone of technological progress for over 130 years. More information is available at the conference and at www.ieee.org.

Again, it is my honor to welcome you to IEDM 2025 – a gathering that will shape the future of electronics and push the boundaries of innovation.

Jan Hoentschel – General Chair



GENERAL CHAIR
Jan Hoentschel
GlobalFoundries Inc.



TECHNICAL PROGRAM CHAIR
Meng-Fan (Marvin) Chang
National Tsing Hua University & TSMC



TECHNICAL PROGRAM VICE CHAIR
Jungwoo Joh
Texas Instruments

CONTENTS

Executive Committee 4
Intro 4
Digital & Social Media 4
Topics of Interest 5
Schedule-at-a-Glance 8
Program 10
Saturday, Dec. 6 10
Tutorials 10
Sunday, Dec. 7 12
Short Courses 12
Short Course 1: Advanced Logic and System Technologies in the AI Era 12
Short Course 2: Advanced Memory Technologies for Energy-Efficient AI and HPC 14
Monday, December 8, 2025 16
Keynote Session 16
Technical Sessions 19
Session 2: ALT | Gate-All-Around Technology, CFET and DTCO 19
Session 3: MT | Embedded Memory and Selector Only Memory 21
Session 4: RSD | Ferroelectric Memory Reliability 23
Session 5: NC | LLM and Search Engines with In-Memory Compute 26
Session 6: MT/ALT/NC | Focus Session | Efficient AI Solutions: Advances in Architecture, Circuit, Device and 3D Integration 29
Session 7: MS | Compact Modeling and DTCO 31
Session 8: ODI | Emerging Photodetectors and SPADs 33
Session 9: PMA | Advanced Power and Cryogenic Devices 36
Session 10: EDT | Advances in 2D Material Devices and Integration 38
Monday Evening Reception 41
Tuesday, December 9, 2025 41
Session 11: NC | Adaptive and On-chip Learning 41
Session 12: MT | Ferroelectric Memories 43
Session 13: SMB | MEMS Resonators and Filters 46
Session 14: EDT/SMB | Focus Session | The Invisible Revolution in Thin-Film Transistors 48
Session 15: EDT | Emerging Spintronic and Cryogenic Devices 50
Session 16: PMA | High Performance RF FET and HBT Technologies 52
Session 17: MS | Thermal Analyses for Novel Transistors, Stacked Designs and Power Devices 54
Session L: FETs Lunch 56
Session 18: MT | NAND Technology 57
Session 19: ALT | Transistors with 2D, Oxide Semiconductor, Novel Materials 59
Session 20: NC | Probabilistic In-Memory Computing 62
Session 21: ODI | Displays and Integrated Photonics 64
Session 22: EDT | Advances in Oxide Semiconductor Devices: Physics, Integration, and Reliability 67
Session 23: RSD | Reliability Across Materials and Mechanisms 69
Session 24: MS | Beyond CMOS and Advanced Transport Model 72
Session 25: SMB | MEMS and Sensors for Touch, Flow, Vision, and Gas Detection 74
Session 26: PMA | Advanced GaN Power Devices: Integration, Reliability and High Voltage 76

Session 27: Evening Panel | FETs at 100: Yesterday's Sparks, Tomorrow's Arcs -- Human Wisdom Meets AI Insight 79
Wednesday, December 10, 2025 79
Session 28: EDT | Monolithic 3D and Novel Device Architectures for Post-Moore Computing 79
Session 29: MT | Oxide Semiconductors for Memory 83
Session 30: MS | Advanced Memories and Ferroelectric Devices 85
Session 31: ALT/EDT/NC | Focus Session | Beyond Von Neumann, from Physics-inspired Solvers to Quantum Hardware Implementation 87
Session 32: RSD | Compute-in-Memory (CiM) Reliability 89
Session 33: ALT | BEOL-Compatible Devices and Interconnect Technology 92
Session 34: SMB | Integrated Photo Sensors and Applications 94
Session 35: NC | Co-Design for Multimodal & Accurate In-Memory Computing 97
Session 36: PMA | GaN and III-V Integration for Next-Generation RF Devices 100
Session 37: EDT | Advances in Memory and Computing Using Hafnia based Ferroelectrics 102
Session 38: ALT | System Optimization and Process Innovation 105
Session 39: MT | MRAM and RRAM 108
Session 40: RSD | Advanced Transistor Reliability 110
Session 41: ALT/ODI | Focus Session | Silicon Photonics for Energy Efficient AI Computing 113
Session 42: ODI | Advanced Image Sensors 115
Session 43: SMB | Bioelectronic Devices and Systems 117
IEDM Committee List 120

Executive Committee



Front row left to right: Vita Pi-Ho Hu, Asian Arrangements Co-Chair; Rihito Kuroda, Asian Arrangements Chair; Geert Eneman, Publications Chair; Jungwoo Joh, Technical Program Vice Chair; Jan Hoentschel, General Chair; Meng-Fan (Marvin) Chang, Technical Program Chair; Michael Waltl, European Arrangements Co-Chair; Srabanti Chowdhury, Publications Co-Chair; Olga Spahn, Focus Session Co-Chair

Second row left to right: Byounghak Hong, ALT Chair; Uygar Avci, Focus Session Chair; Jing Wang, MS Chair; Kang-ill Seo, Courses Co-Chair (Short Courses); Bonnie Weir, RSD Chair; Hongping Zhao, PMA Chair; John Paul Strachan, Publicity Co-Chair; Sandy Liao, Courses Chair; Elena Gnani, European Arrangements Chair; Shawn Boon, Conference Manager; Jun Ogi, ODI Chair; Tania Roy, EDT Chair; Polly Hocking, Conference Manager; Gaudenzio Meneghesso, Publicity Chair

Third row left to right: Nanbo Gong, MT Chair; Dechao Guo, Virtual Arrangements Chair; Daniele Ielmini, NC Chair; Sangbum Kim, Courses Co-Chair (Tutorials); Gail Sparks-Riegel, Conference Manager; Sheng-Shian Li, SMB Chair
Missing: Phyllis Mahoney, Conference Manager;

Intro

IEEE International Electron Devices Meeting (IEDM) is the world's preeminent forum for reporting technological breakthroughs in the areas of semiconductor and electronic device technology, design, manufacturing, physics, and modeling. IEDM is the flagship conference for nanometer-scale CMOS transistor technology, advanced memory, displays, sensors, MEMS devices, novel quantum and nano-scale devices and phenomenology, optoelectronics, devices for power and energy harvesting, high-speed devices, as well as process technology and device modeling and simulation.

Digital & Social Media

- LinkedIn: https://www.linkedin.com/groups/7475096/
- Twitter: https://twitter.com/ieee_iedm
- Facebook: https://www.facebook.com/IEEE.IEDM
- YouTube: https://www.youtube.com/@iedminternationalelectrond4318
- Wikipedia: https://en.wikipedia.org/wiki/International_Electron_Devices_Meeting

Topics of Interest

ADVANCED LOGIC TECHNOLOGY (ALT)

Papers are solicited in the areas of CMOS platform technologies and applications (e.g., HPC, LOP, mobile, automotive, low temperature CMOS, etc.), logic devices and circuits, process integration schemes for advanced nodes, innovations in material, process and metrology techniques, and design technology co-optimization (DTCO) and system technology co-optimization (STCO). Platform technologies include state-of-the-art Si and beyond-Si channel devices, gate-all-around devices, stacked devices with different polarity transistors, advanced interconnect, novel power distribution integration schemes, heterogenous 2.5D/3D integration schemes, and BEOL compatible transistors. Device architecture, device design and analysis, process integration, module advancements in process and patterning, metrology, physical layout effects, techniques for reduced variability, yield, thermal management, methodologies, and solutions for DTCO/STCO in the solicited areas are of high interest.

EMERGING DEVICE and COMPUTE TECHNOLOGY (EDT)

Papers are solicited on emerging nano-electronic devices and physics. This includes devices based on novel transport and control mechanisms such as tunnel FET, negative capacitance FET, cold-source FET, cryogenic devices, topological materials and devices, phase transitions, ferroelectrics and quantum effects. Devices based on low-dimensional systems including 2D materials, CNTs, nanowires, single electron transistors and quantum dots are welcomed. Exploratory devices with novel device functions and/or novel materials for neuromorphic compute, approximate and analog compute, and non-charge-based compute such as spintronics are key topics. Furthermore, emerging state machines and time dynamical compute systems are also of interest. Qubit devices as well as devices and systems designed to enable quantum computing, quantum simulation and quantum annealing are of high interest. Papers in EDT focus primarily on device physics and novel elaboration concepts.

MEMORY TECHNOLOGY (MT)

Papers are solicited across various domains in memory and storage technologies. This encompasses advancements and scaling in established technologies like SRAM, DRAM, and Flash, as well as breakthroughs in emerging innovations such as MRAM, PCM, FRAM, RRAM, ECM, Cross-Point memory, organic memory, and NEMS-based memory. These include research pertaining to materials and devices for memory and selector in electrostatic and atomistic switching mechanisms, as well as the design and implementation of memory cells and arrays in 3D constructs, including stacking and tiering, alongside read/write access mechanisms. Additionally, submissions of novel approaches to homogenous or heterogenous integration and manufacturing techniques of memory in semiconductor fabrication or packaging assembly are welcome. Authors are also invited to present papers focusing on reliability characterization and physics exploration, as well as improvements through manufacturing and usage algorithm, such as wear leveling. Furthermore, contributions addressing the span-expanding and gap-bridging of tiering and hierarchy in memory subsystem are encouraged, as are those addressing memory subsystems tailored to a broad spectrum of computing needs, ranging from mobile to large-scale platforms for client to cloud computing, encompassing both general-purpose and AI-based architectures. Papers disclosing novel physics and computing architecture may be transferred to other technical sections such as EDT or NC at the discretion of the committee.

MODELING and SIMULATION (MS)

Papers are solicited on theoretical or computational studies of electronic devices, including logic, memory, optical components, interconnects, and (bio)sensors. The approaches used in the study may include analytical, numerical, statistical, and machine-learning/AI-based methods applied to structures ranging from atomistic to device dimensions, and up to full-chip dimensions, including physics-based compact modeling. Central to submissions is the innovation of devices, whether through predictive insight into novel device concepts, predictive analysis demonstrating significant device improvements, breakthroughs in theoretical understanding of device operation, advancements in knowledge of device processing facilitating enhanced device performance, novel insights into variability, reliability, and yield issues, or breakthroughs in device optimization based on DTCO. Topics also include ab-initio/atomistic materials modeling, neuromorphic computing modeling, quantum computing, spintronics, low-dimensional devices, ferroelectrics, thermal modeling, 3D/heterogeneous integration, electro-chemical/mechanical devices, and compact modeling methodology innovation. Encouragement is given for comparison with experimental data, model calibration, and utilization of multi-scale simulation chains.

NEUROMORPHIC and NOVEL COMPUTING (NC):

Papers are solicited in areas related to memory and logic devices, circuits, and algorithms for new and unconventional compute paradigms, including but not limited to neuromorphic computing and other artificial intelligence (AI) acceleration techniques. This includes, for example, analog/mixed-signal computing, in-memory computing devices and circuits for deep learning, probabilistic computing, bio-inspired computing such as spiking neural networks, reservoir computing, combinatorial optimization, content-addressable memory, etc. Use cases in the datacenter as well as on the edge, e.g. in-sensor computing, are contemplated. Demonstration of novel device concepts improving computational efficiency, full hardware integration, device-algorithm co-optimization to mitigate non-ideal device properties, and real-world applications are of high interest. Papers on near-memory computing may be transferred to MT, and papers centered on optical or quantum computing may be transferred to EDT, at the discretion of the committee.

OPTOELECTRONICS, DISPLAYS, and IMAGING SYSTEMS (ODI)

Papers are solicited on optoelectronics, displays, and imaging systems. This includes novel devices, structures, and integration for image sensors, displays, light sources, photonic devices, and high-speed photodetectors and modulators. New technologies on heterogeneous integration of optoelectronics devices as well as on photonic-electronic integration for optical interconnects, on-chip networks and sensing are welcomed. Papers on quantum photonics, neuromorphic photonics, and plasmonics for quantum computation, sensing and encryption are also of interest. Furthermore, ODI includes CMOS imagers, high-speed and high-time resolution imagers, stacked imagers sensors, single-photon sensitivity, Time-Of-Flight and nonvisible image sensors. In addition, papers on displays of all types, for augmented or virtual reality, holography, TFTs for photonics applications, flexible, stretchable, and/or printed electronics, in-display sensors are encouraged. Papers on displays or light emitting devices with novel materials such as perovskites or quantum dots are also of interest.

POWER, MICROWAVE/MM-WAVE and ANALOG DEVICES/SYSTEMS(PMA)

Contributions are sought on novel circuit topologies, manufacturing processes, supporting modeling (TCAD and compact models), device physics, reliability, and materials (Si, III-Vs, SiC, (Al)GaN, Ga_2O_3, AlScN, $LiNbO_3$, diamond, LTO, etc.) along with fundamental studies on doping, deep-level traps, interface state densities, and device reliability for power and/or high frequency devices. Papers are solicited on discrete and integrated power and/or high frequency (micro, mm-wave and THz) devices and physics, modules and systems. Topics of interest include devices (diodes, BJTs, FETs, super-junction devices, heterostructures, IGBTs, HEMTs, HBTs, light-triggered structures for galvanic isolation and faster switching, bi-directional switches, vertical geometry devices, RF acoustic resonators, SAW and BAW filters etc.) and device/package/circuit interactions, including thermal management. Wide variety of applications are also of interest (power conversion, supply, regulation and conditioning for computers and data centers, motor drives, transportation, solar, wind, smart grid applications, wireless power harvesting/transfer, filters, beam formers, power amplifiers, tunable passives, antenna arrays).

RELIABILITY OF SYSTEMS and DEVICES (RSD)

Papers are solicited that focus on component-level FEOL/MEOL/BEOL characterization and reliability modeling, reliability evaluation, both experimental and modeling, of devices dedicated to analog, logic, and memory applications, interconnects, circuits, and systems. In addition to Si-based technologies, authors are encouraged to submit their recent achievements made employing other material systems, such as SiGe, IGZO, ferroelectric materials, 2D materials, etc. The reliability topics include, for FEOL, transistor degradation due to hot carriers, bias temperature instabilities, random telegraph noise, dielectric SILC and breakdown as well as modeling the aging and wearout behavior. For MEOL/BEOL, topics include the breakdown of MEOL spacers and BEOL dielectrics, electromigration, and stress migration failures of contacts and interconnects. For product, system, and circuit reliability, topics include latch-up, ESD, soft error mechanisms, variability-aware design, and design for reliability, robustness, and security of electronic circuits and systems. Of particular interest are investigations of degradation mechanisms for devices, circuits, and systems in the following areas: conventional and emerging memories; beyond CMOS devices; 3D IC package reliability; more-than-Moore applications; biomedical devices and systems; automotive and aerospace.

SENSORS, MEMS, and BIOELECTRONICS (SMB)

Papers are solicited in the areas of sensors, transducers, micro/nano electromechanical systems (MEMS and NEMS), organ-on-chip, microfluidics/lab-on-chip, BioMEMS, and bioelectronic devices and systems, with particular emphasis on new device concepts, integrated and highly parallel CMOS implementations, CMOS-MEMS, embedded machine learning, organic-inorganic hybrid microfabrication, flexible devices, soft devices, and multimodal sensors on a chip for applications in health, medicine, communication, mobility, and energy. Sensors include chemical, molecular and biological detection based on acoustic, electrical, electrochemical, magnetic, mechanical, and optical principles. Topics of interest in the MEMS area include actuators, physical and biochemical sensors (BioMEMS), acoustic resonators, integrated inertial measurement units, RF MEMS, SAW and BAW filters, SAW and BAW sensors, optomechanical devices, micro-power generators, and devices for energy harvesting as well as on-chip energy storage. Bioelectronics covers organic-inorganic hybrid devices, point-of-care biomedical devices, bio-electronic interfaces, integrated biomedical sensing, and implantable sensors and neural interfaces.

Schedule-at-a-Glance

SATURDAY - TUTORIALS (Separate Registration Required)				
12:00 PM	**-**	**6:00 PM**	**East Lounge**	**Registration** (12-3:30 pm: Tutorial registrants only, Open to all at 3:30 pm)
1:30 PM	-	2:50 PM	Cont 1-5	T1 | Advances in CMOS Technologies and Cell Height Scaling Considerations
1:30 PM	-	2:50 PM	Cont 6-9	T2 | In-memory Computing Approaches for Large Language Model Acceleration
3:05 PM	-	4:25 PM	Cont 1-5	T3 | Atomic-layer-deposited atomically thin In2O3 transistors for BEOL logic and memory applications
3:05 PM	-	4:25 PM	Cont 6-9	T4 | Quantum computing – Concepts, status quo and challenges ahead
4:40 PM	-	6:00 PM	Cont 1-5	T5 | Developing a reliable and open chiplet eco-system for wafer-scale systems
4:40 PM	-	6:00 PM	Cont 6-9	T6 | High-Speed Memory Solutions for ML and High-Performance Compute
SUNDAY - SHORT COURSES (Separate Registration Required)				
8:00 AM	**-**	**5:00 PM**	**East Lounge**	**Registration**
9:00 AM	-	6:00 PM	Cont 1-5	SC1 | Advanced Logic and System Technologies in the AI Era
9:00 AM	-	6:00 PM	Cont 6-9	SC2 | Advanced Memory Technologies for Energy-Efficient AI and HPC
MONDAY				
8:00 AM	**-**	**5:00 PM**	**East Lounge**	**Registration**
9:00 AM	**-**	**12:00 PM**	**Grand B**	**1 | Keynote**
12:00 PM	**-**	**5:30 PM**	**Yosemite**	**Exhibits Open**
1:30 PM	**-**	**5:20 PM**	**Technical Sessions**	
1:30 PM	-	5:20 PM	Grand A	2 | ALT | Gate-All-Around Technology, CFET and DTCO
1:30 PM	-	4:30 PM	Grand B	3 | MT | Embedded Memory and Selector Only Memory
1:30 PM	-	5:20 PM	Cont 1-3	4 | RSD | Ferroelectric Memory Reliability
1:30 PM	-	4:55 PM	Cont 4	5 | NC | LLM and Search Engines with In-Memory Compute
1:30 PM	-	5:20 PM	Cont 5	6 | Focus Session | Efficient AI Solutions: Advances in Architecture, Circuit, Device and 3D Integration
1:30 PM	-	4:30 PM	Cont 6	7 | MS | Compact Modeling and DTCO
1:30 PM	-	4:55 PM	Cont 7-9	8 | ODI | Emerging Photodetectors and SPADs
1:30 PM	-	5:20 PM	Imperial A	9 | PMA | Advanced Power and Cryogenic Devices
1:30 PM	-	5:20 PM	Imperial B	10 | EDT | Advances in 2D Material Devices and Integration
2:45 PM	**-**	**3:45 PM**	**Yosemite**	**Coffee Break with Exhibitors**
6:30 PM	**-**	**8:00 PM**	**Grand B**	**Monday Evening Reception**
TUESDAY				
8:30 AM	**-**	**5:00 PM**	**East Lounge**	**Registration**
8:30 AM	**-**	**5:00 PM**	**Yosemite**	**Exhibits Open**
9:00 AM	**-**	**12:25 PM**	**Technical Sessions**	
9:00 AM	-	11:35 AM	Grand A	11 | NC | Adaptive and On-chip Learning
9:00 AM	-	12:25 PM	Grand B	12 | MT | Ferroelectric Memories
9:00 AM	-	12:25 PM	Cont 1-3	13 | SMB | MEMS Resonators and Filters
9:00 AM	-	12:25 PM	Cont 4	14 | Focus Session 2 - The Invisible Revolution in Thin-Film Transistors
9:00 AM	-	12:00 PM	Cont 5	15 | EDT | Emerging Spintronic and Cryogenic Devices
9:00 AM	-	12:25 PM	Cont 6	16 | PMA | High Performance RF FET and HBT Technologies
9:00 AM	-	12:00 PM	Cont 7-9	17 | MS | Thermal Analyses for Novel Transistors, Stacked Designs and Power Devices
10:15 AM	**-**	**11:15 AM**	**Yosemite**	**Coffee Break with Exhibitors**
12:20 PM	**-**	**2:00 PM**	**Plaza**	**FETs Lunch (Ticket Required)**

2:15 PM	**-**	**6:05 PM**	**Technical Sessions**	
2:15 PM	-	6:05 PM	Grand A	18 | MT | NAND Technology
2:15 PM	-	6:05 PM	Grand B	19 | ALT | Transistors with 2D, Oxide Semiconductor, Novel Materials
2:15 PM	-	5:15 PM	Cont 1-3	20 | NC | Probabilistic In-Memory Computing
2:15 PM	-	5:40 PM	Cont 4	21 | ODI | Displays and Integrated Photonics
2:15 PM	-	5:40 PM	Cont 5	22 | EDT | Advances in Oxide Semiconductor Devices: Physics, Integration, and Reliability
2:15 PM	-	6:05 PM	Cont 6	23 | RSD | Reliability Across Materials and Mechanisms
2:15 PM	-	5:15 PM	Cont 7-9	24 | MS | Beyond CMOS and Advanced Transport Model
2:15 PM	-	5:40 PM	Imperial A	25 | SMB | MEMS and Sensors for Touch, Flow, Vision, and Gas Detection
2:15 PM	-	6:05 PM	Imperial B	26 | PMA | Advanced GaN Power Devices: Integration, Reliability and High Voltage
3:30 PM	**-**	**4:30 PM**	**Yosemite**	**Coffee Break with Exhibitors**
8:00 PM	**-**	**10:00 PM**	**Cont 4-5**	**EP | FETs at 100: Yesterday's Sparks, Tomorrow's Arcs -- Human Wisdom Meets AI Insight**
WEDNESDAY				
8:30 AM	**-**	**3:00 PM**	**East Lounge**	**Registration**
9:00 AM	**-**	**11:30 AM**	**Yosemite**	**Exhibits Open**
9:00 AM	**-**	**12:50 PM**	**Technical Sessions**	
9:00 AM	-	12:50 PM	Grand A	28 | EDT | Monolithic 3D and Novel Device Architectures for Post-Moore Computing
9:00 AM	-	12:25 PM	Grand B	29 | MT | Oxide Semiconductors for Memory
9:00 AM	-	12:00 PM	Cont 1-3	30 | MS | Advanced Memories and Ferroelectric Devices
9:00 AM	-	12:00 PM	Cont 4	31 | Focus Session | Beyond Von Neumann, from Physics-inspired Solvers to Quantum Hardware Implementation
9:00 AM	-	12:50 PM	Cont 5	32 | RSD | Compute-in-Memory (CiM) Reliability
9:00 AM	-	12:50 PM	Cont 6	33 | ALT | BEOL-Compatible Devices and Interconnect Technology
9:00 AM	-	12:25 PM	Cont 7-9	34 | SMB | Integrated Photo Sensors and Applications
9:00 AM	-	12:25 PM	Imperial A	35 | NC | Co-Design for Multimodal & Accurate In-Memory Computing
9:00 AM	-	12:25 PM	Imperial B	36 | PMA | GaN and III-V Integration for Next-Generation RF Devices
10:15 AM	**-**	**11:15 AM**	**Yosemite**	**Coffee Break with Exhibitors**
1:30 PM	**-**	**5:20 PM**	**Technical Sessions**	
1:30 PM	-	5:20 PM	Grand A	37 | EDT | Advances in Memory and Computing Using Hafnia based Ferroelectrics
1:30 PM	-	5:20 PM	Grand B	38 | ALT | System Optimization and Process Innovation
1:30 PM	-	5:20 PM	Cont 1-3	39 | MT | MRAM and RRAM
1:30 PM	-	5:20 PM	Cont 4	40 | RSD | Advanced Transistor Reliability
1:30 PM	-	5:20 PM	Cont 5	41 | Focus Session | Silicon Photonics for Energy Efficient AI Computing
1:30 PM	-	5:20 PM	Cont 6	42 | ODI | Advanced Image Sensors
1:30 PM	-	4:55 PM	Cont 7-9	43 | SMB | Bioelectronic Devices and Systems
2:45 PM	**-**	**4:00 PM**	**East Lounge**	**Coffee Break**

Program

Saturday, Dec. 6

Tutorials

The tutorials are in their fifteenth year and are standalone presentations on specialized topics taught by world-class experts. These tutorials will provide a brief introduction to their respective fields and facilitate understanding of the technical sessions. In contrast, the traditional short courses are focused on a single technical topic.

Co-Chairs: Sangbum Kim, Seoul National University
and Sandy Liao, TSMC
1:30 PM - 6:00 PM

Tutorial 1: Advances in CMOS Technologies and Cell Height Scaling Considerations, Yan Bin Luo, Intel
1:30 PM– 2:50 PM
Continental 1-5
Chair: Sangbum Kim, Seoul National University

Artificial Intelligence is seeing a tremendous jump in capability and drastically expanded application area in the last few years. The complex AI models require high performance computing (HPC) systems. While Artificial General Intelligence is still being developed, the technology will have even higher performance needs. CMOS technology is essential in building HPC systems for AI applications. This tutorial focuses on the latest advances on CMOS scaling from Fin-FET technology, nano-sheet or Gate All Around, dielectric separated N-P sheet or fork-sheet to stacked N-P nano sheet or CFET. Due to the increased pace and demand, additional scaling considerations such as cost and time to market have been added on top of power/performance/Area. The overall scaling considerations will strongly affect process optimization of all critical dimensions around the active transistor and BEOL dimensions. Back-side power delivery network (BSPDN) offers additional power/performance scaling/optimization point. Within a process node, Design Technology Co-Optimization (DTCO) methodology is playing an increasingly important role for performance offering and design for manufacturing optimization.

Tutorial 2: In-memory Computing Approaches for Large Language Model Acceleration, Manuel Le Gallo, IBM Research Europe
1:30 PM– 2:50 PM
Continental 6-9
Chair: Sangbum Kim, Seoul National University

Recent advancements in large language models (LLMs) have significantly revolutionized natural language processing, empowering machines to produce text akin to humans and participate in meaningful dialogues. As these systems grow in complexity, demanding substantial computational and memory resources, there is an urgent need for enhanced speed, efficiency, and accessibility in LLM inference. Due to their size, LLMs now exceed the capacity of single GPUs, necessitating intricate, expert-level configurations for parallel operation. The cost of memory access surpasses that of computations, creating a barrier to efficient scaling. In-memory computing (IMC) technologies present a promising avenue to alleviate the growing data communication for LLMs by performing computations directly within memory. By merging memory and computing components seamlessly, IMC alleviates the von Neumann bottleneck, lessening data transfers, and enhancing energy efficiency.

In this tutorial, we will first present the Transformer architecture at the core of modern LLMs and outline the challenges it poses for hardware acceleration. This will be followed by a brief presentation of conventional approaches to run LLMs on general-purpose and special-purpose digital accelerators. We will then cover the DRAM processing-in-memory (PIM) accelerators that have recently emerged to improve throughput and energy efficiency over GPU baselines. Finally, we will review various IMC architectures with volatile and non-volatile memristive devices, as well as Flash memory, and discuss how they can address the memory access bottlenecks through large on-chip weight capacity. Emphasis will be made on the device-level and system-level requirements of LLM architectures for IMC acceleration, as well as challenges in maintaining high inference accuracy with this approach.

Tutorial 3: Atomic-layer-deposited Atomically Thin In_2O_3 Transistors for BEOL Logic and Memory Applications, Peide Ye, Purdue University
3:05 PM– 4:25 PM
Continental 1-5
Chair: Sangbum Kim, Seoul National University

In this talk, we will review the recent progress on oxide semiconductor thin-film transistor (TFT) technologies for back-end-of-line (BEOL) 3D integrations. We will focus on the first experimental demonstration of atomically thin In2O3 channel for logic and memory devices by a BEOL-compatible atomic layer deposition (ALD) process [1]. High performance planar In2O3 transistors with high mobility of 153 cm2/V·s and record high maximum drain current of near 20 mA/um are achieved by gate-all-around structure and thermal engineering [2]. ALD In2O3 3D Fin transistors are also demonstrated, benefiting from the conformal deposition capability of ALD. High reliability and record low contact resistance/resistivity reaching Landauer limit are demonstrated experimentally [3]. High-performance In2O3 ferro-electric transistors are demonstrated using ALD HfZrO2 gating with >2.2V large memory window, >10 years retention and >109 endurance [4]. These results suggest ALD oxide semiconductors and devices have unique advantages and are promising toward BEOL-compatible monolithic 3D integration [5].

1. M. Si et al. IEEE EDL 42(2), 184-187, 2020; M. Si et al. Nano Lett. 21 (1), 500-506, 2020.
2. M. Si et al. VLSI, T2.4, 2021; Z. Lin et al., VLSI, 2024. Z. Zhang et al. IEEE EDL 43 (11), 1905, 2022; Z. Lin et al. ACS Nano 2022.
3. Z. Zhang et al. VLSI, T11.3, 2023; C. Niu et al., IEDM, 37-2, 2023.
4. Z. Lin et al. IEDM, TF17.4, 386, 2021; Z. Lin et al. VLSI, TF13-2, 2022.
5. M. Si, Z. Lin et al. Nature Electronics 5(3), 164-170, 2022. A. Charnas et al., Advanced Materials, 2023.

Tutorial 4: Quantum Computing – Concepts, Status Quo and Challenges Ahead, Hendrik Bluhm, RWTH Aachen University, Forschungszentrum Jülich, ARQUE Systems GmbH
3:05 PM– 4:25 PM
Continental 6-9
Chair: Sangbum Kim, Seoul National University

Quantum computing aspires to solve some computational problems that are currently intractable for problem sizes of practical interest by leveraging quantum superposition states and entanglement as a fundamentally new way to encode information. Significant progress has been made in realizing quantum computing systems with impressive capabilities, yet a utility advantage so far remains elusive. Current implementations start to reach the qubit numbers needed for “noisy intermediate scale quantum computing” (NISQ), but significant improvements in the error rates are needed for applications. Quantum error correction can correct unavoidable errors and thus enable large scale “universal quantum computing” but requires orders of magnitude more qubits than available today.

This tutorial will survey the key concepts and some platforms of quantum computing and compare the state of the art to the expected requirements for both NISQ and error corrected quantum computers. I will conclude with an outlook on how semiconductor-based qubits in combination with cryoelectronic control systems might reach the required system scale in the spirit of large-scale integration. Addressing the associated challenges will likely benefit from smart adaptions of classical semiconductor technology, such as circuit components optimized for low temperature, ultra-low-power circuits and 3D-integration solutions with tailored heat management capabilities.

Tutorial 5: Developing a Reliable and Open Chiplet Eco-system for Wafer-scale Systems, Subramanian Iyer, UCLA
4:40 PM– 6:00 PM
Continental 6-9
Chair: Sangbum Kim, Seoul National University

Advanced packaging at wafer-scale rests on three foundations: 1. scale down and scale out; 2. An open chiplet eco-system and 3. An autonomous built-in test and repair methodology. Making these work in concert especially when there is no viable rework capability is challenging and needs to be addressed for advanced packaging to realize its full potential.

In this tutorial we explain how these three come together when we have an abundance of wires and connections, relatively small chiplets assembled at fine pitch on smart substrates with a high wiring density with a granular design including power delivery and thermal dissipation, utility dielets that monitor and repair the system during its lifetime, the concept of graceful failure, along with the ability to assemble and functionally interconnect multiple wafer scale systems without mechanical constraints.

Tutorial 6: High-Speed Memory Solutions for ML and High-Performance Compute, Eric Yih Wang, TSMC
4:40 PM– 6:00 PM
Continental 6-9
Chair: Sangbum Kim, Seoul National University

High-speed embedded memory is essential in meeting the growing demand for higher computational efficiency in next-generation AI and High-Performance Computing (HPC) chips. This tutorial will introduce the fundamentals of cache and memory system for Machine Learning (ML) and HPC and provide an overview of current memory technologies and architectures, such as SRAM and digital compute-in-memory. Additionally, it will explore alternative high-density embedded memory solutions with the potential to enhance the performance, energy efficiency, density and functionality of embedded memory in future AI and HPC system-on-chip.

Sunday, Dec. 7

Short Courses

IEEE IEDM will offer two short courses with in-depth coverage of highly relevant topics from world experts.

Co-Chairs: Kang-ill Seo, IBM
and Sandy Liao, TSMC
1:30 PM - 6:00 PM

Short Course 1: Advanced Logic and System Technologies in the AI Era

9:00 AM, Continental 1-5
Co-Chairs: Kang-ill Seo, IBM and Tenko Yamashita, IBM

9:00 AM - 10:15 AM
SC1-1 | Device Technologies for GAA Scaling and Emerging Alternatives, Myungil Kang, Samsung Electronics
The increasing demand for enhanced power efficiency, performance, and area optimization in mobile and high-performance computing devices—such as smartphones, wearables, and HPC servers—has driven continuous advancements in logic technology. Traditional planar Si-MOSFETs have successfully followed Moore's law through key innovations, including stress engineering and high-k metal gate, eventually evolving into 3D structures such as FinFET and GAA (Gate-All-Around) nanosheet FET.

Notably, the transition to GAA nanosheet device technology is nearing completion at the 3nm and 2nm nodes, alongside the adoption of the Back Side Power Delivery Network (BSPDN), which enhances platform technology for high-performance computing applications. Similar to FinFET generations, GAA devices are expected to remain the mainstream technology for over a decade. However, continued cell height scaling will necessitate a shift toward Forksheet and 3D stacked FET architectures.

This short course will explore the development of GAA nanosheet device technology, its evolution, and the potential successor technologies that aim to overcome current architectural limitations.

10:45 AM - 12:00 PM
SC1-2 | Advanced Interconnect Technologies for Cu Extension and Beyond, Koichi Motoyama, IBM
Copper has been used as the BEOL interconnect metal for many generations since it was introduced in1997. However, due to continued scaling of BEOL Cu interconnect dimensions, several challenges have arisen in recent technology nodes such as an increase in line resistance and EM performance degradation. Therefore, post-Cu alternative conductors have been studied extensively as an approach to enable lower line resistance and excellent

reliability for sub-20 nm pitch technologies due to their smaller scattering effect, less stringent barrier/liner requirements, and higher melting point compared to Cu. In this short course, we will review what the industry has been developing to overcome challenges for extending Cu interconnects towards future technology nodes. Moreover, we will discuss various aspects of post-Cu alternative metal interconnects such as metal candidates, current proposed processes (patterning scheme and interconnect structure), and key challenges to provide the outlook of the future advanced interconnect technology.

1:00 PM - 2:15 PM

SC1-3 | Process Innovations and Materials Sustaining Moore's Law, Nicolas Breil, AMAT

As CMOS scaling continues its relentless journey, process technology and material innovations become foundational to keep delivering power and performance improvements with scaling. We will introduce this short course by identifying the key bottlenecks for performance improvements in the various transistor modules (superlattice, channel, source-drain, contact module and interconnects). We will present a first principles analysis of the performance and/or capacitance limitations of the transistor elements and will connect these with solutions relying on innovative materials and/or process technologies, in the context of the current hGAA and upcoming CFET CMOS technologies.

2:45 PM - 4:00 PM

SC1-4 | Silicon Photonics for High-Speed, Energy-Efficient Data Transfer, Di Liang, University of Michigan

Silicon photonics has rapidly emerged as a transformative integration platform for applications spanning data communications, sensing, and computing. This short course will provide a comprehensive introduction to the fundamentals of silicon photonics, covering key material systems, core passive and active device building blocks, and integration schemes. Special emphasis will be placed on the practical challenges and solutions in large-scale manufacturing, advanced packaging, wafer-level testing, and device variability. The course will review the state of the art in high-performance interconnects, currently the leading application, as well as explore emerging frontiers such as quantum photonics, LiDAR, and neuromorphic computing. Designed for device engineers, circuit designers, and system architects, this course aims to deliver both foundational knowledge and practical insights into one of the most dynamic and rapidly evolving fields at the intersection of microelectronics and integrated photonics.

4:00 PM - 5:00 PM

SC1-5 | Advanced Packaging and Chiplet Technologies for AI and HPC, James Chen, TSMC

As computational demands surge, traditional monolithic chip with 2D package designs face limitations in scalability, power efficiency and thermal management. Advanced packaging techniques, including 2.5D and 3D integration, offer promising solutions by allowing heterogeneous integration. These innovations not only address power consumption and bandwidth but also provide a scalable path to meet the growing computational needs. This talk will delve into the cutting-edge advancements in packaging and chiplet technologies, emphasizing their pivotal role in enhancing AI and High-Performance Computing (HPC) applications. Additionally, we'll examine the critical aspects of thermal solutions and stress management strategies essential for maintaining performance and reliability in these dense and complex configurations. The discussion will also provide insights into future directions and the significant impact on next-generation computing systems.

5:00 PM - 6:00 PM

SC1-6 | System Technology Co-Optimization (STCO) for Thermal Management and Power Delivery in Heterogeneous Integration, Kenneth Larson, Synopsys

Heterogeneous integration is pivotal for overcoming throughput, bandwidth, latency, power efficiency, and thermal challenges in high-performance computing systems, especially for emerging AI workloads demanding unprecedented performance density. This work presents a comprehensive System Technology Co-Optimization (STCO) methodology focused on thermal management and power delivery for 3D heterogeneous chiplet platforms. The approach addresses critical issues of power density and hotspot mitigation inherent in 3D stacked systems by integrating co-design principles across cooling, packaging, and power delivery. The methodology enables the optimization of system-level cost, performance, energy efficiency, and thermal key performance indicators (KPIs), facilitating the development of next-generation 3D chiplet platforms with significantly enhanced volumetric performance.

Short Course 2: Advanced Memory Technologies for Energy-Efficient AI and HPC

9:00 AM, Continental 6-9
Co-Chairs: Kang-ill Seo, IBM and Duygu Kuzum, UC San Diego

9:00 AM - 10:15 AM
SC2-1 | Memory-Centric AI Architectures Via Advanced Packaging and Heterogeneous Integration, Nuwan Jayasena, AMD
AI workloads place multi-faceted demands on compute architectures and memory systems. Advanced packaging techniques have played a crucial role in meeting these demands by enabling aggressive scaling of compute capabilities and tighter integration of memory. In this short course, we will briefly review some important workload characteristics driving AI architectures and look at historical trends in the evolution of AI architectures and their memory systems. We will also go over a case study of a chiplet-based AI accelerator that makes extensive use of advanced packaging and heterogenous integration techniques. Finally, we will look at some impending challenges to the continued scaling of AI architectures and potential solutions that may be necessary to address them.

10:45 AM - 12:00 PM
SC2-2 | Near- and In-Memory Computing for AI, Hai Li, Duke University
As artificial intelligence (AI) transforms industries, state-of-the-art models have exploded in size and capability. However, deploying these models on resource-constrained edge devices remains a significant challenge. Smartphones, wearables, and IoT sensors face stringent limitations on computing, memory, power, and communication, creating a big gap between demanding AI models and edge hardware capabilities that hinders the deployment of intelligence. In this talk, we will re-examine techniques to bridge this gap and embed big AI on small devices. We will begin by discussing how the properties of various hardware platforms impact the design strategies of efficient deep neural network (DNN) models, such as quantization and pruning. Next, we will discuss techniques aimed at reducing the inference and training costs of distributed collaborative edge AI systems. Finally, we will delve into the underlying design philosophies and their evolution toward efficient, scalable, robust, and secure edge computing systems.

1:00 PM - 2:15 PM
SC2-3 | High-Performance Memory Technologies: DRAM and HBM, Dae-hyun Moon, Samsung Electronics
In this short course, we'll explore the fundamental structure and operation of DRAM, along with the rationale behind shrinking its size to achieve high-performance capabilities. We'll then examine the reliability challenges that arise from miniaturization, covering phenomena such as pass-gate effects due to shorter wordline (WL) distances, row hammering, increased defects caused by traps in smaller DRAM cells, and failures triggered by thermal neutrons from cosmic radiation. Lastly, we'll highlight HBM (High Bandwidth Memory), born out of the demand for high-performance computing DRAM in advanced applications like AI accelerators, as a fitting conclusion to our exploration.

2:45 PM - 4:00 PM
SC2-4 | High-Density 3D Flash Memory in the CMOS Directly Bonded to Array (CBA) Era and Beyond, Katsuyuki Sekine, Kioxia
Rapid advances in AI technology are driving market demand for 3D flash memory that offers higher bit density, lower cost, and higher-speed operation. Innovation in 3D flash memory structures, processes, and materials is currently crucial and will become even more vital in the future. CMOS directly bonded to array (CBA) technology, in which CMOS and the cell array are made on separate wafers and joined by Cu direct bonding, enables us to open the door to a new era of 3D flash memory. CBA technology also allows us to realize multi-wafer stacking, such as multi-stack CMOS and multi-stack cell arrays that can be combined into a single die, enabling the realization of new products. In this short course, newly developed architecture scaling techniques, including crossed bit line (CBL) and on-pitch select gate (OPS), will be presented in addition to the CBA technology itself. Moreover, process and material innovations to realize high-density and high-speed 3D flash memory in the CBA era will be introduced. These include: cryo etching for increasing the number of WLs, merging of various high-aspect etches for cost reduction, and molybdenum (Mo) WL for high-speed reading and improving reliability in the scaled cell arrays. We will also discuss other techniques like metal-induced lateral crystallization (MILC) channel and Schottky source contact (SSC), which are used to supply sufficient electrons and holes for reliable and high-speed cell operation even in cell arrays with a

high layer count. Finally, one of the promising alternative candidates, advanced horizontal channel flash (HCF) will be presented.

4:00 PM - 5:00 PM

SC2-5 | Process and Material Innovations for Memory Scaling, Noboru Ooike, TEL

Memory technologies play a crucial role in the era of AI, where enhancing memory performance and increasing capacity are essential in the rapidly evolving landscape of artificial intelligence. This short course will cover recent advancements in process technologies and material innovations specifically designed for the ongoing evolution of memory technologies. In DRAM scaling, key focus areas include high-aspect-ratio (HAR) capacitor etching and improvements in capacitor and electrode materials. Efficient capacitor scaling requires higher-k, lower-leakage film stacks. Furthermore, EUV patterning plays a significant role in simplifying process integration and improving variability. As scaling limitations arise, emerging processes such as wafer bonding in VCT and 3D DRAM will be explored. For 3D NAND, both vertical and lateral scaling present challenges in increasing bit density. The merging of HAR etching processes and high-throughput HAR etching promote vertical scaling and cost reduction. Innovations such as new channel materials, word line (WL) metals, and isolated charge trap (CT) structures are expected to advance scaling efforts further. Additionally, the shift from CMOS under Array (CuA) to CMOS bonded to Array (CbA) enhances performance and bit density. Continuous innovations in processes and tools —including film deposition, etching, lithography, wet processing, wafer bonding, and thinning—are crucial to support these advancements.

5:00 PM - 6:00 PM

SC2-6 | Emerging Memory Technologies for Next-Generation AI and HPC Applications, Agostino Piroviano, Micron

Starting from the evolving landscape of Artificial Intelligence (AI) and High-Performance Computing (HPC) applications, we will explore and analyze the evolution and challenges of the incumbent technologies for memory (i.e. DRAM) and storage (i.e. NAND). We will then extend our analysis to the characteristics of several promising emerging memory technologies (Resistive RAM, Phase-Change RAM, Magnetic RAM, Single-Chalcogenide Xpoint Memory, etc.), in order to comprehend their strengths and weaknesses against the attributes required to succeed in the AI and HPC arena.

Monday, December 8, 2025

Keynote Session

Session 1: Welcome and Awards Presentations
9:00 AM, Grand Ballroom B
Co-Chairs: Jan Hoentschel, GlobalFoundries and Meng-Fan (Marvin) Chang, National Tsing Hua University and TSMC

2024 Roger A. Haken Best Student Paper Award

Awarded To: Shuhan Liu
For the paper entitled: *"Edge Continual Training and Inference with RRAM-Gain Cell Memory Integrated on Si CMOS"*

2024 EDS Paul Rappaport Award

Awarded To: Kim, Sang-Hyeon; Kim, SeongKwang; Lim, Hyeong-Rak; Jeong, Jaejoong; Lee, Seung Woo; Jeong, Ho Jin; Park, Ju-Hyuk; Kim, Joon Pyo; Jeong, Jaeyong; Kim, Bong ho; Ahn, Seung-Yeop; Park, Youngkeun; Geum, Dae-Myoung; Kim, Younghyun; Baek, Yongku; Cho, Byung Jin
For the paper entitled: "*Heterogeneous 3-D Sequential CFETs With Ge (110) Nanosheet p-FETs on Si (100) Bulk n-FETs*"

2024 EDS Leo Esaki Award

Awarded To: Vachon, Frédéric; Parent, Samuel; Gauthier, Valérie; Lamoureux, Steve; Paquette, Alexandre; Deschamps, Jacob; Rossignol, Tommy; Roy, Nicolas; Arsenault, Philippe; Dautet, Henri; Charlebois, Serge; Pratte, Jean-François
For the paper entitled: "*Wafer-level Characterization and Monitoring Platform for Single-Photon Avalanche Diodes*"

2025 IEEE/EDS Fellows

**This is a complete listing of the 2025 IEEE/EDS Fellows. Not all Fellows will be recognized at the 2025 IEDM*

Ana Arias - for contributions to printed and flexible large-area electronics
Shujun Cai - for contributions to GaN microwave devices for RF systems applications
Yang Chai - for contributions to bioinspired in-sensor computing
Larry Dunleavy - for contributions to commercial development of microwave device models and microwave education leadership
L. Jay Guo - for contributions to nanoimprint, scalable nanopatterning
Cherie Kagan - for contributions to colloidal nanocrystals and their integration in optical and optoelectronic devices
Max Lemme - for contributions to the development of graphene-based electronic devices
Santiago Marco - for contributions to the use of signal processing and machine learning in chemical sensing systems
Yuri Masuoka - for contributions to leading-edge foundry transistor development
Matteo Meneghini - for contributions to the reliability of wide bandgap electronic and optoelectronic devices
Shayan Mookherjea - for contributions to optoelectronic devices and their applications
Tse Nga Ng - for contributions to flexible organic electronics used in large-area imagers and sensing systems
Jamie Phillips - for contributions to quantum dot infrared photodetector and microscale photovoltaic energy harvesting
Angelo Pinto - for technical leadership and contributions to FinFET technology adoption into volume manufacturing of mobile products

Malin Harindhu Premaratne - for contributions to theory, modelling, and simulations of optical quantum devices
Makoto Takamiya - for contributions to digitally controlled integrated power management circuits
Kea-tiong Tang - for contributions to smart miniature electronic nose circuits and systems
Han Wui Then - for the development and heterogeneous integration of GaN electronics on 300-mm Silicon wafers
Dina H. Triyoso - for contributions to high-k metal gate complementary metal-oxide-semiconductor technology
Regan A. Zane - for contributions to modular power conversion in energy harvesting, lighting, and electrified transportation systems

IEEE Technical Field Awards

2025 IEEE Andrew S. Grove Award

Awarded To: Sayeef Salahuddin
"For pioneering contributions to physics of ferroelectrics and integrated ferroelectric devices"

2025 IEEE Frederik Philips Award

Awarded To: Deirdre Hanford
"For visionary leadership in electronic design automation for secure and energy-efficient microelectronics"

Keynote Papers

Technical Program Chair: Meng-Fan (Marvin) Chang, National Tsing Hua University & TSMC

9:45 AM, Grand Ballroom B

1-1 **Electron Device Challenges to Enable Individual Centric Edge AI (Invited)**

Chidi Chidambaram, Senior Vice President and Fellow at Qualcomm

Although many pundits have forecasted the growth of edge AI device demand would far surpass the cloud [1] it has not happened yet. AI is widely used in the edge already, but the key killer applications have not evolved. The progress on cloud AI was enabled by AI acceleration with massively parallel compute and high bandwidth access to the memory. On the other hand, the edge device systems are more complex. Four key attributes of the consumer-oriented edge devices are low power consumption, robust connectivity, low cost and large manufacturing scale and these attributes make the edge AI implementations very challenging. Elegant integration and system level optimization of the core compute, AI acceleration, memory access, connectivity, sensing and power delivery blocks are key for edge devices to excel in each of these attributes. This talk will elucidate both the opportunities in front of the community to squeeze more from scaling/power reduction and the need to extend the bandwidth improvements from recent years into the edge devices to achieve the optimal edge system.

10:30 AM, Grand Ballroom B

1-2 Inside Enterprise Mission-Critical Computing (Invited)

Hillery Hunter, CTO, IBM Infrastructure and GM, Innovation; IBM Fellow at IBM

Behind the scenes, mainframes (IBM Z) and other enterprise servers like IBM Power Systems fuel banking, manufacturing, healthcare, insurance, and many other critical industries. While we each touch multiple systems with silicon daily, enterprise systems are the ones you never see but always rely on. In this talk, we'll look inside these systems, taking a technology-centric view to their evolution and discuss the specific silicon and packaging focus areas which come together to deliver eight nines of system-level resiliency (99.999999% uptime). We'll explore the changes in workloads on these systems and discuss how new drivers like AI for fraud detection, AI for insurance claims processing, and AI-driven system operations are changing enterprise system silicon and microarchitecture requirements.

11:15 AM, Grand Ballroom B

1-3 Essential Semiconductors - Innovation and Impact (Invited)

Ted Letavic, Senior Vice President, GlobalFoundries

Essential semiconductor technology comprises solutions which are optimized for performance, power, and reliability rather than transistor density alone. This segment represents the majority of the semiconductor non-memory total available market and plays a foundational role in enabling a wide range of applications which include communications and infrastructure, automotive and industrial systems, smart mobile devices, and autonomous systems and physical AI. Innovation trends for essential semiconductor technology discussed in this paper include silicon photonics, embedded non-volatile memory, advances in RF, embedded power, ultra-low power bulk silicon and silicon-on-insulator(SOI), and wide band gap semiconductor technology. Examples of the societal impact of essential semiconductor technology are presented – underpinning the importance of continued investment and innovation in this dynamic segment for decades to come.

Session 2: ALT | Gate-All-Around Technology, CFET and DTCO

1:30 PM, Grand Ballroom A
Co-Chairs: Chung-Hsun Lin, Intel and Changhwan Shin, Korea University
This session presents critical advances in gate-all-around (GAA) and 3D stacked transistor technologies as the semiconductor industry transitions beyond FinFET scaling limits at the 3nm node. The session covers eight papers addressing both near-term GAA implementation and future CFET architectures. Samsung demonstrates GAA technology's readiness for sustainable AI applications, emphasizing its role as the foundation for future logic architectures including forksheet and 3D-stacked FETs. IMEC presents multiple contributions: hybrid channel monolithic CFET with different crystal orientations for n/p-type optimization, aggressive cell height scaling through omega-gate forksheet structures from A7 to A3 nodes, comprehensive DTCO achieving 46% area reduction at iso-performance, and ultra-low temperature (<350°C) SiGe:B epitaxy achieving 100% performance gains through enhanced dopant activation. TSMC showcases circuit-level CFET maturity with the first functional ring oscillator and smallest SRAM bit-cells at sub-48nm pitch. IBM contributes two papers on critical scaling enablers: sub-100nm cell height library optimization with backside power delivery demonstrating viable performance at aggressive dimensions, and strained SiGe channel implementation delivering 17% current improvement through 34% channel resistance reduction. Together, these papers demonstrate how GAA serves as the pragmatic near-term solution maintaining compatibility with existing processes, while CFET technology with backside power delivery and heterogeneous integration capabilities provides the pathway for continued Moore's Law scaling through enhanced area efficiency and performance benefits.

1:35 PM
2-1 Gate-All-Around Technology for Sustainable AI: A Foundation for Future Logic Architectures (Invited)
Wookhyun Kwon[1], Jaehong Lee[2], Byounghak Hong[2], Myunggil Kang[2], Jinsu Jeong[2], Seunghwan Lee[2], Hidenobu Fukutome[2], Byung-Sung Kim[2], Ken Rim[2,2], Jaihyuk Song[2]
[1]Logic TD, Samsung, [2]Samsung
The transition from FinFET to Gate-All-Around (GAA) FET architectures is a necessary step driven by the physical limitations of FinFETs at advanced technology nodes, and GAA FETs offer superior electrostatic control, enabling steeper sub-threshold slopes and reduced leakage currents. GAA FETs provide additional design freedoms, such as continuous nanosheet width patterning, allowing for broader coverage across the speed-power spectrum and enabling the implementation of design technology co-optimization techniques to maximize chip design utility. The GAA architecture serves as a foundational platform for future innovations, including Forksheet, 3D-stacked FET, and their integration with backside power delivery networks, paving the way for energy-efficient computation and sustainable AI development.

2:00 PM
2-2 Hybrid Channel monolithic-CFET with Si (110) pMOS and (100) nMOS
Anne Vandooren[1], Serena Iacovo[1], Vincent Brissonneau[1], Thomas Chiarella[1], Pierre Eyben[1], Karen Stiers[1], Cassie Sheng[1], Camila Toledo de Carvalho Cavalcante[1], Dmitry Batuk[1], Maryam Hosseini[1], Andy Peng[1], Daisy Zhou[1], Sujan Sarkar Kumar[1], Andrea Mingardi[1], Rajendra Kumar Saroj[1], Subhobroto Choudhury[1], Rami Chukka[1], Violeta Georgieva[1], Roger Loo[1], Clement Porret[1], Thomas Dursap[1], Anjani Akula[1], Emmanuel Dupuy[1], Antony Peter[1], Nicolas Jourdan[1], Kevin Vandersmissen[1], Daniel Montero Alvarez[1], Evi Vrancken[1], Farid Sebaai[1], Pallavi Puttarame Gowda[1], Kenneth Lai[1], Venkat Sunil Kumar Channam[1], B.T. Chan[1], Alfonso Sepulveda Marquez[1], Robert Langer[1], Steven Brems[1], Efrain Altamirano Sanchez[1], Il.Gyo Koo[1], Katia Devriendt[1], Jerome Mitard[1], Sujith Subramanian[1], Conor Cullen[2], Daniel Casey[2], Gianluca Rengo[2], Anabela Veloso[1], Jishnu Ganguly[1], Ritam Sarkar[1], Naoto Horiguchi[1], Steven Demuynck[1], Lucas Petersen Barbosa Lima[1], Rami Khazaka[2], Serge Biesemans[1]
[1]imec, [2]ASM
This work reports on the demonstration of a monolithic-CFET (mCFET) integrating different channels for nMOS and pMOS. We enhanced our conventional mCFET process with a (110) Si pMOS - (100) Si nMOS combination and a single film embedded middle dielectric insulator (eMDI) by means of layer transfer. Using this advanced substrate,

these two new features can be integrated in any mCFET baseline with minimal modifications and reduced process complexity compared to a replacement MDI (rMDI) integration.

2:25 PM

2-3 3.5T CFET Block-Level DTCO for Superior PPA in A7 Node by Split Power, hDR Cells, Optimized Pins and BEOL

Ji-Yung Lin[1], Halil Kükner[1], Sheng Yang[1], Lynn Verschueren[1], Juergen Bömmels[1], Francesco Dell'Atti[1,2], Anita Farokhnejad[1], Maarten Van de Put[1], Odysseas Zografos[1], Naoto Horiguchi[1], Marie Garcia Bardon[1], Geert Hellings[1], Julien Ryckaert[1]
[1]imec, [2]KU Leuven

Complementary FETs (CFET), with the structure of stacked n-/p-FETs, hold promises for continuing shrinking device footprints after the nanosheet era. To realize CFET block-level designs with superior power-performance-area (PPA), several enhancements are introduced through design technology co-optimization (DTCO), such as a double-row split-power CFET structure, half-height double-row cells, as well as optimization in pins and back-end-of-line (BEOL). Results show that A7 3.5T CFET designs reach -46% area and iso-performance compared to N2 nanosheet designs.

3:15 PM

2-4 Block-Level Design Optimization for Sub-100-nm Cell Height Libraries with Stacked Transistor

Nick Lanzillo[1], Jim Mazza[1], Shahrukh Khan[1], Hyojong Shin[2], Jinyoung Lim[2], Young Gook Park[2], Utkarsh Bajpai[1], Janet Wilson[1], Henry Trombley[1], Jason Stuffle[1], Nick Perez[1], Susan Fan[1], Jingyun Zhang[1], Tenko Yamashita[1], Ghavam Shahidi[1], Dechao Guo[1], Kang-ill Seo[1], Dan Dechene[1], Ken Rim[2], Huiming Bu[1]
[1]IBM Research, [2]Samsung Electronics

We benchmark block-level power, performance and area scaling for a scaled architecture with 80 nm cell height based on stacked, stepped nanosheet transistors with backside power delivery. In addition to a base 5-track (5T) library, we evaluate several variant libraries, including high-density 4T, high-performance, double height, low-power and cross-couple (XC)-enabled. We demonstrate library optimization with these design variants enable attractive power, performance and area scaling for the Angstrom era of transistor scaling.

3:40 PM

2-5 First Demonstration of CFET Ring Oscillator and SRAM Bit-Cell Functionality at Gate Pitch Smaller Than 48 nm for Future Logic and SRAM Technology

Szuya Sandy Liao, Lu Yang, Tsung-Kai Chiu, Wei-Xiang You, Ting-Yun Wu, Yi-Che Lee, Wei-De Ho, Hsin Yang Hung, Kelvin Elphick, Zhi-Chang Frank Lin, June Fan James Hsu, Ku-Feng Yang, Chu-Hsuan Sha, Guan-Ren Wang, Chi-Ming Chang, Jia-Syuan Chen, Rui-Fu Chen, Szu-Hua Chen, Yung-Ta Chen, Yu-Hsien Chiang, Shih-Jung Ho, Tung-Wu Hsieh, Sheng-Yu Hsu, I Fan Hu, Kuan-Kan Hu, Jui-Chien Huang, Shao-Tse Huang, Jin-Hao Jhang, Ching Yen Lee, Yi-Hsuan Li, Cheng-Ming Lin, Chih-Pin Lin, Chun-Yu Liu, Hung-Kun Lo, Shao-Yu Lo, Ying-Chun Shen, Che Chi Shih, Chia Jung Tsai, Min-Chang Tsai, Chun-Yun Wang, Je-Ruei Wen, Wei-Yen Woon, Kai-Chieh Yang, Yuan-Chi Yang, Chien-Ning Yao, Chien-Lin Yu, Szu-Chia Huang, Chih-Liang Chen, Xiaomeng Chen, Fred Lai, Yao Ching Ku, S.S. Lin, L.C. Lu, Simon S.M. Jang, Min Cao
Taiwan Semiconductor Manufacturing Company

This study presents groundbreaking advancements in complementary field-effect transistor (CFET) technology, achieving two key milestones: the demonstration of the first fully functional 101-stage CFET ring oscillator (RO) and the development of the world's smallest operational SRAM bit-cell, offered in both high-density (HD) and high-current (HC) designs. Building upon our nanosheet-based monolithic CFET process architecture, this work further reduces the gate pitch to below 48 nm and introduces novel design features, including a nanosheet cut isolation (NCI) for effective isolation of neighboring devices within RO circuit and two versions of butted contact (BCT) local interconnects for cross-coupling inverters within a 6T SRAM bit-cell. These advancements signify a pivotal transition in CFET development, moving beyond device-level optimization to circuit-level integration.

4:05 PM

2-6 Multi-node Scaling Potential of Monolithic CFET

Sheng Yang, Lynn Verschueren, Jürgen Bömmels, Halil Kükner, Ji-Yung Lin, Fabian Bufler, Venkateswarlu Sankatali, Anita Farokhnejad, Maarten Van de Put, Geert Hellings
IMEC
This paper explores design strategies and Power-Performance-Area (PPA) trade-offs for scaling monolithic Complementary FETs (CFETs) from the A7 to A3 technology nodes. At A7 node, we demonstrate gate and Middle-of-Line (MOL) parasitic capacitance optimization to match N2 node performance. At A5 node, an Omega-gate Forksheet (O-FS) architecture keeps performance at the same level at lower sheet width. At A3 node, asymmetric crystal orientation engineering enhances drive current and supports aggressive area scaling targets.

4:30 PM
2-7 SiGe Channel for Scaled Gate-All-Around Nanosheet pFET Transistor for Advanced Logic Applications
Shogo Mochizuki, Sushant Kumar, Andrew Greene, Karsu Kilic, Xin Wang, Huimei Zhou, Andres Bryant, Juntao Li, James Demarest, Effendi Leobandung
IBM Research
Stacked Gate-All-Around (GAA) nanosheet p-type Field Effect Transistors (FETs) incorporating Si1-xGex channel have been successfully fabricated on scaled devices to investigate their electrical performance for next-generation logic applications. The $Si_{1-x}Ge_x$ NS channels demonstrate high crystalline quality and substantial compressive stress (0.6-1 Gpa), enabling enhanced carrier transport in the scaled pFET device with 15 nm gate length and 54 nm Contacted Poly Pitch (CPP). The influence of Ge fraction and junction overlap on the device characteristics of $Si_{1-x}Ge_x$ NS channel pFETs has been systematically analyzed to assess their impact on device scalability and transport characteristics. It is found that the $Si_{1-x}Ge_x$ NS channel provides a 17% uplift in I_{EFF} due to a corresponding channel resistance reduction of 34% while maintaining an excellent subthreshold slope of below 75 mV/dec.

4:55 PM
2-8 Junction-engineered Scaled High-performance GAA Nanosheet FETs with Ultra-low Temperature (< 350 °C) SiGe:B Source/Drain
Ritam Sarkar[1], Daniel Casey[2], Arka Dutta[1], Pierre Eyben[1], Andrea Pondini[1], Hans Mertens[1], Thomas Dursap[1], Clement Porret[1], Anabela Veloso[1], Jishnu Ganguly[1], Rutger Duflou[1], Conor Cullen[2], Parth Ashish Rathi[1], Min-Soo Kim[1], Rami Soo Khazaka[1], Jerome Soo Mitard[1], Lucas Petersen Barbosa Lima[1], Serge Biesemans[1], Naoto Horiguchi[1]
[1]IMEC, [2]ASM
We present an ultra-low-temperature (ULT) boron-doped SiGe (SiGe:B) epitaxial (epi) layer as PMOS junction in a gate-all-around (GAA) Si nanosheet (NS) transistor at 48 nm contacted poly-pitch (CPP) and 14 nm measured gate length (LG). We investigate the impact of dopant concentration and diffusion on NS performance at different RTA conditions. We find that the ULT junction (with S/D epi growth temperature <350 °C) with controlled RTA (at 800 °C) delivers over 100% improvement in performance (ID,LIN , gm,LIN) over our reference epitaxy process (at 500 °C) by significantly increasing the active dopant concentration and carefully position the junction under the inner spacer without degrading the short-channel effects (SCE). Moreover, contact resistivity reduces by 3.5X compared to reference epi process.

Session 3: MT | Embedded Memory and Selector Only Memory

1:30 PM, Grand Ballroom B
Co-Chairs: Mattia Boniardi, STMicroelectronics and Shimeng Yu, Georgia Institute of Technology
This session features three embedded papers and three SOM papers. The first paper by Fraunhofer/GlobalFoundries demonstrates a high scaled FRAM using Al co-doped $Hf_xZr_{1-x}O_2$ (HZAO). It offers low leakage and stable data retention at 150 °C, verified for 1 h. The second paper from Macronix presents an innovative architecture for 3D Stackable Split-Gate 1.5T NOR Flash, achieves 120Mb/mm^2 for 32-layer 1.5T 3D NOR embedded flash applications that can perform very low energy 0.2pJ/bit. The third paper from STMicroelectronics/CEA-Leti/IBM Research explores in-memory computing utilizing 28nm FD-SOI Ge-GST technology for edge AI. By using a Rheostatic cell architecture, the performance can be improved. The fourth contribution is from POSTECH/Kyungpook National University and focuses on SOM, reporting on a multi-scale simulation framework, bridging atomic scale Density Functional Theory (DFT) calculations with device-scale stochastic dynamics in In–Te OTS systems. The fifth talk is from POSTECH too and proposes an innovative hybrid

SOM architecture that separates memory and selector functionalities within a single nanoscale device, addressing the inherent trade-off between a memory window (MW) and other key characteristics such as threshold voltage drift and variability. Micron Technology closes the session and focuses on the Single-chalcogenide X-point Memory (SXM) technology, showing a robust scheme to increase memory density from two levels per cell to three levels per cell. Further, it demonstrates MLC capability (2bit/cell) on large statistics.

1:35 PM

3-1 Highly Scaled FRAM Technology in 22FDX→ Achieving Sub 1 V Operation

Franz Müller[1], Vadym Iurchuk[2], Alison Viegas[1], Nandakishor Yadav[1], Jacob Syndikus[1], Robert Seidel[2], David Lehninger[1], Raik Hoffmann[1], Ricardo Olivo[1], Johannes Mutter[2], Halid Mulaosmanovic[2], Konrad Seidel[1], Thomas Kämpfe[1], Johannes Müller[2], Steven Soss[2], Sven Beyer[2], Maximilian Lederer[1]
[1]Fraunhofer IPMS, [2]GlobalFoundries

We report on the successful integration of a ferroelectric random-access memory (FRAM) in GlobalFoundries 22FDX® node. The ferroelectric capacitor (FeCap) of the 1T1C-bitcell is located in the Back-End-of-Line (BEoL) between M3 and M4. Using Al co-doped $Hf_xZr_{1-x}O_2$ (HZAO), remanent polarization ($2P_R$) above 40 µC/cm^2 and coercive field ($2E_C$) of 1.4 MV/cm are achieved for scaled FeCaps of 0.01-0.04 µm^2. Compared to state-of-the-art HZO, the lower leakage of HZAO results in improved data retention, even at elevated temperatures of 150°C for 1h. The FRAMs are combined in segments of 2 bitlines (BL, BL_{bar}) with a latch sense amplifier (LSA). In the segments sub-1V switching at 8nm HZAO film thickness is demonstrated. Operating at switching voltages of 1.5V, endurance cycling of >10^{12} is shown. Sub-E_C operation and maintained charge in the FRAM node enables subsequent non-destructive readout capability.

2:00 PM

3-2 Architecture Design and Simulation of a Novel 3D Stackable Split-Gate 1.5T NOR Flash for Gb-Density Embedded Flash Chiplets with Low Latency (<30 ns) and Low Power Sensing (<0.2 pJ/bit)

Hang-Ting Lue[1], Weichen Chen[2], Teng-Hao Yeh[2], Tzu-Hsuan Hsu[2], Keh-Chung Wang[2], Chih-Yuan Lu[2]
[1]Emerging Central Lab., Macronix, [2]Macronix

We propose a novel 3D stackable split-gate (1.5T) NOR Flash memory. This work builds on our previously developed 32-layer 3D NOR 4Gb chip parameters, which include device dimensions, poly-Si channel, array R/C parameters for the 16Mb tile, and SPICE models for peripheral transistors. The 1.5T split-gate 3D NOR Flash offers advantages in faster read latency. By adopting Cu hybrid bonding technology to connect the 1.5T 3D NOR wafer to an advanced CMOS logic wafer, we can position sense amplifiers (SAs) near each 16Mb tile. This configuration enables low-latency random access reads (<30 ns) while consuming very low power (<0.2 pJ/bit).

2:25 PM

3-3 A Novel A-IMC Ge-GST ePCM Cell for Edge AI Applications on 28nm FD-SOI Platform

Matteo Baldo[1], Mario Allegra[1], Mattia Boniardi[1], Luca Scotti[1], Jeremie Jasse[1], Riccardo Zurla[1], Jacopo Bertolini[1], Emanuela Calvetti[1], Marco Pasotti[1], Christian Boccaccio[1], Latifa Desvoivres[1,2], Yannick Le-Friec[1], Alain Ostrovsky[1], Pascal Gouraud[1], Simon Jeannot[1], Laurent Favennec[1], Jesse Luchtenveld[3], Athanasios Vasilopoulos[3], Prasad Jonnalagadda[3], Ghazi Syed[3], Abu Sebastian[3], Andrea Redaelli[1]
[1]STMicroelectronics, [2]CEA-Leti, [3]IBM Research - Zurich

A new Ge-GST embedded PCM cell architecture, optimized for Deep Neural Network (DNN) acceleration, is presented. Its integration flow is discussed, and its dimensions are optimized through TCAD simulations. Extensive electrical characterization of conductance levels is performed to assess analog programming and level stability over time and temperature. Simulations of large DNN demonstrate superior performance of the optimized cell when compared with standard Wall structure ePCM. Finally, functionality on a large statistic is demonstrated with the validation of a test-vehicle showing good process yield. Silicon implementation of a NN proves the excellent A-IMC characteristics of this architecture.

3:15 PM

3-4 Revealing Switching Mechanisms and Design Strategies for Selector-Only Memory Devices through Integrated DFT-Stochastic Modeling

Dongmin Kim[1], Yoori Seo[2], Ohhyuk Kwon[2], Yubin Park[3], Tae Hoon Lee[3], Hyunsang Hwang[2]
[1]Materials Science and Enegineering, POSTECH, [2]POSTECH, [3]Kyungpook National University
In this work, we present a multi-scale framework that, for the first time in Ovonic Threshold Switch (OTS), provides a direct physical basis for the locally stochastic switching mechanism by organically integrating atomic-scale Density Functional Theory (DFT) with device-level stochastic dynamics. Our DFT-informed Stochastically Induced Defect Percolation (SIDP) model consistently explains fundamental threshold switching, intrinsic variability, and memory operations. Experimental validation with In-Te OTS devices shows excellent agreement with our model. We explore a multilayer approach that assigns memory functionality to the interface region and stable switching behavior to the bulk and also analyses scaling impacts on both switching and memory performance. Our simulation reveals an ideal performance memory window (MW = 4.34 V), read window margin (RWM = 2.83 V), providing practical guidelines for the future development of high-performance OTS devices.

3:40 PM
3-5 Hybrid-Stacked Sb-based Selector-Only Memory with Large Memory Window (> 1.3 V) and Excellent Endurance (> 10^{10}) via Decoupled Memory and Selector Functions
Jangseop Lee, Yoori Seo, Ohhyuk Kwon, Laeyong Jung, Hyunsang Hwang
Pohang University of Science and Technology (POSTECH)
In this study, we proposed an innovative hybrid selector-only memory (SOM) architecture that strategically separates memory and selector functionalities within a single nanoscale device. By combining chalcogenide layers with different materials, our design addresses the inherent trade-off in conventional SOMs between a memory window (MW) and other key characteristics such as threshold voltage drift and variability. Detailed optical and electrical analyses reveal that memory and selector functions require fundamentally distinct bonding environments. Based on these results, our hybrid stack integrates a high-distortion Sb-InTe layer for memory switching and a stable InTe layer for selector operation. The optimized hybrid device exhibited an enhanced MW (> 1.3 V), excellent endurance (> 10^{10}), and ultra-low write current (I_{write} ~ 10 μA), confirming its potential for an energy-efficient memory device.

4:05 PM
3-6 Multilevel Single Chalcogenide X-point Memory
Innocenzo Tortorelli, Mattia Robustelli, Hernan Castro, Jeremy Hirst, Fabio Pellizzer, Agostino Pirovano, Nevil Gajera, Russ Meyer
Micron
The new Single-chalcogenide X-point Memory (SXM) has been proposed as a potential candidate for next generation high-density storage class memories thanks to superior performances and cell stack simplification with respect to 3DXP. To keep cost competitiveness in terms of $ per bit in the 3D NAND ballpark, a suitable scaling roadmap together with the possibility to store more than one bit per cell seems mandatory. This is the first time that multilevel programming has been demonstrated on large statistics on SXM technology paving the way to enlarge application space to high performance storage products.

Session 4: RSD | Ferroelectric Memory Reliability

1:30 PM, Continental 1-3
Co-Chairs: Motoyuki Sato, Tokyo Electron and Kai Ni, University of Notre Dame
The first paper investigates the mechanisms underlying access-rate-dependent degradation in ferroelectric memory and proposes operation–device co-optimization together with hybrid polar recovery strategies, achieving a 100× improvement in endurance and robust polarization recovery. The second and third papers address the reliability of ultra-thin ferroelectric HfO_2. The second paper focuses on pulse-voltage engineering combined with interfacial and thermal preconditioning, demonstrating fatigue-free endurance beyond 1.2×10^{12} cycles at 1 V. In contrast, the third paper emphasizes process optimization and theoretical insights to enhance ferroelectric performance, reporting fully ALD-deposited 9 nm MFM stacks with 3 nm HZO and TiN electrodes that exhibit excellent performance. The fourth and fifth papers investigate imprint-related degradation in HfO_2 ferroelectric capacitors. The fourth paper examines retention-loss mechanisms and corresponding mitigation strategies, revealing that imprint and charge-assisted oxygen vacancy migration drive the degradation. The fifth paper

advances the understanding of intrinsic imprint, universal de-imprint, and AC-enhanced imprint effects through a unified oxygen-vacancy dynamics model, supported by comprehensive DC/AC measurements and DFT calculations. The sixth and seventh papers explore double-gate FeFETs for memory applications. The sixth paper experimentally demonstrates a double-gate nanosheet FeFET, showing that separate read/write paths reduce variability and mitigate subthreshold swing degradation, making it a strong candidate for embedded memory. Meanwhile, the seventh paper leverages the disturb-free characteristics of double-gate FeFETs for vertical NAND. It addresses challenges of increased leakage and reduced ON current through a combined experimental and simulation-based study. Finally, the eighth paper proposes a method to decouple and quantify polarization and trapped charges in FeFETs with gate-side injection for vertical NAND applications. It reveals distinct charge dynamics during program and erase states, identifies reliability challenges, and maps charge-loss pathways, providing guidance for improved FeNAND memory design.

1:35 PM

4-1 First Experimental Insights into Access-Rate-Dependent Ferroelectric Memory Degradation: Mechanisms and Mitigation through Hybrid Polar Recovery and Operation-Device Co-Optimization

Cheng-Hung Wu[1], Tszying Pun[2], Jia-Tai Wu[1], Chih-Wei Sung[3], Yi-Ming Tseng[1], Po-Lun Lin[2], Wei-Tung Chen[1], Peng-Wei Chu[3], Chien-Ting Wu[4], Chun-Jung Su[2,4], Vita Pi-Ho Hu[1]
[1]National Taiwan University, [2]National Yang Ming Chiao Tung University, [3]National Tsing Hua University, [4]Taiwan Semiconductor Research Institute

We present the first experimental decoupling and quantification of oxygen vacancy (V_o) migration and generation mechanisms in access-rate (AR)-dependent degradation of ferroelectric (FE) memory. Cryogenic analysis reveals V_o migration as the dominant degradation path under slow AR conditions. To address this, we propose a robust strategy combining Nb electrodes and an AFE-FE bilayer HZO stack to suppress V_o activity, along with asymmetric write operations and a hybrid polar recovery scheme. This operation-device co-optimization (ODCO) approach achieves a 100× improvement in endurance under realistic random AR scenarios. Beyond unveiling the mechanistic origin of AR-dependent reliability loss, this work provides a practical and time-efficient pathway toward robust FE memory operation.

2:00 PM

4-2 Fatigue-Free Ferroelectric HZO Operable at 1V with >10^{12} Endurance via Precycling Pulse Engineering

Kyungmee Song[1], Donghyun Kim[2], Donghoon Kim[2], Sangjun Lee[2], Hyun Hwi Lee[3], Seunggeol Nam[2], Sunghyun Kim[2], Jooho Lee[2], Sijung Yoo[2], Yoonsang Park[2], Seung Min Lee[2], Kyunghwan Lee[2], Intak Jeon[2], Chang Hwa Jung[2], Daewon Ha[2], HanJin Lim[2], Jinseong Heo[2], Kyooho Jung[2], Duk-Hyun Choe[2]
[1]Samsung Advanced Institute of Technology, Samsung Electronics, [2]Samsung Electronics, [3]POSTECH

Overcoming polarization fatigue while enabling low-voltage operation remains a key challenge in HfO_2-based ferroelectric devices. Here, we introduce an asymmetric pulse
voltage engineering approach, combined with interfacial and thermal preconditioning, to achieve both fatigue immunity and robust ferroelectricity in ultrathin $Hf_{1-x}Zr_xO_2$ (HZO) capacitors. Our device exhibits >1.2×10^{12} endurance cycles (limited by measurement time) with excellent $2P_r$ stability (<5% variation) under 1V operation. Moreover, high-temperature wake-up further boosts remnant polarization (P_r) while preserving fatigue-free characteristics. This represents the first demonstration of fatigue-free, high-endurance operation in sub-5 nm HZO at 1V, offering a promising pathway for low power ferroelectric memories.

2:25 PM

4-3 Sub-10 nm Fully Atomic-Layer-Deposited MFM Stack with Enhanced Reliability: Record-low $I_{leakage}$ and Record-high E_{BD} in 3 nm Ferroelectric HZO

Xiaolin Wang[1], Yang Feng[1,2], Jiezhi Chen[2], Xiao Gong[1,3]
[1]National University of Singapore, [2]Shandong University, [3]A*STAR

We demonstrate fully ALD-deposited metal-ferroelectric-metal (MFM) capacitors featuring the smallest total thickness of 9 nm. Our ultra-thin devices not only achieve a record-low coercive voltage ($2V_c$ <0.65 V), but also set new benchmarks in several key reliability metrics for 3 nm ferroelectric (FE) layers: an ultralow leakage current density (<10^{-3} A/cm^2 at 3 MV/cm), an ultrahigh breakdown electric field (E_{BD}) exceeding 9.5 MV/cm, and exceptional

endurance exceeding 10^{11} cycles without ferroelectricity failure (final $2P_r$ > 18 mC/mm^2 and no electrical breakdown). Our comprehensive characterization and analysis reveal that these breakthroughs are driven by two critical performance-enhancing benefits unlocked by thinning TiN electrodes below 8 nm: (1) the transition from polycrystalline to amorphous TiN promotes o-phase formation in HZO, allowing simultaneous downscaling of both HZO and TiN layers; and (2) smoother TiN surfaces and suppressed oxygen scavenging collectively minimize leakage current and enhance breakdown strength.

2:50 PM

4-4 Imprint-centric Retention Loss in HfO_2-based Ferroelectric Thin Films: Charge-assisted Defect Migration and Trap Neutralization Strategy

Xiaopeng Li, Xiaoyu Dou, Yaoyu He, Yilin Hou, Pengpeng Sang, Xuepeng Zhan, Jixuan Wu, Jiezhi Chen
Shandong University

Retention loss in HfO_2-based ferroelectric (FE) thin films is investigated systematically through wide-temperature characterizations. For the first time, it is observed that retention loss has an extremely strong correlation with imprint, which can be well explained by temperature-activated imprint-centric (TI) model. Based on first-principles calculations, it is revealed that charge-assisted defect migration and switching (CDMS) could be the intrinsic mechanism, that is, imprint originates from V_o^{2+} migration to the interface while localized depolarization is caused by V_o^{2+}-induced switching barrier reduction. To mitigate this, a defect-engineering with trap neutralization scheme (DTNS) is proposed. It is demonstrated that, by adopting process optimization to lower V_o and electron pre-injection to neutralize V_o^{2+}, ~28% imprint reduction and ~21% retention enhancement can be achieved. Our results shed light on retention optimization strategies in HfO_2-based FE devices.

3:40 PM

4-5 New Insights into the Imprint Effects in Hf-based Ferroelectric Capacitors: Intrinsic Imprint, Universal De-Imprint, AC Enhanced Imprint, and Unified Modeling of Defect Dynamics

Hao Li[1], Keying Sun[2], Chen Zhang[3], Haoran Wang[2], Yue-Yang Liu[3], Zongwei Shang[1], Ming Li[1], Lang Zeng[2], Huixiong Deng[3], Runsheng Wang[1]
[1]Peking University, [2]Beihang University, [3]Chinese Academy of Sciences

In this paper, by comprehensive DC and AC characterizations with the proposed subtraction methods that decouple various polarization components, we successfully reveal the underlying physics of imprint effects with a unified manner in Hf-based ferroelectric capacitors. It is found that the imprint is an intrinsic result regardless of full or partial polarization conditions. Then, for the first time, the universal de-imprint behavior and the AC-enhanced imprint are observed and analyzed in detail. It is also found that the split-up I-V curve which previously was taken as fatigue, is merely the result of AC-enhanced imprint. Assisted by DFT calculations, the origin of imprint is attributed to asymmetrical double well of interface V_O. Therefore, the imprint, de-imprint and AC-enhanced imprint can be quantitively modeled within a unified framework of V_O dynamics, which can be coupled with NLS model. The simulation results agree well with all the experiments, for the first time.

4:05 PM

4-6 3D-Stackable Double-Gate Nanosheet (DGNS) FeFETs with Separate Read/Write Paths: Suppressed Variability with Sub-2 % Bit Error Rate and Mitigated Subthreshold Swing Degradation

Fang Wu[1], Chun-Yung Chiu[2], Cheng-Hung Wu[3], Yueh-Lin Tsai[2], Shih-Hang Chen[2], Horng-Chih Lin[2], Vita Pi-Ho Hu[3], Chun-Jung Su[2,2], Pin Su[2]
[1]Institue of Electronics, National Yang Ming Chiao Tung University, [2]National Yang Ming Chiao Tung University, [3]National Taiwan University

We present the first experimental validation of a double-gate nanosheet (DGNS) FeFET architecture that suppresses variability and enhances reliability through read/write path separation. By assigning the top-gate (TG) for programming and the side-gate (SG) for sensing, the architecture minimizes read disturbance, mitigates subthreshold swing (SS) degradation under endurance, and improves immunity to ferroelectric phase non-uniformity. Compared to TG-only sensing, this design reduces the bit error rate from 25.7 % to 1.78 % and suppresses device-to-device (D2D) and cycle-to-cycle (C2C) variability by over 22 %. This work fills a critical experimental gap left by prior simulation-based studies on DG FeFET variability. The DGNS scheme further

supports 3D stacking and low-power operation. These improvements demonstrate that architectural decoupling and interface control, rather than merely intrinsic ferroelectric (FE) properties, are key to achieving scalable and reliable FeFETs for non-volatile memory and logic-in-memory applications.

4:30 PM

4-7 Comprehensive Evaluation of Challenges in Realizing Disturb-Free Ferroelectric Vertical NAND via String-Compatible Independent Pass Gates

Xuezhong Niu[1], Lance Fernandes[2], Zijian Zhao[1], Moonyoung Jung[3], Kijoon Kim[3], Suhwan Lim[3], Kwangsoo Kim[3], Wanki Kim[3], Daewon Ha[3], Halid Mulaosmanovic[4], Stefan Duenkel[4], Sven Beyer[4], Vijaykrishnan Narayanan[5], Asif Khan[2], Kai Ni[1]

[1]University of Notre Dame, [2]Georgia Institute of Technology, [3]Samsung Electronics, [4]GlobalFoundries Fab1 LLC Company KG, [5]Pennsylvania State University

This work presents comprehensive experimental and modeling studies of our dual-port FeFET architecture for pass disturb-free vertical NAND storage. The architecture incorporates a string-compatible independent pass gate dedicated for pass operation rather than relying on the write gate. Key findings include: i) The non-ferroelectric pass gate ensures pass disturb immunity through its geometry—pass voltage aids HVT state retention while adverse effects on LVT state are screened by the channel; ii) A critical challenge is increased HVT leakage current from weak back-channel regions not fully depleted by the front gate, verified through TCAD modeling and experimental demonstration; iii) We explore pass gate oxide and channel thickness design space via TCAD simulations and experimentally validate modeling using dual-port gate-side injection FeFETs; iv) The implemented design demonstrates pass disturb-free operation while managing leakage trade-offs; v) Reduced LVT ON-current when unselected cells remain in HVT states can be mitigated by proper pass gate biasing.

4:55 PM

4-8 Decoupling and Quantifying Polarization and Trap Charges in MIFIS FeNAND: Insights into Charge Dynamics and Reliability

Hongrae Joh[1], Hyunjun Kang[2], Junhyeok Kwak[2], Giuk Kim[2], Hyojun Choi[2], Hoon Kim[2], Sanghyun Park[3], Kwangyou Seo[3], Kwangsoo Kim[3], Wanki Kim[3], Daewon Ha[3], Jinho Ahn[4], Sanghun Jeon[2]

[1]School of Electrical Engineering, Korea Advanced Institute of Science and Technology, [2]Korea Advanced Institute of Science and Technology, [3]Samsung Electronics Co., Ltd, [4]Hanyang University

We propose novel measurement methodology to decouple and quantify spontaneous polarization (ΔP_S), channel-injected (ΔQ_{it}) &gate-injected trap charge ($\Delta Q_{it}'$) in metal-insulator-ferroelectric-insulator-semiconductor FeNAND devices. We extracted charge dynamics with delay time, separating contributions from charge components. Using time- and bias-dependent measurements, distinct memory window (MW) behaviors were identified: read-after-write delay dominates MW degradation in PGM state, while charge loss governs ERS state. Additionally, we confirmed that FE and Ch.IL act as primary tunneling paths. Finally, endurance analysis reveals significant reduction in ΔP_S, while ΔQ_{it}&$\Delta Q_{it}'$ remain stable, indicating increasing role of trap charges in MW control with cycling. This framework enables design of FeNAND with improved reliability characteristics.

Session 5: NC | LLM and Search Engines with In-Memory Compute

1:30 PM, Continental 4

Co-Chairs: Valeria Bragaglia, IBM Research and Christopher Bennett, Sandia National Laboratories

This session includes 7 papers that focus on novel in-memory compute systems for Large Language Models (LLM) and search engines accelerators. In the first paper by Feixiong Wang et al., from the Chinese Academy of Science, the authors demonstrate experimentally energy efficient object detection by combining digital and analog emerging devices. The second paper by Yu-Yu Lin et al., from Macronix, explores a novel 3D-NAND-based mixed mode in Storage-Computing architecture and demonstrates the execution of a compact LLM. In the third paper by Won-Tae Koo et al., from SK Hynix, the authors present a holistic co-optimization of a 3D ferroelectric NAND for efficient analog compute. The fourth paper by Weikai Xu et al., from Peking University, presents a 3D integrated ferroelectric capacitor implementing efficient reads and writes of values from attention layers. In the fifth paper by Omkar Phadke, et al., from Georgia Tech, the authors implement ternary content addressable memory arrays using back-end of the line ferroelectric capacitor and show that the system can implement genomic search. In the sixth

paper by Rui Qiu et al., from Hong Kong University, the authors show that the integration of two different memory technologies enables energy efficient transformer inference. Finally, in the seventh paper by Mingcheng Shi et al., from Tsinghua University, a parallel read mechanism is implemented for accelerating attention layers using ferroelectric technology.

1:35 PM

5-1 First demonstration of Monolithic 3D SiSiC: A Hybrid Digital-Analog CIM with Vertically Stacked Dual-Tier Si-FinFETs and CNT-FETs for Intelligent Vision Processing

Feixiong Wang[1], Yadong Zhang[1], Ze Wang[2], Shuang Liu[1], Yanqing Li[1], Zhaohao Zhang[1], Yanzhao Wei[1], Zhiyao Wang[1], Yu Wang[1], Hong Yang[1], Junfeng Li[1], Yuankun Li[2,2], Mingcheng Shi[2], Yanbo Su[2], Song Qiu[3], Heyi Huang[1], Jianshi Tang[2], Huaxiang Yin[1]

[1]Institute of Microelectronics, Chinese Academy of Sciences, [2]School of Integrated Circuits, Beijing Innovation Center for Future Chips, BNRist, Tsinghua University, [3]Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences

We present the first experimental demonstration of vertically stacked dual-Tier Si FinFETs for computing-in-memory, integrated with a CNTFET array, enabling intelligent vision processing in the M3D-SiSiC architecture. On 1st-Tier of standard Si-FinFETs, the 2nd-Tier was formed by low-thermal-budget Si-FinFETs with back-end-of-line compatible Schottky Barrier techniques for basic CMOS logic functions. Specially, through unconventional VDS bias approach, the Low-TB Si-FinFETs exhibits reliable memristive behavior both in P/NMOS configurations, demonstrating fast set/reset speed (40 ns), and scalability down to 28 nm Lg. Furthermore, we proposed a hybrid digital-analog CIM system (FinFET CMOS circuits for DMAC + Low-TB Si-FinFET for analog CIM), stacking with 3rd-Tier CNT-FET sensor array, to enhance intelligent vision processing. It achieves a 45.7% improvement in energy efficiency over DMAC-based NPUs while maintaining high precision. This work indicated the innovated M3D-SiSiC system with advanced nodes and a cost-effective fabrication process, offer a promising solution for high-performance AI systems integration.

2:00 PM

5-2 3D NAND-based Mixed-mode in-Storage-Computing Architecture for Large Language Model

Yu-Yu Lin, Feng-Min Lee, Ming-Hsiu Lee, Keh-Chung Wang, Chih-Yuan Lu

Macronix International Co., Ltd.

This paper explores a novel 3D-NAND-based mixed-mode in-Storage-Computing architecture, which provides a promising solution to address AI's growing computing demands. The iSC architecture leverages NAND Flash's cost-effectiveness and high parallelism to enable high-performance, energy-efficient vector-matrix multiplication operations. The Qwen 2.5 Math 7B model is used as a case study to explain the weight and input quantization, and an iSC + ALU hybrid system for matrix computations. The effects of iSC-induced errors and the accuracy of the iSC-integrated system are assessed with the MATH500 dataset, showcasing the potential of the proposed mixed-mode iSC architecture as a transformative technology for AI applications.

2:25 PM

5-3 Co-Optimizing Cell, Non-Cell, and Page Schemes for Energy Efficient Analog Computing in 3D FeNAND

Won-Tae Koo, Jihun Kim, Jae-Gil Lee, Seungyeol Oh, Hyung Dong Lee, Kyusung Kim, Joongsik Kim, Soo Gil Kim, Seho Lee, Jaeyun Yi, Youngman Cho, Seon Yong Cha

SK hynix Inc.

This work presents a holistic co-optimization of 3D ferroelectric NAND (FeNAND) for energy-efficient and high-throughput analog computation-in-memory. The co-optimization covers cell properties, non-cell peripherals, and computational schemes. Our approach enhances multi-level capability through write algorithm tuning, reduces word-line transition power, and improves analog multiply-and-accumulate efficiency via in-page pipelining and splitting. As a result, computational throughput and energy efficiency of 3D FeNAND are improved by up to 16× and 4950×, respectively, compared to conventional 2D arrays.

2:50 PM

5-4 First Experimental Demonstration of Disturb-Free 3D Vertical 1T-nC-1T Ferroelectric-based KV Cache with Co-Optimization of Hybrid Analog-Digital CIM and Token-Wise Dynamic Pruning for Efficient Long-Context LLM Inference

Weikai Xu, Danyun Luo, Minyue Deng, Shengjie Cao, Shuzhang Zhong, Meng Li, Qianqian Huang, Ru Huang
Peking University

Through device-architecture-algorithm co-optimization, a novel 3D vertical 1T-nC-1T FeRAM structure for KV cache, featuring orthogonally aligned word-lines/bit-lines and shared FeCap strings, as well as a hybrid analog-digital CIM architecture with token-wise dynamic pruning algorithm, are proposed and experimentally demonstrated. The designed 1T-nC-ACIM with small-signal non-destructive read can evaluate similarity for efficient token selection, and the nC-1T-DCIM with robust destructive read supports accurate attention computation with a subset of selected KV cache, enabling efficient and accurate KV cache workloads. Besides, by optimizing data mapping scheme, parallel computation across nC strings is enabled, leading to enhanced throughput and avoided write disturbance. Based on the above design, a 3D 3×32×32 FeCap array is experimentally fabricated with ~10ns switching, 10-year retention, 10^{16} endurance and good consistency, and KV cache-based attention is demonstrated with 6×/315× improved performance and energy-efficiency over the state-of-the-art designs, along with high accuracy, showing its great potential for long-context LLM inference.

3:40 PM

5-5 Array-Level Demonstration of Charge-Domain In-Memory Search using Back-End-of-Line Compatible Ferroelectric Nonvolatile Capacitors

Omkar Phadke, Junmo Lee, Po-Kai Hsu, Chengyang Zhang, Asif Khan, Suman Datta, Shimeng Yu
Georgia Institute of Technology

In this work, we experimentally demonstrate a charge-domain Ternary Content Addressable Memory (TCAM) array using back-end-of-line (BEOL) compatible ferroelectric non-volatile capacitors (nvCAPs) for in-memory search. The geometrically optimized nvCAP provides: 1) reduced static power consumption due to charge-domain compute with charge-transfer mechanism, 2) non-destructive readout with a 100mV pulse, 3) >1000s (or 10^{12} equivalent read cycles) read-stress resilience and retention @85°C temperature. The pattern matching measurements and linear dependence of output response of TCAM array on the level of similarity between stored and searched pattern are conducted for the functional verification of TCAM array. The proposed TCAM is evaluated for bioinformatic workloads in genome sequencing by performing SPICE and system-level simulation. The benchmarking for scaled nvCAP devices exhibits ~40× improvement in search energy-delay-product over the conventional 16T CMOS TCAM implementations.

4:05 PM

5-6 Monolithic 3D Integration of MoS_2 eDRAM and RRAM for Analog In-Memory Attention Computing

Rui Qiu[1], Guoyun Gao[1], Zhiyuan Du[1], Ni Yang[1], Zefan Li[1], Dongsheng Zuo[2], Yunwei Tong[1], Peiyi He[1], Hang-Ming ZHANG[1], Hao Jiang[3], Lain-Jong Li[4], Han Wang[1], Can Li[1]
[1]The University of Hong Kong, [2]The Hong Kong University of Science and Technology, [3]Fudan University, [4]National University of Singapore

We demonstrate a monolithically 3D compute-in-memory chip that accelerates Transformer attention by vertically integrating TaO_x RRAM and MoS_2 eDRAM layers on Si-CMOS. The BEOL-fabricated 2T0C MoS_2 eDRAM exhibits sub-nanosecond write latency and > 100 s retention, significantly outperforms AOS-based cells, overcoming write bottlenecks and eliminating frequent memory refresh. This M3D architecture provides massive interlayer bandwidth, eliminating data movement bottlenecks between static (RRAM) and dynamic (eDRAM) CIM with SRAM. Hardware-calibrated simulations show a 500× speedup over GPU and an 85% memory area reduction over 2D-based CIM, while maintaining high accuracy on language benchmarks. Our work demonstrates a scalable path for efficient LLM acceleration.

4:30 PM

5-7 High-Throughput Monolithic 3D Multi-bit Vertical 2TnF Ferroelectric Gain Cells for Computing-in-Memory to Accelerate Attention Mechanism in Transformer

Mingcheng Shi[1], Yi Ding[2], Yuyan Wang[2], Bowen Shen[2], Benjamin Yang[2], Ruofei Hu[2], Yiming Zhou[2], Jiaming Li[2], Wen Sun[2], Yuankun Li[2], Eunhye Baek[2], Heyi Huang[3], Qingtian Zhang[2], Bin Gao[2], He Qian[2], Jianshi Tang[2], Huaqiang Wu[2]
[1]School of IC, Tsinghua University, [2]Tsinghua University, [3]Chinese Academy of Sciences
For the first time, we experimentally demonstrate a monolithic 3D (M3D) integrated $Hf_{0.5}Zr_{0.5}O_2$ (HZO)-based 2TnF ferroelectric gain cell (Fe-GC), featuring vertical transistors and stackable storage nodes (SNs) for parallel operations. A co-designed parallel read computing-in-memory scheme is proposed, aiming to accelerate the attention mechanism in Transformer. The write transistor employs selfoxidized tantalum gate to form its gate oxide, while the read transistor showcases a vertical top-gate IGZO FeFET with quasi-destructive read capability for the frequently accessed K matrix in attention. Utilizing the same ferroelectric film, highdensity multi-bit SNs constructed on the sidewalls enable efficient one-time buffering of high-volume Q matrices. This co-optimization achieves high throughput and reduced read cost by minimizing data movement for dynamic matrix-vector multiplication. Array-level characterizations validate the proper function of 2TnF Fe-GC, highlighting its potential as a high-throughput KV cache for M3D integration. Performance benchmarks indicate a 13.0× overall performance improvement over conventional planar DRAM architectures.

Session 6: MT/ALT/NC | Focus Session | Efficient AI Solutions: Advances in Architecture, Circuit, Device and 3D Integration

1:30 PM, Continental 5
Co-Chairs: Katherine H. Chiang, TSMC and Seyoung Kim, POSTECH
The rapid advancements in AI have ushered in a transformative era, essentially reshaping all aspects of our lives. While progress in models and algorithms has garnered significant attention, hardware innovations are equally critical. Advances in custom silicon, high-bandwidth memory, storage technologies, and energy-efficient architectures are redefining the capabilities of AI systems. This synergy between software and hardware continues to propel the AI revolution forward. In this focus session, we will begin with a system-level overview to provide a comprehensive perspective before delving into the critical technologies that underpin AI advancements. Given the data-intensive nature of AI computation, innovations in memory technologies are becoming increasingly vital. Among these, oxide semiconductors have emerged as a promising solution due to their unique ability to meet the growing demands for power efficiency, bandwidth, and scalability to 3D architectures. Finally, we will conclude the session by exploring material innovations that are essential to driving future breakthroughs in AI hardware.

1:35 PM
6-1 Architecting for Power-Efficiency and Performance: A CPU Architect's Perspective (Invited)
Debbie Marr, AheadComputing
As process technology advances, the promise of performance and transistor density improvements is increasingly constrained by power and thermal limits. While innovations in materials, interconnects, and device structures remain essential, they must now be tightly coupled with architectural strategies to fully realize system-level efficiency. At the same time, the explosion of computational demands in AI has outpaced traditional scaling curves, intensifying pressure on both architecture and process technology to deliver unprecedented performance within tight power and thermal envelopes. This paper highlights how microarchitecture and process technology co-design can address escalating thermal density, power consumption challenges, and performance needs — urging process researchers to consider architectural implications in their scaling roadmaps.

2:00 PM
6-2 Enabling Efficient AI and HPC through Specialization and Advanced Packaging (Invited)
John Shalf, Computer Science, Lawrence Berkeley National Laboratory (LBNL)
The rapid growth of artificial intelligence (AI) is driving an exponential increase in computing demand, with processing requirements doubling every three months rather than the historical two-year interval, significantly surpassing traditional advances in energy efficiency. Since the proximal end of Dennard Scaling in 2005, reducing energy consumption per transistor has become increasingly challenging. Hyperscale datacenters are contributing significantly to an emerging global energy crisis. Current projections indicate that AI-related power consumption may exceed worldwide energy production within the next decade. This paper examines AI's contribution to this crisis and explores the transition from general-purpose processors to specialized architectures. These specialized

solutions include chiplet-based designs, domain-specific accelerators, advanced packaging techniques, and co-integrated systems combining cutting-edge and legacy technologies. These innovations are designed to enhance energy efficiency, balance performance demands, and promote sustainability within resource-constrained environments.

2:25 PM

6-3 Enabling Enterprise AI with IBM Spyre™ (Invited)

Jeffrey L. Burns, Leland Chang, Vijayalakshmi Srinivasan, IBM Research

AI advancements, in Generative AI and Foundation Models, can significantly enhance mission-critical enterprise workloads. The high computational demands of AI require AI accelerator hardware solutions and full-stack AI infrastructure optimization. IBM Spyre is an enterprise AI accelerator hardware and software solution for IBM systems. This paper describes the design and technology approaches taken in the creation of IBM Spyre, and a look towards options for future versions.

2:50 PM

6-4 Trends in Embedded Volatile and Non-Volatile Memories for Scalable Edge AI Applications (Invited)

Tomoya Saito, Koichi Nose, Tomoya Ogawa, Yasuhiko Taito, Takahiro Shimoi, Kosei Fujishiro, Takashi Ito

Renesas Electronics

We present trends in embedded memory technologies, including both volatile and non-volatile memories, for edge AI applications. In this paper, we focus on Edge AI implemented on embedded system powered by MCUs or MPUs, in contrast to AI running on servers in data center. In edge AI applications, all input information from vision, voice and other real-time sensors is processed by embedded systems according to various inference tasks. Therefore, MPUs/MCUs for edge AI require a scalable products lineup to cost-effectively achieve real-time response, higher performance, and low power operation. To expand edge AI applications from MPUs/MCUs system perspective, we implement efficient memory access and memory size reduction for SRAM, as well as high-speed access and low power write for MRAM.

3:40 PM

6-5 Oxide Semiconductor: Technology and Applications (Invited)

Yu-Ming Lin, Chun-Chieh Lu, Cheng-Hsien Wu, Elia Ambrosi, Yu-Chuan Shih, Cheng-Chen Kuo, Hao-Kai Peng, Sheng-Chih Lai, Ken-Ichi Goto, Huai-Ying Huang, Feng-Ching Chu, Chih-Yu Chang, Yen-Chung Ho, Ming-Yen Chuang, Gerben Doornbos, Georgios Vellianitis, Paulina Rincon, Mark van Dal, Katherine H. Chiang, Xinyu Bao, Feng-Cheng Yang, Chung-Te Lin, Min Cao

TSMC

Oxide semiconductors have recently gained significant attention as promising materials for next-generation memory architectures by enabling back-end-of-line (BEOL)-compatible logic and memory devices. This study reports recent progress and challenges in BEOL memory devices that utilize oxide semiconductor channels, including DRAM-like 1T-1C cells, capacitorless gain cells, and non-volatile Ferroelectric FETs. Key characteristics of oxide channels are analyzed, with a focus on advancements in materials and device process technology that improve essential memory metrics, such as endurance, retention and scalability. These findings offer valuable insights into the optimization of oxide semiconductor-based memory devices to address the demand of next-generation applications.

4:05 PM

6-6 Monolithic 3D SRAMs using Thin-Film Transistors: Opportunities and Challenges (Invited)

Hussam Amrouch[1], Mahdi Benkhelifa[1], Girish Pahwa[2], Albi Mema[1]

[1]Technical University of Munich (TUM), [2]National Yang Ming Chiao Tung University

Monolithic 3D Integration (M3D) is a compelling pathway to overcome scaling stagnation. Thin-Film Transistor (TFT) technology has significant potential for integration into the Back-End-of-Line (BEOL), enabling vertical stacking. We evaluate 6T SRAM on-chip memory cells designed utilizing BEOL Indium-tungsten-oxide (IWO)-based TFTs in tandem with Front-End-of-Line (FEOL) Nanosheet Field Effect Transistor (NSFET) technology. A complete 32-bit SRAM array, inclusive of peripheral components, is simulated using SPICE while employing a physics-based

compact TFT model calibrated against experimental data. We comprehensively evaluate the array on noise margins, delay, power, and resilience to process variations. Various M3D SRAM designs incorporating different combinations of NSFETs and TFTs are explored. We demonstrate that BEOL pass-gates lead to a 5x increase in read delay, while substituting all n-type devices with BEOL TFT technology results in a 10x increase in write delay. Nevertheless, SRAM reliability during read operations is enhanced with TFT pass-gates, as the read static noise margin improves by 150%.

4:30 PM

6-7 Oxide-semiconductor Channel Transistor DRAM (OCTRAM) for High-density and Low-power AI memory (Invited)

Shoichi Kabuyanagi[1], Tseng Fu Lu[2], Shosuke Fujii[1], Szu Yao Chang[2], Shinji Miyano[1], Chung Peng Hao[2], Akihiro Kajita[1], Chung Lin Huang[2], Takeshi Fujimaki[1], Chiang-Lin Shih[2]
[1]Kioxia Corporation, [2]Nanya Technology Corporation

We show the potential of oxide semiconductor materials (InGaZnO) in $4F^2$ DRAM by demonstrating high on-current (>10 μA/cell), extremely low leakage current (< 1 aA/cell), and sufficiently long TDDB and BTI lifetime. The results were obtained using an InGaZnO vertical channel transistor designed for reliability. We also showcase a 275Mbit oxide semiconductor DRAM (OCTRAM) array with long retention (>100s), making it a viable option for high-density and low-power AI memory.

4:55 PM

6-8 Material Innovations and Solutions Enabling AI Architectures fueling the AI Era (Invited)

Milan Pesic
Applied Materials

Current AI revolution is increasingly driven by power hungry compute and memory devices. In addition, moving data between compute and memory decreases performance and increases power consumption. Current industry solutions for this involve primarily lower precision compute, sparsity of deep learning models and software optimizations. In this paper key materials engineering innovations that can drive intrinsic transistor power reduction while increasing bit density for storage and reduce latency of memory access will be outlined. Several technology solutions outlined involve integration or co-optimization of diverse unit processes that are building blocks for Si CMOS and DRAM.

Session 7: MS | Compact Modeling and DTCO

1:30 PM, Continental 6

Co-Chairs: Changwook Jeong, UNIST and Lado Filipovic, TU Wien

This session includes 6 papers that describe advances in the area of compact modeling and design-technology co-optimization (DTCO). The first paper, by Chung-Kai Lin of TSMC, introduces a machine learning-based compact modeling platform to address the limitations of existing physics-based models. The work highlights neural network approaches for MOSFET modeling, CNN applications for MEOL/BEOL, and the integration of ML techniques into parameter extraction, model construction, and design sign-off flows (invited). The second paper, by Yijing Feng from the University of Waterloo, presents WAVSiC, a scalable physics-based compact model for SiC MOSFETs that accurately captures key device effects, matches experimental data across voltages and temperatures, and is ready for PDK adoption. The third paper, from Tanmoy Kumar Paul of Purdue University, introduces a physics-based compact model for In2Se3-based ferroelectric semiconductor FETs, calibrated to capture experimental hysteresis trends and provide insights into device physics and array-level optimization for memory applications. The fourth paper, by Sungju Kim from Seoul National University, reports a physics-based SPICE-compatible simulator for threshold voltage distribution in 3D NAND flash, enabled by a closed-form GIDL-assisted erase compact model, achieving quantitative prediction of erase operation. The fifth paper, by Carlo Gilardi of TSMC, describes an analytical framework that bridges materials, devices, and memory macro design to enable fast exploration of a wide technology space for gain cells. The final paper, by Minji Shon from Georgia Institute of Technology, reports on a DTCO of multi-tier CFETs as a novel extension beyond the industry-standard 2-tier CFETs.

1:35 PM

7-1 Machine Learning Oriented Compact Model Platform for Technology Development and Circuit Design (Invited)

Chung-Kai Lin[1], Meng-Lin Lu[2], Jim Liang[2], Cheng Hsiao[2], Yi-Hung Tsai[2], Orion Yu[2], Gary Lee[2], Wen-Che Chang[2], Pei-Wen Liu[2], Guan-Ming Huang[2], Chih-Wei Yeh[2], Yi-Ting Chiang[2], Jin-Yu Lin[2], Ke-Wei Su[2]
[1]TMLD, Taiwan Semiconductor Manufacturing Company, [2]Taiwan Semiconductor Manufacturing Company

A comprehensive solution regarding compact modeling has been proposed. Regarding the primitive MOSFET model, a neural network (NN)-based compact modeling methodology will be introduced to augment existing physics-based model equations. Beyond the scope of MOSFET model, CNN-based MEOL and BEOL modeling approaches will be evaluated in this work. More than that, the proposed work will also address the advantages of the proposed ML-oriented solutions on several scenarios such as parameter extractions, model construction and design sign-off flow. To achieve the new era of SPICE platform, TSMC model interface (TMI/TMI-RC) has been introduced then benefit the whole semiconductor community.

2:00 PM

7-2 WAVSiC: A Physics-based Compact Model for SiC MOSFETs

Johan Alant[1], Yijing Feng[1], Xin Yang[2], Rebecca Fang[1], Qihao Song[3], Bixuan Wang[3], Han Wang[2], Yuhao Zhang[2], Ujwal Radhakrishna[4], Lan Wei[1]
[1]University of Waterloo, [2]University of Hong Kong, [3]Virginia Polytechnic Institute and State University, [4]Texas Instruments

The paper proposes WAVSiC, a comprehensive and user-friendly physics-based compact model for SiC MOSFETs. The model formulation, including the details of the modular approach and the construction of each basic module are explained. The model can automatically account for scalability with device channel-length and other dimensions, as well as critical effects from body diode, JFET depletion, and non-linear drift resistance. Good agreement has been achieved between simulation using fitted model and experimental device characterization and circuit testing, for both a commercial 650V device (at multiple temperatures) and a commercial 1200V device, showing the accuracy, scalability and flexibility of the proposed model. The model has also been validated for computational efficiency, symmetry, robustness, and circuit simulator compatibility, ready for PDK adoption.

2:25 PM

7-3 Physics-Based Compact Model of Ferroelectric Semiconductor Field Effect Transistors

Tanmoy Kumar Paul, Mengwei Si, Peide Ye, Sumeet Kumar Gupta
Purdue University

The presence of ferroelectric and semiconducting properties in two dimensional (2D) In_2Se_3 has made it an interesting material. In particular, Ferroelectric Semiconductor Field Effect Transistors (FeSFET) offer great potential for memory and synaptic applications. In this work, we propose a physics-based compact model of In_2Se_3-based FeSFET. We also present our experimental findings of clockwise (CW) & counter-clockwise (CCW) hysteresis which shows that CW loop is more favorable for high effective oxide thickness (EOT) and thinner In_2Se_3. We show that our calibrated compact model is able to capture these experimental trends. The model also provides insights into the physics of CW and CCW hysteretic characteristics of FeSFET, shedding light onto the interplay between polarization switching, semiconductor electrostatics and multi-layer transport in the 2D FeS. Using our model, we analyze FeSFET-based memory array and perform array-driven optimization of EOT and In_2Se_3 thickness to achieve high distinguishability of binary memory states.

3:15 PM

7-4 Physics-Based Closed-Form GIDL-Assisted Erase Compact Model: SPICE Framework for V_{th} Distribution Simulator of 3D NAND Flash

Sungju Kim[1], Haechan Choi[1], Sechun Park[2], Hyungcheol Shin[1]
[1]Seoul National University, [2]SK Hynix Inc

This work presents the first physics-based SPICE-compatible simulator for threshold voltage (V_{th}) distribution in 3D NAND flash, enabled by a closed-form GIDL-assisted erase compact model. The model integrates GIDL physics and cell operation dynamics, capturing time-dependent hole density (p_{GIDL}). Validated against TCAD and

experimental data, the simulator achieves quantitative prediction of erase distribution and establishes a physics-driven SPICE framework for precise erase V_{th} distribution prediction and optimization.

3:40 PM

7-5 COMET: Co-Optimization of Memory with Exploratory Technologies for Hybrid and Back-End Gain Cell Arrays

Carlo Gilardi[1], Goutham Arutchelvan[2], Xiaoyu Sun[2], Chih-Yu Chang[2], Jen-Chieh Liu[2], Aslan Wei[2], D. Mahaveer Sathaiya[2], Jui Jen Wui[2], Win-San Khwa[2], Meng-Fan Chang[2], Katherine H. Chiang[2], Kerem Akarvardar[2], Iuliana Radu[2], Min Cao[2]
[1]Corporate Research, Taiwan Semiconductor Manufacturing Company, [2]Taiwan Semiconductor Manufacturing Company

AI accelerators require high-density, high-performance on-chip memory. Refresh-free Gain Cell (GC) memory has been proposed as an addition or alternative to SRAM. We present COMET, an analytical framework that bridges materials, devices, and memory macro design to enable rapid evaluation of a wide technology space for GCs. COMET is designed to derive device- and module-level targets necessary for technology development. We analyze multiple GC options, including hybrid FEOL-BEOL and fully BEOL GCs. We explore key trade-offs between read delay, retention, and array size and propose BEOL FET optimizations to meet requirements. Retention time requirements for activation buffer range from 20 us to 20 ms, necessitating distinct GC optimizations. Our analysis shows that hybrid GCs with additional capacitance (~0.5 fF) achieves up to 1 ms retention time, while optimized fully BEOL GCs can exceed 100 ms retention time at the expense of 10x slower read delay than hybrid GCs.

4:05 PM

7-6 DTCO of Multi-Tier CFETs: Ultimate Logic and SRAM Scaling of A5 Technology and Beyond

Minji Shon[1], Seongkwang Lim[2], Shimeng Yu[2]
[1]Electrical and Computer Engineering, Georgia Institute of Technology, [2]Georgia Institute of Technology

This paper proposes a comprehensive Design-Technology Co-Optimization (DTCO) framework that benchmarks "multi-tier" Complementary FET (CFET) integration as a novel extension beyond industry-standard 2-tier CFETs to overcome scaling limitations. Unlike prior CFETs stacking n-type and p-type devices to reduce footprint, our approach proposes 3-tier/4-tier CFETs to enhance vertical integration. 3D layout complexity is solved by standard-cell library compiler (SLiC™) allowing efficient and consistent layout automation. By co-optimizing device architecture and BEOL design, we demonstrate 3-tier and 4-tier CFETs for logic and SRAM cells. The proposed 4-tier CFETs achieve significant area scaling, reducing logic and SRAM cell footprints by 40% and 43%, respectively, compared to 2-tier CFETs. Our framework provides the early-stage power, performance, and area (PPA) insights for next-generation CFET technologies.

Session 8: ODI | Emerging Photodetectors and SPADs

1:30 PM, Continental 7-9

Co-Chairs: Susanna Thon, Johns Hopkins University and Gauri Karve, IMEC

This session includes seven papers that report on advances in emerging photodetector technologies and Single Photon Avalanche Diodes (SPADs). In the first paper - a highlight paper - by Yoshida et al. from Sony a Ge-on-Si SPAD array achieving record photon detection efficiency of 33.8% in the SWIR and outdoor LiDAR operation is demonstrated. The second paper by Shinohara et al. from Canon reports on a prototype SPAD array with a novel double cathode structure to reduce excess charge per event and dark count rate for low light level applications. The third paper, by Kim et al. from Yonsei University, discusses a lateral SPAD with the avalanche region extending over a broad vertical depth, allowing for relatively broadband detection with high photon detection probability. In the fourth paper, Wang et al. from the University of Science and Technology of China, report on ultra-wide bandgap Ga_2O_3/AlN/Si transistors with DUV responsivity for in-sensor memory and computing, for applications in flame detection and space exploration. The fifth paper by Kara et al. from IMEC is about co-integration of Si/SiO_2 meta-surfaces with colloidal quantum dot photodiodes for CMOS compatible multi-spectral SWIR imagers. The sixth paper by Guo et al. from Nanjing University is about double side polished SiC substrates to enable radiation-hardened X-ray spectroscopy. The seventh paper, by Gao et al. from the University of Science and Technology of

China, reports on a miniaturized UV hyperspectral imager based AlGaN/GaN on back-to-back p-i-n photodiodes, capable of acting as both a chip-based UV spectrometer and logic gates for silicon photonics.

1:35 PM

8-1 10 μm Pitch Ge-on-Si SPAD Pixel Array with PDE of 33.8% at 1300 nm and 23.3% at 1550 nm Under Room Temperature Environment

Satoru Yoshida[1], Shohei Shimada[1], Atsushi Matsuzaki[1], Sylvia Mei Lin Loo[1], Yusuke Otake[1], Tetsuji Nagayama[1], Soichi Ochiai[1], Takashi Tange[1], Ken Sasaki[1], Atsuhito Yasui[1], Kaito Yokochi[1], Hiroaki Takase[1], Katsuaki Sasaki[1], Masaru Miyashita[1], Tetsu Ono[1], Kazuma Takahashi[2], Hirotaka Tanaka[2], Ryu Sasaki[2], So Mitoma[2], Toshifumi Wakano[1]
[1]Sony Semiconductor Solutions, [2]Sony Semiconductor Manufacturing

This paper presents a 10μm pitch Back-Illuminated Ge-on-Si SPAD pixel array achieving PDE of 33.8% at 1300nm and 23.3% at 1550nm at room temperature while suppressing DCR. A LiDAR with our SPAD successfully demonstrates no difference in distance measurement performance between indoor and outdoor conditions within a range of 18m.

2:00 PM

8-2 A Novel Double Cathode SPAD for Drastic Reduction of Charge per Event and Dark Noise

Mahito Shinohara, Hiroshi Sekine, Daiki Shirahige, Hajime Ikeda, Takeru Sakamoto, Ayman Tarek Abdelghafar, Kizashi Mikami, Kazuma Inoue, Alice Ehara, Yoshiki Kato, Kazuhiro Morimoto, Junji Iwata, Yasushi Matsuno, Katsuhito Sakurai, Takeshi Ichikawa
Canon Inc.

We introduce a newly developed SPAD realizing significant reduction of charge per avalanche event and dark noise. A proposed pixel is composed of double cathode vertically coupled by a resistive micro-path. Prototype sensors with the double cathode structure and a conventional single cathode structure, both based on 6.39μm pixel pitch, 3.2 megapixels, were evaluated for comparison. The double cathode pixel reduced charge per event by 44%, DCR at 25°C by 82%, and hot pixel population at 60°C by 63%, respectively. The measured DCR is the lowest among the SPAD image sensors, bringing SPAD technology to a new stage.

2:25 PM

8-3 Novel Lateral p-i-n SPAD With High and Broad Spectral Response for Pixel Miniaturization

Joo-Hyun Kim[1], Doyoon Eom[2], Eunsung Park[2], Hyo-Sung Park[2], Seyoung Yook[2], Doo-Hee Son[2], Woo-young Choi[2], Myung-Jae Lee[2,3]
[1]Department of Electrical and Electronic Engineering, Yonsei University, [2]Yonsei University, [3]TruPixel, Inc.

We propose a novel lateral p-i-n SPAD (L-SPAD) fabricated in a back-illuminated 110 nm CIS foundry process. Thanks to the vertically extended avalanche region, L-SPAD features broadband sensitivity, achieving photon detection probability of 33.27% at 940 nm, 95.26% at 625 nm, and 41.9% at 475 nm under 22.7 V breakdown voltage and 3 V excess bias. The L-SPAD can eliminate a guard ring layer, making it highly advantageous for pixel miniaturization. With the integration of 3D stacking and using an advanced fabrication process, the pixel pitch can potentially be scaled down to sub-1.5 μm levels for ultra-high-resolution SPAD arrays.

3:15 PM

8-4 UWBG Ga_2O_3/AlN DUV Phototransistors on Si Platform for In-sensor Memory and Computing Based on Polarization-Induced Carrier Separation Strategy

Yilin Wang[1], Wentao Wu[1], Keju Han[1], Hanyan Zhan[1], Shenyi Huang[1], Peiwen Tong[2], Xiaolong Zhao[1], Juhong Liu[1], Xiao Feng[1], Xiaohu Hou[1], Hong Huang[1], Wei Wang[2], Guangwei Xu[1], Xuanze Zhou[1], Shibing Long[1]
[1]University of Science and Technology of China, [2]National University of Defense Technology

Integration of sensing, memory, and computing in one unit on commercial platform caters to future development of intelligent photodetection systems. Here, ultra-wide bandgap (UWBG) Ga_2O_3/AlN/Si transistor (GAST) with high deep-ultraviolet (DUV) responsivity (275 A/W) is developed for both in-sensor memory and computing. Polarization in Al-polar AlN contributes to both depletion of the Ga_2O_3 channel and non-volatile trapping of holes at the Ga_2O_3/AlN interface. Such effects endow the GAST with ultra-low leakage current (0.26 pA) and long-term multilevel storage. The GAST also harvests short-term memory for synapse-like computing as a reservoir based on

tunable persistent photoconductivity (PPC) dominated by Ga_2O_3 body traps. Physical reservoir computing (PRC) system, constructed with such all-in-one GAST devices, achieves a 96.75% accuracy in motion direction recognition. This work demonstrates the great potential of UWBG transistors with polarization effect on mature Si platform for advanced in-sensor memory and computing application.

3:40 PM

8-5 Metasurface Cointegration with Colloidal Quantum Dot Photodiodes Enabling Facile Spectral Sensors

Gokhan Kara[1], Renaud Puybaret[2], Antonia Malainou[2], Deniz Sabuncuoglu Tezcan[2], Bruno Figeys[2], Roelof Jansen[2], Xavier Rottenberg[2], Klaas Tack[2], Itai Lieberman[2], David Cheyns[2], Pawel Edmund Malinowski[2], Paul Heremans[2,3], Vladimir Pejovic[2]
[1]LEAN, imec, [2]imec, [3]KU Leuven

We demonstrate the cointegration of meta-optics and colloidal Quantum Dot Photodiodes (QDPDs) in microcavities for compact, next-generation short-wave infrared spectrometers and spectral imagers. Microcavities act as narrowband filters, while QDPDs provide broadband sensitivity up to 1600 nm. Amorphous silicon metasurfaces, nanopatterned via DUV lithography, enable tunable optical filtering through effective refractive index variation. We introduce a novel meta-optic design that integrates optical and electrical functionality, enabling direct monolithic integration onto CMOS readout circuits. This scalable platform offers a new approach for high-throughput fabrication of miniaturized spectral detector arrays for industrial applications in environmental, biomedical, food analysis, and other domains.

4:05 PM

8-6 Radiation-Hardened 35-μm Silicon Carbide Film-Type Lateral Drift Detector for Single-Photon X-ray Spectroscopy and High-Energy Transients Profiling

Yue Guo[1], Weizong Xu[2], Tao Miao[2], Renjie Xu[2], Jie Jin[2], Hang Yu[2], Cheng Mao[2], XiangShun Kong[2], Dong Zhou[2], Feng Zhou[2], Fangfang Ren[2], Jiandong Ye[2], Dunjun Chen[2], Rong Zhang[2], Youdou Zheng[2], Hai Lu[2]
[1]School of Electronic Science and Engineering, Nanjing University, [2]Nanjing University

We address the fabrication challenges of conventional single-side-processed silicon carbide detectors by introducing a mechano-chemical hybrid polishing technique, achieving a 35-μm-thick SiC film with sub-micron dimensional uniformity and atomic-scale surface flatness. Leveraging this advancement, we present a lateral drift chamber architecture that monolithically integrates photoelectric conversion and signal processing functionalities into a double-side processed device. This device outperforms traditional vertical structures, featuring extremely low leakage current (40 fA), minimal junction capacitance (0.6 pF at 40 V), and reduced integrated noise charge. When interfaced with readout electronics, the detector demonstrates single-photon time resolution of 132 ns and enhanced energy spectroscopy capability.

4:30 PM

8-7 A Back-to-Back Photodiode for Miniaturized Ultraviolet Hyperspectral Imager and Optical Logic Gates

Zhixiang Gao[1], Huabin Yu[1], Qianqian Zhan[1], Yuanmin Luo[1], Mingjia Yao[2], Zongyin Yang[2], Sheng Liu[3], Haiding Sun[1]
[1]University of Science and Technology of China, [2]Zhejiang University, [3]Wuhan University

Ultraviolet (UV) hyperspectral imagers are critical analytical instrument. Inspired from the silicon-based visible digital camera-on-a-chip concept, here, we proposed a miniaturized on-chip UV hyperspectral imager with logic functions based on the back-to-back photodiodes (BtB-PDs) array. The BtB-PD-based miniaturized UV spectrometer is composed of vertically integrated n-i-p-i-n photodiode structures, featuring electrically tunable spectral response and bidirectional photo-response. High responsivities (450 mA/W at 355 nm and -155 mA/W at 255 nm) with an ultra-fast response time of 12 ns have been achieved. Finally, a miniaturized UV hyperspectral imager is demonstrated using BtB-PD array, providing a peak wavelength resolution of <2 nm and a broad operational bandwidth of 255-365 nm within an ultra-small footprint, with capability of real-time hyperspectral imaging. Furthermore, thanks to the unique bipolar photocurrent output of the BtB-PDs under different optical input, optically controlled logic gates have been demonstrated with “AND”, “OR”, “NAND”, “NOR”, and “NOT” functionalities.

Session 9: PMA | Advanced Power and Cryogenic Devices

1:30 PM, Imperial A

Co-Chairs: Munetaka Noguchi, Mitsubishi Electric Corp. and Christian Koller, Infineon

This session includes eight papers that showcase the latest technology innovation in wide/ultra-wide bandgap power devices and cryogenic devices. The first paper by Hehe Gong and Xin Yang from the University of Hong Kong demonstrates the first ultra-wide bandgap power module with high power capability through device-package, electro-thermo-mechanical co-optimization. The second paper, an invited contribution by Andrew Green from Air Force Research Laboratory, presents recent progress toward the development and performance of lateral Ga_2O_3 field-effect transistors for power conversion applications. The third paper by Xiangjie Xing from Chinese Academy of Sciences demonstrates a novel single-crystal 4H/3C-SiC engineered substrate to achieve the ultra-low resistivity substrate toward next-generation low-voltage SiC power electronics. The fourth paper by Hitoshi Fujioka from MIRISE Technologies presents the first experimental demonstration of dynamic switching and short-circuit performance in 1.2 kV, 500 A vertical 4H-SiC fin field-effect transistors. In the fifth paper of this session, TaeGi Lee from Asahi Kasei Corporation presents advances in AlN/GaN/AlN HEMTs grown by MOVPE on AlN substrates. The following invited paper by Shinichi Takagi from Teikyo University discusses both the electron transport properties in nanosheet MOSFETs and their behaviour at cryogenic temperatures. The seventh paper by Mohammad Abu Zahra from Infineon Technologies demonstrates the effectiveness of field plates for next-level cryogenic LDMOS devices for trapped ion quantum computers. In the last paper of this session, Yijia Fan from Tsinghua University demonstrates IGZO-based DrMOS with vertical-channel power transistors for monolithic 3D-integrated active PDNs for on-chip voltage conversion.

1:35 PM

9-1 First Demonstration of an Ultra-Wide Bandgap Power Module through Device-Package, Electro-Thermo-Mechanical Co-Optimization

Hehe Gong[1], Xin Yang[1], Boyan Wang[2], Zichen Zhang[2], Qingrui Yuchi[2], Zineng Yang[1], Zhengpeng Wang[1], Matthew Porter[2], Hongchang Cui[1], Yuan Qin[2], Yibo Wang[3], Xiaosheng Wang[3], Chaoqiang Jiang[3], Rong Zhang[4], Han Wang[1], Dong Dong[2], Jiandong Ye[4], Guoquan Lu[2], Yuhao Zhang[1]

[1]The University of Hong Kong, [2]Virginia Polytechnic Institute and State University, [3]City University of Hong Kong, [4]Nanjing University

We demonstrate, for the first time, an ultra-wide bandgap (UWBG) power module, which integrates 6 packaged Ga_2O_3 dies and achieves 1000 V and 200 A switching – over 10× higher power capacity than previously reported UWBG devices. To address Ga_2O_3's low thermal conductivity, a junction-side-cooling (JSC) package is employed. To overcome the electric field crowding in JSC package, we introduce two novel device-package interface designs: a post interface and a high-κ dielectric interface. The post interface preserves the device's *BV*, while the high-κ interface further enhances *BV*. Thermally, the high-κ design reduces thermal resistance by 50% compared to the post design. Mechanically, both the post and high-κ interface provide stress relief, which are validated experimentally by power cycling tests. Overall, these electro-thermo-mechanical co-optimizations offer critical guidance for UWBG module development and enable a key milestone in power scaling for UWBG power technologies toward industrial applications.

2:00 PM

9-2 Progress Towards Demonstration of Ga_2O_3 based Power Modules (Invited)

Andrew J Green[1], Ahmad Islam[2], Nolan Hendricks[2], Daniel Dryden[2], Elizabeth Sowers[2], Kyle Liddy[2], Kevin Leedy[2], Dennis Walker Jr.[2], Jacob Breakfield[2], Laura Davidson[3], Aaron Adams[3], Weisong Wang[4], Michael Harrington[3], Zach Weber[5], Shin Mou[2], Adam Neal[2], Thaddeus Asel[2], Joshua Piel[2]

[1]Sensors Directorate, Air Force Research Laboratory, [2]Air Force Research Laboratory, [3]KBR, [4]Wright State University, [5]The Ohio State University

Abstract—Gallium Oxide (Ga_2O_3) device technology has been under development for just over a decade and significant progress has been made towards demonstration of relevant power switches in comparison with the incumbent GaN and SiC technologies. Here, we report recent progress towards the development and performance of lateral Ga_2O_3 field effect transistors for power conversion applications. Further, we investigate device periphery scaling which enables on-resistances (R_{ON}) required for power module integration. Finally, we review the

integration of these large periphery devices into the first DC/DC converter utilizing a cascode configuration topology integrating Ga_2O_3 transistors in the high and low-side switches.

2:25 PM

9-3 Single-Crystal 4H/3C-SiC Engineered Substrate: A Novel Platform Enabling High-Performance Low-Voltage SiC Devices

Xiangjie Xing[1], Yidan Tang[1], Xin Yang[2], Zhen Dong[1], Zheng Hu[1], Zineng yang[2], Hehe Gong[2], Fuchao Liu[3], Huaping Song[4], Xinglin Xiao[5], Jinan Shi[6], Yipei Lei[1], Yongqiang Zhang[1], Yifei Feng[1], Zhou Wu[6], Xinhua Wang[1], Chao Yuan[5], Fengwen Mu[3], Yuhao Zhang[2], Xiaolong Chen[7], Xinyu Liu[1], Sheng Liu[5], Yue Hao[8]
[1]Institute of Microelectronics, Chinese Academy of Science, [2]University of Hong Kong, [3]TJ Innovative Semiconductor Substrate Technology Co., Ltd, [4]PowerEpi Semiconductor Co., Ltd, [5]Wuhan University, [6]University of Chinese Academy of Sciences, [7]Institute of Physics, Chinese Academy of Science, [8]Xidian University

4H-SiC devices are usually considered uncompetitive for low-voltage applications due to substrate resistance limitations. This work challenges this long-held belief by demonstrating a novel 150 mm single-crystal 4H/3C-SiC engineered substrate with resistivity over 45× lower than conventional 4H-SiC. The engineered substrate leverages the ultra-low resistivity of 3C-SiC and overcomes its high defect density by integrating a prime-grade 4H-SiC template. This template is transferred via surface-activated bonding and ion-cut. A sharp, defect-free bonding interface was achieved with a potential barrier below 30.4 mV and a record thermal boundary resistance of 1 + 0.7/–0.6 m^2·K/GW. 200 V SBD fabricated on this substrate demonstrates a record low specific on-resistance ($R_{on,sp}$) of 0.50 mΩ·cm^2 and a surge current tolerance over 300 A—representing a 47% reduction in $R_{on,sp}$ and a 20% increase in surge current robustness compared to identical 4H-SiC devices. It establishes a scalable hetero-integration material platform for enabling next-generation low-voltage SiC power electronics.

2:50 PM

9-4 Breaking the Trade-off between On-Resistance and Short-Circuit Withstand Time in 1.2 kV - 500 A SiC MOSFETs Using a Narrow Fin Channel

Hitoshi Fujioka[1], Teruaki Kumazawa[1], Hidemoto Tomita[1], Masaru Okuda[1], Hirokazu Fujiwara[1], Qian Wang[2], Tsunenobu Kimoto[3], Florin Udrea[2]
[1]MIRISE Technologies Corporation, [2]University of Cambridge, [3]Kyoto University

We present the first experimental demonstration of switching and short-circuit performance in 1.2 kV–500 A vertical 4H-SiC FinFETs. Devices with different channel fin widths (W_{fin}) of 144, 199, and 224 nm were investigated. The narrowest FinFET (144 nm) achieved a low on-resistance (1.64 mΩ·cm^2) and breakdown voltage over 1400 V. Under short-circuit conditions, the peak drain current was suppressed by 15%, and short-circuit withstand time improved significantly as W_{fin} decreased. These results clearly demonstrate that optimizing the fin width of SiC FinFETs can significantly enhance the on-state performance, switching performance, and short-circuit ruggedness, highlighting the strong potential of SiC FinFETs.

3:40 PM

9-5 Low-Resistance and Current-Collapse-Free MOVPE-Grown Pseudomorphic AlN/GaN/AlN HEMTs on AlN Substrates

TaeGi Lee[1,2], Akira Yoshikawa[3,4], Yoshihito Hagihara[3], Sho Sugiyama[3,4], Manabu Arai[4], Yuji Ando[4], Jun Suda[4], Hiroshi Amano[4]
[1]Research Laboratory of Advanced Science & Technology, Asahi Kasei Corp., [2]Institute of Materials and Systems for Sustainability, Nagoya University, [3]Asahi Kasei Corp., [4]Nagoya University

We report MOVPE-grown pseudomorphic AlN/GaN/AlN HEMTs on AlN substrates, achieving low contact resistance and high breakdown fields. TEGa-based epitaxy suppresses carbon incorporation, improving 2DEG density and mobility. Recess etching of the AlN barrier enables low contact resistances without regrowth. Devices exhibit breakdown fields up to 2 MV/cm and specific on-resistance competitive with leading GaN HEMTs. Stress-engineered SiN passivation effectively suppressed surface traps, enabling current-collapse-free operation without degrading mobility or increasing gate leakage. The behavior is further supported by the overall high-quality epitaxial structure, which includes a AlN substrate with low threading dislocation density (TDD) (<10^4 cm^{-2}), non-doped AlN buffer, and fully strained AlN/GaN/AlN without lattice relaxation. These findings validate MOVPE as a scalable and

manufacturable approach for ultra-wide bandgap HEMTs, offering a promising route for next-generation power and RF electronics.

4:05 PM

9-6 Electron Transport Properties in Nanosheet MOSFETs – From Assessment of Optimum Channel System to Operation at Cryogenic Temperatures (Invited)

Shinichi Takagi[1,2], Kei Sumita[3], Xueyang Han[3], Zhao Jin[3], Yutong Chen[3], Min-soo Kang[3], Mitsuru Takenaka[3], Kasidit Toprasertpong[3]

[1]Advanced Comprehensive Research Organization, Teikyo University, [2]Department of Electrical Engineering and Information Systems, The University of Tokyo, [3]The University of Tokyo

We propose a strategy to enhance current drive in nano-sheet MOSFETs to suppress thickness fluctuation scattering, which determines significant mobility reduction in a thin-body channel region, by using an improved surface roughness scattering model. We point out the importance of employing valleys with anisotropic effective mass, heavy confinement effective masses and light parallel effective mass. From this viewpoint (111) GOI and InAs channels with L_1 valley electrons, and (100) GOI one with Δ_2 valley electrons are promising for n-FETs. We also present the recent progress of experimental achievements on these devices. Also, the present analysis is extended to mobility at low temperatures. The mobility at cryogenic temperatures is well represented by the present model, though refinement on the screening model is needed. Systematic experiments and modeling of sub-threshold swing (SS) at cryogenic temperatures are also performed, revealing the critical roles of mobile and localized tail states on SS.

4:30 PM

9-7 Improved On-Resistance of Cryogenic LDMOS Devices Utilizing a Field Plate for Scaling Up Trapped Ion Quantum Computers

Mohammad Abu Zahra[1,2], Andrey Kravchenko[3], Jens Repp[1], Anirban Kar[4], Matthias Brandl[1], Ralf Brederlow[2], Hussam Amrouch[4]

[1]Infineon Technologies AG, Oberhaching, [2]Chair of Circuit Design, Technical University of Munich, [3]Infineon Technologies Dresden GmbH & Co. KG, [4]Chair of AI Processor Design; Munich Institute of Robotics and Machine Intelligence, Technical University of Munich

At cryogenic temperatures, LDMOS devices exhibit a diode-like behavior, leading to nonlinear on-resistance at low V_{DS}. This work explores the use of a field plate to improve the on-resistance performance of LDMOS devices at cryogenic temperatures. Our key findings from measuring the fabricated device include: 1) a field plate is essential at cryogenic temperatures; 2) it reduces on-resistance by six orders of magnitude at low VDS; 3) it mitigates the breakdown voltage; 4) device nonlinear region span decreases to 40 mV compared to the previous threshold of 400 mV. TCAD simulation at cryogenic temperatures are provided supporting the presented results.

4:55 PM

9-8 High-Performance IGZO-based DrMOS with Vertical Channel Power Transistor for Monolithic 3D Integrated Active PDNs with On-Chip Voltage Conversion

Yijia Fan[1], Jianshi Tang[1], Yanbo Su[1], Hu Zhou[2], Rongmei Chen[2], Xiaosen Liu[1], Hongsen Zhang[1], Kexin Wang[1], Zhidong Tang[1], Yibei Zhang[1], Jiaming Li[1], Jianing Li[1], Linqiu Wang[2], Yiwei Du[1], Peng Yao[1], Bin Gao[1], He Qian[1], Huaqiang Wu[1]

[1]Tsinghua University, [2]Peking University

We demonstrate the first BEOL-compatible DrMOS using vertical-channel oxide semiconductor power transistors for monolithic 3D (M3D) integrated active power delivery network enabling on-chip voltage conversion. By employing a vertical IGZO power transistor structure and engineered channel composition, a low $R_{ON,sp}$ of 0.27 $m\Omega \cdot mm^2$ and a high V_{BD} of 31.6 V was achieved. An IVR was realized via IGZO NFET and CNT PFET, validated by electrical measurements showing ~190 ns switch-on time. A simulated 3nm HP CPU using the M3D active PDN exhibited a 32.7% lower IR drop and a 68.8% reduced routing loss, offering a viable M3D-IC system design architecture.

Session 10: EDT | Advances in 2D Material Devices and Integration

1:30 PM, Imperial B

Co-Chairs: Farnaz Niroui, MIT and Chao Ching Cheng, TSMC

This session features eight papers highlighting the advances in 2D material processing, device engineering and integration. The session begins with IMEC introducing a fab-compatible selective etch process for damascene-type top contacts and replacement oxide integration. This is followed by three talks highlighting advances in materials growth, device performance, and reliability in extremely scaled MoS_2 devices. Samsung Electronics reports EOT-scaled devices using direct channel growth on 200 mm wafers. Purdue University presents contact extension doping. A comprehensive reliability study of MoS_2 FETs will then be discussed by MIT. In the second half of the session, performance enhancements in WSe_2 p-type devices will be covered by Peking University and IMEC. The session concludes by TU Wien/Intel and TSMC highlighting stability, reliability, and critical modules in nanosheet transistors.

1:35 PM

10-1 Selective Etch Process for Fab-compatible Top Contacts, Replacement Oxide and Interlayer Removal in 2D FETs

Quentin Smets[1], Tom Schram[1], Dennis van Dorp[1], Yuchao Jiang[1], Ann Opdebeeck[1], Daire Cott[1], Farid Sebaai[1], Pierre Morin[1], Bogdan Govoreanu[1], Luca Panarella[1,2], Azimkhan Kozhakhmetov[3], Chelsey J. Dorow[3], Adedapo A. Oni[3], Kevin P. O'Brien[3], Uygar E. Avci[3], Cesar Javier Lockhart de la Rosa[1], Gouri Sankar Kar[1]
[1]imec, [2]KULeuven, [3]intel

Two rarely used features of Transition Metal Dichalcogenides (TMDs) are their exceptional chemical stability in some wet etchants and their unique anisotropic van der Waals structure. These are leveraged to selectively recess the oxide cap, and fabricate 2D FETs with damascene-type top contacts on monolayer WS_2, MoS_2, and multilayer WSe_2 in a 300mm pilot line. In addition, the process is extended to a replacement oxide stack, and interlayer selective removal is demonstrated using liquid intercalation, reducing the top gate CET from 2.5 to 1.5nm. These form three new fundamental building blocks for 2D integration technology.

2:00 PM

10-2 EOT-scaled 2D FETs on Embedded Bottom Gates Using Direct Channel Growth on 200 mm Wafers

Huije Ryu, Min Seok Yoo, Alum Jung, Minsu Seol, Baekwon Park, Eun-Kyu Lee, Eunji Yang, Junyoung Kwon, Joung Eun Yoo, Taejin Choi, Kyung-Eun Byun, Changhyun Kim
Samsung Advanced Institute of Technology (SAIT), Samsung Electronics

This work demonstrates equivalent oxide thickness (EOT)-scaled, direct-grown MoS_2 channel two-dimensional (2D) field-effect transistors (FETs). A robust fabrication process with oxide passivation enables nearly 100% yield on 200 mm wafers and clean contact interfaces. Direct-grown devices exhibit lower interface trap density (D_{it}) and reduced variation compared to transferred devices. Moreover, we engineer the crystal structure of the gate dielectrics using Al-doped HfO_2 (HfAlO) and ZrO_2 as substrates during direct MoS_2 growth.The optimized devices deliver an on-current (I_{ON}) of 820 μA/μm at an EOT of 1.15 nm, the highest among direct-grown 2D FETs.

2:25 PM

10-3 First Demonstration of DG Monolayer MoS_2 FETs with 0.3 nm-thin Contact Extensions Achieving Near Immunity to SCEs at L_{CH}=20 nm and g_m=206 μS/μm

Jun Cai[1], Hao-Yu Lan[2], Yuanqiu Tan[2], Zhihong Chen[2], Joerg Appenzeller[2]
[1]School of Electrical and Computer Engineering, Purdue University, [2]Purdue University

In this work, we present the first demonstration of ultrathin (0.3 nm) contact extensions in ultra-scaled dual gate (DG) MoS_2 FETs to address short-channel effects (SCEs) induced by 'contact gating' - a key limitation in existing 2D FETs. The ultra-scaled DG devices feature channel lengths (L_{CH}) down to 20 nm and a combined equivalent oxide thickness (EOT) of 0.75 nm, with on-currents reaching 460 μA/μm and maximum transconductance values reaching 206 μS/μm at V_{DS}=1 V, the highest reported for 20 nm channels. Moreover, due to the ultra-thin contact extensions, our devices exhibit near immunity to SCEs. The off-state performance specs outperform prior reported data on 20 nm FETs, with minimal subthreshold slope (SS) degradation at 20 nm (< 25% increase compared to long-channel device), SS= 97 mV/dec, DIBL=30 mV/V, and a near zero threshold voltage. This work marks a key step toward aggressively scaled logic devices based on 2D materials.

2:50 PM

10-4 Reliability Study of MoS_2 Transistors with Channel Length Scaling and Sub-1-nm EOT

Zhongyunshen Zhu[1], Aijia Yao[2], Jiadi Zhu[2,3], Yixuan Jiao[2], Jing Kong[2], Tomás Palacios[2]
[1]EECS, Massachusetts Institute of Technology, [2]Massachusetts Institute of Technology, [3]CDimension Inc.

In this work, we present a systematic reliability study of monolayer MoS_2-channel MOSFETs with a scaled EOT (~0.8 nm) and channel length (L_{ch}) down to 40 nm. By carefully engineering gate and drain bias conditions, we correlate V_T instability with L_{ch} and identify a transition in dominant degradation mechanisms from long-channel to short-channel devices. TCAD simulations validate the experimental results, revealing a strong coupling between bias-temperature instability (BTI) and hot-carrier injection (HCI) under simultaneous gate/drain stress, primarily induced by non-uniform electrical field distribution in the channel and gate dielectric. Our findings demonstrate that highly scaled MoS_2 transistors exhibit reliability metrics — such as PBTI and HCI degradation — comparable to those of planar Si CMOS with similar channel lengths.

3:40 PM

10-5 First Demonstration of Monolayer WSe_2 p-FETs with 53 nm Ultra-scaled Pitch (L_C=20 nm, L_{CH}=33 nm, W=100 nm) and I_{on} of 1195 μA/μm at V_{ds}= -1 V

Aocheng Qiu, Xiong Xiong, Xin Wang, Yuhang Chen, Yuzhe Zhu, Shiyuan Liu, Tianyue Fu, Yanqing Wu
Peking University

Two-dimensional (2D) semiconductors offer great promise for ultimate transistor scaling, with significant progress already demonstrated in n-FETs based on MoS_2. However, high-performance p-type 2D transistors under aggressive pitch scaling remain insufficiently explored. In this work, we systematically investigate ultra-scaled monolayer WSe_2 p-type transistors, achieving full-dimensional scaling of channel length, width, and contact length. A record contact pitch of 53 nm is demonstrated, with L_C=20 nm and L_{CH}=33 nm. The scaled device with a 100 nm channel width exhibits a high on-current of 350 μA/μm at V_{ds}=-1 V. Meanwhile, devices with a contact length of 200 nm achieve a lower contact resistance (R_C) of 0.61 kΩ·μm and an on-current exceeding 1200 μA/μm at V_{ds}=-1 V. This work indicates the tremendous potential of WSe_2 p-type transistors for ultra-scaled CMOS technologies.

4:05 PM

10-6 Towards a Low-Disorder pFET Gate Stack for Monolayer WSe_2 Channels

Dat Tien Ngo[1], Goutham Arutchelvan[2], Sui-An Chou[2], Zaoyang Lin[3], Robert Kimes Grubbs[3], Souvik Ghosh[3], Daire Cott[3], Kaustuv Banerjee[3], Xiangyu Wu[3], Ming-Yang Li[2], Mark Van Dal[2], Iuliana Radu[2], Cesar Javier Lockhart de la Rosa[3], Gouri Sankar Kar[3], Bogdan Govoreanu[3]
[1]Leuven, imec, [2]TSMC, [3]imec

We address challenges of pFET gate-stack fabrication by leveraging synthetic bilayer WSe_2, which combines low-defectivity monolayer WSe_2 with optimized transfer process. This approach achieves low subthreshold swings of 65 mV/decade, I_{on}/I_{off} of over 10^7 and median mobility of 38 $cm^2/V{\cdot}s$, with maximum of 85 $cm^2/V{\cdot}s$. These improvements lead to record high on-currents of 108 μA/μm at a technology-relevant overdrive voltage of 1 V and maximum current of 690 μA/μm. By optimizing WSe_2 oxidation, we achieve controlled V_t shifts while maintaining device performance and enabling top gate EOT of ~1.4nm – the smallest EOT reported for pFET devices with monolayer WSe_2.

4:30 PM

10-7 From Planar to Gate-All-Around: Stability and Reliability of Scaled 2D FETs

Alexandros Provias[1], Rittik Ghosh[2], Mohammad Rasool Davoudi[2], Alexander Karl[2], Theresia Knobloch[2], Dominic Waldhoer[2], Seyedmehdi Sattari-Esfahlan[2], Tobias Zinsler[2], Wouter Mortelmans[3], Chelsey Dorow[3], Ashish Verma Penumatcha[3], Kevin P O'Brien[3], Pratyush Buragohain[3], Marc Jaikissoon[3], Ande Kitamura[3], Sudarat Lee[3], Rahul Ramamurthy[3], Scott Clendenning[3], Michael Waltl[2,2], Uygar Avci[3], Tibor Grasser[2]
[1]Institute for Microelectronics, Technical University of Vienna, [2]Technical University of Vienna, [3]Intel Corporation

Two-dimensional (2D) channels could potentially outperform silicon in the ultimately scaled complementary field-effect transistor (CFET) configuration. However, forming a high-quality interface between layered van der Waals (vdW) channels and the gate oxide remains a major challenge for 2D FETs. In this study, we comprehensively test

the stability and reliability of two transistor technologies: Planar and Gate-All-Around (GAA) FETs. Both feature a 1-L MoS_2 channel and a gate stack of amorphous hafnium dioxide (HfO_2): 4.3 nm for Planar and 4.5 nm for GAA. The GAA, however, are scaled 2D FETs with nanometer-range dimensions, enabling observation of single atomic charge trapping events. The main objective of this study is to compare these technologies in terms of hysteresis, bias temperature instability (BTI), and random telegraph noise (RTN) measurements, which together offer a better understanding of the trap physics within the oxide and at the channel/insulator interface.

4:55 PM

10-8 Mechanics of Integration and Component Performance Step-up for Nanosheet Transistors with Ultrathin 2D Channel

Terry Hung[1], Shu-Jui Chang[1], Yun-Yan Chung[1], Ang-Sheng Chou[1], Wei-Sheng Yun[1], Wen-Chia Wu[1,2], Meng-Zhan Li[1,3], Shuer Ku[1], Po-Sen Mao[1,2], Sui-An Chou[1], Edward Chen[1], Sheng-Kai Su[1], Yu-Tung Lin[1], Bo-Jhih Chou[1,2], Yu-I Chang[2], Chueh-Cheng Yang[4], Shin-Yuan Wang[2], Kris K.H. Lin[2], Liang-Ching He[2], Yi-Fan Ho[2], Kuan-Bo Lin[2], Wei-Tung Liu[2], Jen-Hao Yang[2], Shu-Wei Huang[1], Erh-Chen Lin[1], Chun-Chun Lin[1], Yen-Hui Liu[1], Chen-Feng Hsu[1], T.Y. Lee[1], Georgios Vellianitis[1], Mark van Dal[1], Bo-Heng Liu[5], Chi-Chung Kei[5], Ying-Mei Yang[6], Tse-Ming Chen[6], Minn-Tsong Lin[3], Chun-Liang Lin[2], Wen-Hao Chang[2], Chao-Hsin Chien[2], Chenming Hu[2,7], Ming-Yang Li[1], Chao-Ching Cheng[1], Iuliana Radu[1], Min Cao[1]

[1]TSMC, [2]NYCU, [3]NTU, [4]NSRRC, [5]NCIR, [6]NCKU, [7]UC Berkeley

Mechanical robustness is critical for nanosheet (NS) integration, especially in ultra-thin 2D channels. Bilayer MoS_2 supports NS integrity at extended gate lengths and enables C-contact geometry with Leff up to 30 nm. Nearly epitaxial ALD-deposited metal provides a ~6.0 eV work function with WSe_2, suitable for PMOS contact, and is compatible with phosphorous doping. Mobility remains dielectric-limited, but appropriate interlayer (IL) insertion and surface pre-treatment reduce remote phonon scattering and interface defects, yielding ~1.6× enhancement. Monolayer WSe_2 mobility thus reaches relevant values at a relevant EOT of 1.4 nm.

Monday Evening Reception

Welcome and Networking Reception

6:30 PM – 8:00 PM, Grand Ballroom B

Tuesday, December 9, 2025

Session 11: NC | Adaptive and On-chip Learning

9:00 AM, Grand Ballroom A

Co-Chairs: Erika Covi, University of Groningen and Ming-Hsiu (Eric) Lee, Macronix

This session contains 5 papers that deal with learning and training on chip. The first paper by Ando et al., from IBM, demonstrates transfer learning tasks by using RRAM technology and the Tiki-Taka algorithm, where device asymmetry and variability are taken into account. The second paper by Yan et al., from Peking University, presents an accelerator for embodied self-supervised learning with an 8Mb RRAM compute-in-memory chip. Two techniques are introduced to improve system performance and adaptive programming functionalities. The third paper by Wang et al., from Nanjing University, presents a compute-in-memory array based on 4T1C IGZO DRAM cells. A time-dependent ADC calibration strategy that maintains high precision over extended periods without refresh is introduced. The fourth paper by Li et al., from IMECAS, performs online learning on an analog computing in-memory chip based on back end of the line metal-ferroelectric-metal FeFET. The modular macro-design enables cascading and pipelining of multiple macros. The fifth paper by Jung et al., from KAIST, presents a tactile system that combines threshold switching and L-AFeFET devices. The system works at bio-compatible timescales and uses time-to-first-spike coding of pressure stimuli.

9:05 AM

11.1 Transfer Learning on Edge Using 14nm CMOS-compatible ReRAM Array and Analog In-memory Training Algorithm (Invited)

Takashi Ando[1], Omobayode Fagbohungbe[1], Kenichi Imakita[2], Hiroyuki Miyazoe[1], Sara Aoki[2], Ruturaj Pujari[1], Sarah Lynch[2], Paul Jamison[1], Kevin Brew[1], Kishore Natarajan[1], Gerald Gibson[1], Robert Bruce[1], Rinus Lee[2], Cory Wajda[2], Takaaki Tsunomura[3], Vijay Narayanan[1]
[1]IBM Research, [2]TEL Technology Center, America, [3]Tokyo Electron Limited
This work presents a holistic approach to enabling energy-efficient on-chip Transfer Learning (TL) via Analog In-Memory Computing (AIMC) using 14nm CMOS-compatible ReRAM arrays. We develop an optimized ReRAM stack featuring H_2 plasma-treated high-k (HfO_2 or ZrO_2) and in-vacuo processing, achieving reverse area scaling of forming voltage for co-integration with advanced-node CMOS technologies. To address non-ideal analog weight updates, we implement and evaluate the latest versions of Tiki-Taka training algorithms—TTv2, c-TTv2, and AGAD—capable of tolerating device asymmetry and variability. TL is demonstrated on hardware using compressed MNIST with on-chip training and extended via simulations to Vison Transformer (ViT)-based TL from CIFAR-10 to CIFAR-100. While analog-only models show sensitivity to weight transfer noise, hybrid analog-digital implementations maintain performance up to 20% noise. Using AGAD with optimized ReRAM devices, we achieve <1% accuracy degradation compared to digital baselines, validating AIMC-based TL as a viable path for low-power, on-chip training at the edge.

9:30 AM
11.2 An 8Mb Learning-Aware RRAM Compute-in-Memory Accelerator for Embodied Self-Supervised Learning
Longhao Yan[1], Zhe Zhan[2], Zelun Pan[2], Bowen Wang[2], Yaoyu Tao[1,2], Yuchao Yang[1,2,3]
[1]School of Integrated Circuits, Peking University, [2]Peking University, [3]Chinese Institute for Brain Research
Embodied Self-Supervised Learning (E-SSL) enables intelligent agents to autonomously adapt to changing environments without human annotations, a key requirement for real-time, edge-based autonomy. However, conventional computing architectures suffer from energy and latency inefficiencies due to frequent data movement, impeding on-device learning. To address this, we present an 8Mb learning-aware RRAM-based Compute-in-Memory (CIM) accelerator tailored for E-SSL. Fabricated in a 40nm process, the CIM chip introduces two significant innovations: a Two-Stage Analog Weight Programming (TSAWP) unit for fast, precise, and relaxation-mitigated device conductance programming, and a Neural-Optimized Gradient-Aware Programming Scheduler (NoGAPS) for lifetime-aware, adaptive programming. The proposed chip achieves 10.2× improvements in programming speed and 11.7× system lifetime improvement compared to conventional schemes. Deployed on a quadruped robot for terrain adaptation, the system demonstrates 347× higher energy efficiency and an 8.7× reduction in latency compared to GPU baselines, showcasing robust, low-power, real-time online learning in dynamic edge scenarios.

9:55 AM
11.3 Precision-Sustained Analog In-Memory Computing using IGZO DRAM Array with Quasi-Nonvolatile and Continuously Tunable Weights
Cong Wang[1], Xuan Feng[2], Yu Shen[3], Zhengwei Liao[3], Shuo Ding[3], Zhuo Li[2], Xing-Jian Yangdong[3], Yichen Zhao[3], Dehe Kong[3], Ding-Chang Zhang[2], Yang Lv[2], Ming-Hua Xuan[2], Xiao-Chuan Chen[2], Shi-Jun Liang[3], Feng Miao[3]
[1]School of Physics, Nanjing University, [2]BOE Technology Group Company Limited, [3]Nanjing University
We present a 58×32 compute-in-memory array based on 4T1C IGZO DRAM cells, demonstrating precise vector-by-matrix multiplication operations with quasi-nonvolatile and continuously tunable weights. Fabricated using a large-scale IGZO flexible panel process, the CiM array achieves ultra-long retention of 100,000 (or 2,000) seconds with 8-level (or 263-level) weight precision. The analog weights in CiM array can be precisely programmed into the target values, via a self-calibrating write scheme immune against device-level variation. Experimental results validate accurate VMM operations performed by our 58×32 IGZO CiM array, achieving a low RMSE of 0.0078 μA between measured and expected results (48,000 data points).

10:45 AM
11.4 A 1763.9 TOPS/W 520kb Scalable FeFET-Based Analog Computing-in-Memory Chip Using BEOL HZO Capacitor for On-chip Learning
Weizeng Li[1,2], Junyu Zhu[2,3], Zhidao Zhou[2,3], Linfang Wang[4], Junzhe Shen[2,3], Bohan Wang[2,3], Zhi Li[2,3], Wang Ye[2,4], Zhongze Han[2,3], Hanghang Gao[2,3], Hongyang Hu[3], Qing Luo[2,3], Chunmeng Dou[2,3], Ming Liu[3]

[1]State Key Laboratory of Fabrication Technologies for Integrated Circuits, Institute of Microelectronics of the Chinese Academy of Sciences, [2]University of Chinese Academy of Sciences, [3]Institute of Microelectronics of the Chinese Academy of Sciences, [4]Columbia University
This work presents a highly energy-efficient learnable computing-in-memory (CIM) chip enabled by ferroelectric FETs (FeFET) based on BEOL-integrated HZO capacitors. It features: (1) manufacturable metal-ferroelectric-metal FeFET arrays (MFAs) with high endurance and variation tolerance in a standard logic process, (2) ADC-free current I/O ports for high area- and energy-efficiency, and (3) a modular macro design enabling cascading and pipelining of multiple macros. A dual-macro 520-kb FeFET CIM chip has been silicon-verified. The MFA supports over 10^9 write cycles and shows 75.76% reduced device variations. The chip achieves a peak energy efficiency of 1763.9 TOPS/W and a peak throughput of 27.7 TOPS. The MNIST test evaluation shows accuracies of 84.10% for forward-forward learning and 91.90% for the multi-layer perceptron approach, respectively. This work first demonstrates ultra-low-power on-chip learning ferroelectric analog CIM hardware.

11:10 AM

11.5 Sub-1 ms Bio-Inspired Time-to-First Spike (TTFS) Neuromorphic Tactile System Using L-AFeFET and Threshold Switch Devices

Taeseung Jung[1], Hyunsik Choi[2], Seonjae Park[3], Seungyeob Kim[3], Jinho Ahn[4], Jiyong Woo[2], Sanghun Jeon[3]
[1]School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), [2]Kyungpook National University, [3]Korea Advanced Institute of Science and Technology (KAIST), [4]Hanyang University
We present a bio-inspired neuromorphic tactile system achieving sub-1 ms time-to-first-spike (TTFS) coding of pressure inputs. Co-integrating a threshold switch and leaky anti-ferroelectric FET, the system converts event-driven analog stimuli into TTFS spikes within 1 ms. By treating the sensor array as a single receptive field and dynamically modulating body bias for lateral inhibition, redundant spiking and local interference are suppressed, reducing spike count by 33.1 % and synaptic operations by 52.8 % versus conventional TTFS. This architecture bridges the latency and energy-efficiency gap for next-generation robotics and prosthetics.

Session 12: MT | Ferroelectric Memories

9:00 AM, Grand Ballroom B
Co-Chairs: Shimeng Yu, Georgia Institute of Technology and Seiyon Kim, SK Hynix
Ferroelectric materials are unlocking new functionalities for next-generation memory technologies. In this session, CXMT presents a 2-tier horizontally stacked 1T1C 3D FeRAM featuring dual-gate poly-Si access transistors. Researchers from the University of Notre Dame introduce a 2T-nC FeRAM architecture with a scalable pathway toward hundreds of layers. National Taiwan University presents a 3 nm HZO FeCap with low switching voltage, integrated on top of CMOS for non-volatile SRAM operation. Georgia Tech demonstrates direct nanoscale observation of polarization switching in HZO ferroelectrics using advanced metrology, and provides new insights into the device physics underlying the fluid imprint effect, which critically impacts FeRAM read reliability. POSTECH reports a record-high on-state current density in a 1 nm-thin HZO FTJ with a 2D material interface. Finally, Seoul National University showcases a record-high 2Pr in HZO FeCap and highlights its potential for neuromorphic applications.

9:05 AM

12-1 First Experimental Demonstration of Monolithically Stacked FeRAM with Dual-Gate Poly-Si Access Transistors and Horizontal Ferroelectric Capacitors

Huihui Li, Yali Tan, Zihao Yuan, Zean Guo, Chen Yang, Yu Li, Zelun Li, Jifang Chen, Jianfeng Xiao, He Wang, Gaofeng Xu, Yiming Ding, Hao Meng, Mingyuan Liu, Kanyu Cao
Changxin Memory Technologies, Inc
Ferroelectric memories have attracted extensive attention due to its high speed, non-volatility, and superior power-efficiency characteristics. In this work, we present the world's first silicon demonstration of a monolithically stacked FeRAM featuring dual-gate poly-Si array transistors and horizontal $HfZrO_x$-based ferroelectric capacitors. The stacked array transistors exhibit promising electrical performance, including an Ion of 214.13 μA/μm, an SS of 76 mV/dec, and an Ion/Ioff ratio greater than 1×10^7 (with Vds = 2 V and Vgs = 3 V). The stacked horizontal capacitors show a good ferroelectric property of $2Pr\geq58$ μC/cm^2 and endurance > 1×10^8.

9:30 AM

12-2 Towards Scalable 3D Integration of 2T-nC FeRAM with Hundreds of Layer Stacking

Jiahui Duan[1], Shan Deng[1], Rudra Biswas[2], Sadik Yasir Tauki[2], Sizhe Ma[1], Rajiv Joshi[3], Thomas Kampfe[4], Xiao Gong[5], Vijaykrishnan Narayanan[2], Kai Ni[1]

[1]University of Notre Dame, [2]Pennsylvania State University, [3]IBM Thomas J. Watson Research Center, [4]Fraunhofer IPMS, [5]National University of Singapore

We study the limits of the number of capacitors and read history dependence in a 2T-nC FeRAM cell, paving the way for its high-density integration toward hundreds of stacked layers. Through a comprehensive experimental and simulation study on the scaling behavior of the 2T-nC FeRAM architecture, we demonstrate: (i) successful fabrication of 2T-64C cells with robust memory operation and clearly distinguishable '0' and '1' states, even in 64-capacitor configurations; (ii) that the parasitic capacitance of the floating node originates predominantly from the linear component of the ferroelectric capacitor, and its impact on n-scaling—due to degraded sense margin—can be mitigated by floating unselected capacitors with enough TΩ isolation; (iii) that sharing write and read transistors among n capacitors introduces a read history dependence issue due to fluctuating floating node voltage; and (iv) that a proposed FN discharge scheme can effectively eliminate read-sequence dependence, at the cost of reduced read endurance.

9:55 AM

12-3 HZO MFM with Record 45μC/cm^2 2P$_r$ @ ≤3nm, Record 60 2P$_r$/V$_{op}$, and Record >8E12 Endurance by β-W Electrodes and Initial Solid Phase Epitaxy with 1.2V NVSRAM Applications

Yu-Tsung Liao[1], Jhih-Yuan Liang[2], Tao Chou[2], Guan-Hua Chen[2], Yun-Wen Chen[2], Yu-Rui Chen[2], Wei-Jen Chen[2], Min-Kuan Lin[2], Yan-Jyun Chen[2], Yu-An Chen[2], Bo-Hui Yu[2], Hidenari Fujiwara[2], Ming-Chang Liu[2], Kuan-Heng Lin[2], Dai-Ying Lee[3], Ming-Hsiu Lee[3], C. W. Liu[2]

[1]Graduate Institute of Electronics Engineering, National Taiwan University, [2]National Taiwan University, [3]Macronix International Co., Ltd

The leakage current(2.6E^{-2}A/cm^2@+0.75V)of 3nm HZO MFM capacitor is reduced by 4orders of magnitude using water quenching and H_2 plasma due to the diffusion suppression of β-W electrode into HZO.ISPE enhances 2P$_r$ by 40%,achieving record high 2P$_r$ of 45μC/cm^2 for the thickness≤3nm and record high 2P$_r$/V$_{op}$ratio of 60 among all MFMs. At 0.75VV$_{op}$ for major loops, the 3nm HZO capacitor achieves record fatigue-free endurance exceeding 8E$_{12}$cycles with final 2P$_r$ of 45μC/cm^2. 3.5nm HZO MFM capacitors with low 2V$_c$ of 1.4V and 2P$_r$ of 27μC/cm^2 are integrated into NVSRAM cells. The reliable non-volatile recall voltage is as low as 1.2V with the store voltage of 1.2V.

10:45 AM

12-4 Direct Observation of Nanoscale Polarization Switching in HZO Ferroelectrics Using STEM-EBIC

Prasanna Venkatesan Ravindran[1], Yueyun Chen[2], Tristan O'Neill[2], Ho Chan[2], Hari Jayasankar[1], Priyankka Ravikumar[1], Zekai Wang[1], Shelby Fields[3], Megan Lenox[3], William Hubbard[4], Andrea Padovani[5], Luca Larcher[6], Gaurav Thareja[6], Shimeng Yu[1], Jon Ihlefeld[3], Wanki Kim[7], Daewon Ha[7], Brian Reagan[2], Asif Khan[1]

[1]Georgia Institute of Technology, [2]University of California Los Angeles, [3]University of Virginia, [4]NanoElectronic Imaging, [5]University of Modena and Reggio Emilia, [6]Applied Materials, [7]Samsung Electronics

We report the experimental observation of discrete ferroelectric switching down to 0.475nm×0.475nm in HfO_2-based ferroelectrics using correlated STEM electron beam induced current(EBIC) imaging and nanoscale voltage pulsing. These high-resolution measurements reveal single-domain switching events characterized by abrupt, bi-state transitions with no intermediate polarization states. By mapping polarization–voltage behavior and switching kinetics, we extract spatially resolved switching thresholds(E*min*) and speeds(ts*w*), and remanent polarization(2P*r*), uncovering scaling trends. Larger domains exhibit nearly twice the polarization and faster switching despite higher E*min*, while smaller domains switch slower at the same field. Sub-domain imaging enables quantification of variation scaling: polarization variation increases logarithmically with area, with inter-domain variation dominating intra-domain variation. Importantly, we show that multi-bit operation arises from statistical distribution of domain-level switching—rather than partial polarization—enabling distinct states in

devices as small as 0.01μm^2. These findings provide a foundational experimental framework for designing scalable, multibit ferroelectric memory based on grain- and domain-level heterogeneity.

11:10 AM

12-5 Fluid Imprint: A Reliability Bottleneck in Ferroelectric RAM Read Operation

Hyeonwoo Park[1], Jaewon Shin[2], Emmanuel Quezada[2], Hyun Jae Lee[2], Eknath Sarkar[2], Shimeng Yu[2], Asif Islam Khan[2], Supratik Guha[3], Wilfried Haensch[3], Suman Datta[2]
[1]School of Electrical and Computer Engineering, Georgia Institute of Technology, [2]Georgia Institute of Technology, [3]University of Chicago

This work presents a comprehensive analysis of fluid imprint as a critical reliability bottleneck in hafnium zirconium oxide (Hf,ZrO_2)-based ferroelectric random-access memory (FeRAM). Unlike conventional, dwell-time-dependent imprint, fluid imprint manifests rapidly under repeated non-destructive read operations, inducing asymmetric shifts in coercive fields. By examining the dependence of fluid imprint on pulse time, voltage, temperature, and duty cycle, we decouple its underlying mechanisms and reveal that oxygen vacancy ionization dynamics govern its behavior. We develop a compact, physics-based model capturing the observed logarithmic time dependence and field/temperature acceleration, integrating it into a SPICE-compatible ferroelectric capacitor platform. Monte Carlo simulations using first-order reversal curve (FORC)-extracted distributions show that fluid imprint can limit device endurance to fewer than 10^{10} cycles, even with a 10-year retention window. These findings underscore the importance of data-aware memory management strategies and establish a new physical framework for understanding fluid imprint mechanism, guiding reliability improvements in FeRAM technologies.

11:35 AM

12-6 Transfer-Free, Lattice-Matched 2D SnS_2 Interfaces for High-Performance 1 nm HZO Ferroelectric Tunnel Junctions with High J_{on} (2.5 × 10^4 A/cm^2)

Seungkwon Hwang[1,2], Hojung Jang[3], Laeyong Jung[3], Byeongjin Park[4], Seock-Jin Jeong[4], Jongwon Yoon[4], Jung-Dae Kwon[4], Yonghun Kim[4], Hyunsang Hwang[3]
[1]Department of Materials Science and Engineering, Pohang University of Science and Technology (POSTECH), [2]Energy and Environment Materials Research Division, Korea Institute of Materials Science (KIMS), [3]Pohang University of Science and Technology (POSTECH), [4]Korea Institute of Materials Science (KIMS)

We report high-performance ferroelectric tunnel junction (FTJ) memory devices enabled by a lattice-matched 2D SnS_2/HZO heterostructure that reliably scales the ferroelectric layer down to 1 nm. By directly synthesizing 2D SnS_2 interface at low temperature (350 °C), we stabilize the orthorhombic (o-) ferroelectric phase in ultra-thin HZO films while suppressing interfacial defects. This engineered heterointerface effectively modulates both tunneling barrier height and width, leading to significantly enhanced electrical performance. Devices with 30 nm × 30 nm electrodes demonstrate high ON-state current density (2.5 × 10^4 A/cm^2 at 0.5 V), large ON/OFF ratio (>3.5 × 10^3), excellent endurance (>10^{10} cycles), and robust 10-year retention at 85 °C. This scalable, transfer-free, low-temperature integration of 2D interfaces overcomes critical scaling limits of conventional HZO FTJs, advancing manufacturable, high-density nonvolatile memory technologies.

12:00 PM

12-7 HZO FeCap with Ultra-High $2P_r$ = 133 μC/cm^2, PZT-Level E_c (1.17 MV/cm @ 6 nm), and CMWε = 9.6 @ 0 V by Adopting HfO_2-TiO_2-HfO_2 Pseudomorphic Interlayer

Sangwoo Ryu[1,2], Ryun-Han Koo[3], Seungwhan Kim[3], Wonjun Shin[4], Jong-Ho Lee[3]
[1]Department of Electrical and Computer Engineering, Seoul National University, [2]R&D, SK hynix, [3]Seoul National University, [4]Sungkyunkwan University

We demonstrate that a pseudomorphic HfO_2-TiO_2-HfO_2 (HTH) interlayer tabilizes the orthorhombic phase in 6 nm thick $Hf_{0.5}Zr_{0.5}O_2$ (HZO), overcoming ferroelectric (FE) limitations. Atomic-scale lattice matching promotes coherent growth, enlarges o-phase grains(>25 nm) and and lowers domain energy barriers by 72%. Domain dynamics and noise analysis confirm enhanced switching efficiency and alignment. HTH-HZO achieves $2P_r$ = 133 μC/cm^2 and E_c = 1.17 MV/cm, matching PZT level E_c. In FeCap mode, it exhibits CMWε = 9.6 at 0 V and retains 7.7 after 10^{10} cycles. BNN simulations using these parameters achieve high inference accuracy, indicating strong potential for low-power, CMOS-compatible neuromorphic systems.

Session 13: SMB | MEMS Resonators and Filters

9:00 AM, Continental 1-3

Co-Chairs: Wan-Thai Hsu, Soundskrit and Taeko Ando, Ritsumeikan University

This session highlights the recent development on MEMS resonators and filters for various applications. The first paper presents highly stable MEMS oscillators with two orthogonal shear-mode Lame resonators. The second paper shows the design and reliability of switchable filters fabricated in a bi-layer AlScN platform. The third paper presents further frequency scaling for RF front-ends by bonding lithium niobate oxide to silicon carbide substrate. The fourth paper shows wideband PMUT by integrating multiple devices for artery ultrasound imaging application. The fifth paper, an invited paper, presents filters and resonators fabricated using microfabrication technology based on anisotropic etching of gadolinium gallium garnet. The sixth paper shows MEMS resonators operating in conductive liquids for bio and chemical applications. The last paper presents nonlinear MEMS resonators with asymmetric structures to obtain two-bit signals for analog-to-digital conversion for encryption applications.

9:05 AM

13-1 A Multi-Shear-Mode Silicon Bulk Acoustic Clock Achieving 102 ns Time Deviation at 8 Hours

Banafsheh Jabbari[1], Shruti Mishra[2], Ratul Kundu[2], Shubham Mondal[3], Ezike Everestus[2], Shaurya Dabas[2], Jinghan Gao[2], Zetian Mi[3], Baibhab Chatterjee[2], Roozbeh Tabrizian[3,3]

[1]Department of Electrical & Computer Engineering, University of Michigan, [2]University of Florida, [3]University of Michigan

This paper presents, for the first time, a silicon MEMS clock that achieves 102 ns time deviation at 8 hours. The system employs two orthogonal shear-mode Lamé resonances in a single silicon BAW resonator: a high-Q fundamental mode (78 MHz, Q=124,011) as the frequency reference, and its first overtone (143.7 MHz, Q=93,732) as an integrated temperature sensor. Duffing nonlinearity in fundamental mode enables real-time compensation of second-order thermal drift via autonomous drive-level control. The resonator uses an MBE-grown AlScN–AlN transducer on degenerately doped silicon, yielding low insertion-loss and –110 dBc/Hz phase noise at 1 kHz offset. The integrated system consumes 130 mW in a 0.25 cm^3 volume and maintains ±0.45 ppb frequency fluctuation over continuous 25°C sweeps, with ±135 ppt drift over 24 hours. Compared to miniature rubidium atomic standards, the system achieves 2–3 orders-of-magnitude lower size and power, and enables long-term frequency stability not previously demonstrated in MEMS-based clocks.

9:30 AM

13-2 Experimental Insights of ScAlN-Based Switchable Resonator and Filter: Operation and Reliability

Chen Liu[1], Ying Zhang[2], Xinghua Wang[2], You Qian[2], Huamao Lin[2], Qingxin Zhang[2], Yao Zhu[2]

[1]MEMS, Institute for Microelectronics (IME), Agency for Science, Technology and Research (A*STAR), [2]Institute for Microelectronics (IME), Agency for Science, Technology and Research (A*STAR)

In this work, $Sc_{0.3}Al_{0.7}N$-based switchable bilayer film bulk acoustic resonators (FBAR) and filters exhibiting electrically switchable bands of 3 GHz and 6 GHz were fabricated and investigated. Four polarization configurations are examined to identify the layer to switch for an optimal performance. For the filter, a 2-step switching is proposed to reduce the operational complexity and was demonstrated on the fabricated 2-stage filters, achieving the switching between a 3 GHz passband with fractional bandwidth (FBW) of ~15% and a 6 GHz passband with FBW of 8%. Additionally, a thorough endurance analysis of the bilayer FBAR was performed to characterize changes in resonator performance and ferroelectric behavior over repeated cycling, and three dominant failure mechanisms were identified. Endurance of the bilayer FBAR can be improved up to 10^7 by adopting a dynamic pulse-amplitude cycling scheme. This study provides insights into the operation, optimization, and long-term reliability of switchable ScAlN-based BAW devices.

9:55 AM

13-3 Cost-Effective Frequency Scaling of Acoustic Devices from Sub-6 GHz to *X*-Band on Pizeo-on-SiC

Peisen Liu[1], Jiajun Gao[1], Sulei Fu[1], Boyuan Xiao[1], Qiufeng Xu[1], Xinchen Zhou[1], Haoqin Ma[1], Shouren Xie[1], Zhibin Xu[2], Shuai Zhang[3], Rui Wang[1], Cheng Song[1], Fei Zeng[1], Weibiao Wang[3], Feng Pan[1]

[1]Tsinghua University, [2]Southeast University, [3]SHOULDER Electronics Limited

Frequency-scalable acoustic devices are essential for 5G/6G multi-band RF front-ends. We demonstrate a sub-6 GHz to X-band solution using a planar SAW process on LN/SiC heterostructure. By selectively exciting three modes (SH-SAW, LLSAW, high-order SV), multi-band functionality is integrated via mode-specific engineering. Resonators/filters (3-12 GHz) fabricated via single-mask KrF lithography cover 5G n77/n79, Wi-Fi 6/6E, and extend SAW applicability to X-band. This provides a low-cost, scalable pathway for broadband acoustic integration.

10:20 AM

13-4 First Codesign of 70%-Bandwidth PMUT Array with ±5V Multi-level CMOS Transmitter for Human Artery Ultrasound Imaging in Wearable Devices

Yucheng Lin[1,2,3], Tong Jin[2,4], Hongyang Chen[2,4], Tongyu Zhang[2,4], Wenchang Zhang[4], Zhiqiang Li[2,4], Chengjun Huang[2,4], Xin Liu[4], Hang Gao[4], Ling Li[2,4]

[1]State Key Laboratory of Fabrication Technologies for Integrated Circuits, Institute of Microelectronics, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]School of Advanced Interdisciplinary Science, University of Chinese Academy of Sciences, Beijing, China, [4]Institute of Microelectronics, Chinese Academy of Sciences, Beijing, China

This paper presents the first co-designed 64-channel broadband PMUT array and multi-level CMOS transmitter (TxCMOS) in 0.18μm BCD platform, achieving a record 70% bandwidth at 5MHz with only ±5V drive voltage. Our innovative integration of multi-PMUT channel architecture and novel multi-level pulser enables high transceiving efficiency for wearables, including 8.4 kPa/V/mm^2 for transmission and 1.2 mV/V for self-reception sensitivity. Integration with a CMOS low-noise amplifier (LNA) boosts received signals to 60mVpp signals (6x voltage gain). Critically, we demonstrate—for the first time—human deep artery imaging via TxCMOS-PMUT, capturing M-mode ultrasound imaging of the brachial artery (5 mm diameter at 15 mm depth). The system achieves 19 μm resolution in simultaneous upper and lower arterial wall tracking, enabling cardiac-cycle-correlated blood pressure calculation. This breakthrough establishes the broadband TxCMOS-PMUT as a transformative technology for wearable, non-invasive cardiovascular monitoring and high-resolution deep-tissue imaging.

11:10 AM

13-5 Modern YIG Micromachining Techniques for Tunable Microwave Components (Invited)

Sunil A Bhave[1], Connor Devitt[2], Sudhanshu Tiwari[2], Anuj Ashok[2], Yiyang Feng[2], Sen Dai[2]

[1]Elmore Family School of Electrical and Computer Engineering, Purdue University, [2]Purdue University

This paper reviews recent advances in the microfabrication techniques for tunable magnetostatic wave (MSW) resonators and filters. Micromachining based on ion milling, conformal metalization, and anisotropic etching of gadolinium gallium garnet (GGG) have enabled chip-scale, tunable, high performance bandstop filters, bandpass filters, and MSW resonators with a maximum $f_0 \cdot Q$ product of 9.376×10^{13} at f_0 = 12.99 GHz with a Q-factor of 7217.

11:35 AM

13-6 Modular Integration of Magnetic and Bottom-Electrode Strategies for High SNR MEMS Resonators in Conductive Liquids

Zhong-Wei Lin, Anurag A. Zope, Sheng-Shian Li

National Tsing Hua University

This work presents a robust strategy for MEMS resonators in conductive liquids. Bottom-electrode signal routing addresses electrical leakage but introduces parasitic feedthrough. We propose a PCB-level electromagnetic feedthrough cancellation technique, decoupled from the MEMS device, that suppresses this feedthrough. This integrated approach significantly improves stopband rejection in saline by three-fold (from ~5 dB to ~15 dB) and enhances resonance clarity. It offers a modular, practical solution for high-fidelity MEMS operation in challenging fluidic environments for biological and chemical sensing.

12:00 PM

13-7 Using Nonlinear CMOS-MEMS Resonators for Self-Encrypting Analog-to-Digital Conversion

Ting-Yi Chen, Wei-Chang Li
National Taiwan University

This work demonstrates a nonlinear MEMS resonator-based analog-to-digital converter (ADC) integrated with an intrinsic hardware encryption function. In particular, the resonator exhibits internal resonance-induced asymmetric frequency combs whose profiles are digitized into binary bits. Due to the high sensitivity of combs to inherent physical variation during the microfabrication process, the MEMS-based ADC can uniquely map the input voltage level into hardware-encrypted bits. The resonator, together with a 16-bit digitization, converts 9 comb teeth into unique 144-bit encrypted digital outputs with a voltage resolution down to 0.006 V. The generated bits exhibit randomness from device to device, confirmed by the inter-Hamming distance (inter-HD) analysis and the NIST 800-22 statistical tests. The proposed resonant MEMS-based ADC provides an intrinsic hardware security solution that eliminates the need for mode-switching or external database management, paving the way for more compact and secure sensor networks.

Session 14: EDT/SMB | Focus Session | The Invisible Revolution in Thin-Film Transistors

9:00 AM, Continental 4

Co-Chairs: Elena Gnani, University of Bologna and Pierpaolo Palestri, University of Modena and Reggio Emilia

When it comes to transistors, high-performance CMOS as well as power devices often dominate attention, while thin-film transistors (TFTs) receive less focus. Yet, their diverse materials and fabrication processes enable highly versatile and precise applications. Their inherent physical flexibility further broadens the range of possible implementations, making TFTs uniquely adaptable to emerging technologies. Ongoing research worldwide is advancing performance with new materials, improving efficiency and sustainability through material-device-circuit co-design, and adding novel functionalities to established technologies. This special session aims to foster cross-disciplinary discussion to accelerate progress by sharing ideas and best practices across largely independent research communities. Innovations in materials, deposition, and device architectures are expanding TFT applications beyond displays and imaging, into fields such as microfluidics, neuromorphic computing, disposable sensors, and human–machine interaction. The broad range of available options makes TFTs increasingly viable as ecological and commercially sound alternatives to mainstream CMOS technology.

9:05 AM

14-1 Recent Advances in Digitally Printed Organic TFTs (Invited)

Hiroyuki Matsui[1], Rei Shiwaku[2], Kazuki Ono[2], Satoru Inoue[2], Yudai Hemmi[2], Mika-Matti Laurila[3], Tomohito Sekine[2], Itsuki Shoji[2], Hideki Wada[2], Tatsuo Hasegawa[4], Radu A. Sporea[5], Matti Mäntysalo[3], Kuniaki Nagamine[2], Toshiyuki Sugimoto[2], Shizuo Tokito[2]
[1]Research Center for Organic Electronics, Yamagata University, [2]Yamagata University, [3]Tampere University, [4]The University of Tokyo, [5]University of Surrey

Digital printing, such as inkjet, is a promising fabrication process for both graphics and electronic circuits due to its minimal material loss and rapid prototyping capabilities. Soluble organic semiconductors and dispersible metal nanoparticles are well-suited for such digital printing techniques. This paper reviews recent advancements in digitally printed organic thin-film transistors, covering their device performances and circuit applications.

9:30 AM

14-2 Embracing Energy Barriers for Improved Stability in Organic Thin-Film Transistors (Invited)

Radu Alexandru Sporea[1,2], Eva Bestelink[3], Andreas Petritz[4], Christine Prietl[4], Herbert Gold[4], Bernhard Kleinburger[4], Vlad Stolojan[3], Ahmed Nejim[5], Barbara Stadlober[4]
[1]Advanced Technology Institute, University of Surrey, [2]Institute for Sustainability, University of Surrey, [3]University of Surrey, [4]Joanneum Research Forschungsgesellschaft mbH, [5]Silvaco Europe Ltd.

This work demonstrates the benefits of low, but dominant, contact barriers in organic thin-film transistors (OTFTs) for threshold stability. Multimodal transistor measurements (≤ -8 V) confirm contact-controlled operation even when single-gate OTFTs produce apparent Ohmic behavior ($dV_{SAT} / dV_{GS} \approx 0.85$). Simulations demonstrate that the contact-controlled nature of the device (contact energy barrier $\Phi_B = 0.4$ eV for hole mobility 2.7 $cm^2V^{-1}s^{-1}$) promotes

drain current stability with respect to interface charge-induced threshold variations, a finding with direct implications on emerging OTFT applications e.g. displays where simplified pixel design may lead to competitive new implementations. The stability and saturation properties also make such devices attractive to life sciences and distributed sensors applications.

9:55 AM

14-3 Fully Printable Oxide TFTs and Memristors on Rigid and Flexible Substrates with Excellent Integration Properties for Hybrid Systems (Invited)

Jasmin Aghassi-Hagmann[1], Shawon Alam[2], Xu Wang[3], Hongrong Hu[2], Alexander Scholz[2,4], Axel Neffe[3], Sehun Seo[3], Francesca Toma[3], Gabriel Marques[5]
[1]Institute of Nanotechnology, Karlsruhe Institute of Technology, [2]Karlsruhe Institute of Technology, [3]Helmholtz-Zentrum Hereon, [4]Infineon Technologies, [5]Karlsruhe Institute of Technology (KIT)

Recent advances in inorganic oxide electronics exploiting metal oxides as semiconducting channel materials show a rich portfolio of functional electronic devices and integration possibilities enabling future applications in security, robotics, the display and health sector. We report on fully additively manufactured, composite polymer electrolyte gated in indium (In_2O_3) oxide thin film transistors combining aerosol-jet and inkjet printing as well as integrated circuits such as physical unclonable functions, ring oscillators and hybrid flexible systems for temperature sensing. In the more complex systems, we show that our printed transistor can be integrated also with silicon discrete components as well as thinned ICs and work reliably at low voltages (<2V). Finally, we present results on printed resistive switching devices (memristors) with fast switching speeds (~ns) and high R_{on}/R_{off} values (~10^7) which will further boost the potential of all printed oxide circuits in the future.

10:20 AM

14-4 Ultra-Precise Dispensing: Microscale Additive Manufacturing for Next-Generation TFTs (Invited)

Filip Granek
XTPL, Poland

XTPL's Ultra-Precise Dispensing enables microscale additive manufacturing for next-generation thin-film transistors (TFTs). With micrometer resolution, high-viscosity ink compatibility, and direct-write deposition on complex 3D substrates, UPD addresses critical challenges in microelectronics. It significantly enhances performance, reduces manufacturing complexity, and opens new design opportunities for future electronics.

11:10 AM

14-5 From Flexible to Green Electronics (Invited)

Marco Fattori, Kyle van Oosterhout, Martijn Timmermans, Enrico Genco
TU Eindhoven

Flexible electronics has achieved remarkable electrical performance and technological maturity, enabling groundbreaking display innovations and promising further applications soon. In this paper, we present our latest achievements in the field and offer a forward-looking perspective on the integration of sustainable materials and the development of solutions towards lower environmental impact and "green electronics".

11:35 AM

14-6 Interface-Engineered Al:IZTO/Tellurium Heterojunction TFTs with Ultra-Thin Al_2O_3 for Tunable Double Turn-On and NDR Characteristics in Logic Applications (Invited)

Sooji Nam[1,2], Jung Hoon Han[1,3], Mirinae Lee[1,2], Chahwan Yang[1,2], Byeong-Kwon Ju[3]
[1]ETRI, South Korea, [2]University of Science and Technology, [3]Korea University

We propose an all-oxide p-n heterojunction thin film transistor (H-TFT) based on Tellurium (Te)/Al-doped In Zn-Sn-O (Al:IZTO) stack with an ultra-thin (sub-2 nm) Al_2O_3 interlayer. In this architecture, n-type TFTs are demonstrated with Te functioning as an electron-blocking and hole conducting layer, and Al_2O_3 acting as a tunneling barrier and interface passivation layer, enhancing interfacial control and Te crystallinity. By engineering these layers, we achieve threshold voltage (V_{TH}) shifts of up to +20 V with increasing Te thickness (4–8 nm), second turn-on modulation via Al doped In–Zn–Sn–O (Al:IZTO) thickness (10–20 nm) with peak-to-valley ratio (PVR) exceeding 100, and distinct negative differential resistance (NDR) in long channel devices. A ternary logic inverter is configured

using this structure, exhibiting three distinct output states (Logic 0, 1, 2) based on the device's dual turn-on behavior, highlighting its potential for MVL applications using fully oxide-based devices.

12:00 PM

14-7 From Display Technology to Life Science Innovations: Exploring Novel TFT Applications (Invited)

Linrun Feng[1], Jun Li[2], Zhe Liu[1], Ye Xiao[1], Wei Tang[2], Xiaojun Guo[2]
[1]Hangzhou LinkZill Technology Co. Ltd., [2]Shanghai Jiao Tong University

In this work, a DNA synthesis biochip with 256 × 256 resolution fabricated by using indium-gallium-zinc-oxide (IGZO) thin-film transistor (TFT) technology is presented. A TFT-based gate-on-array (GOA) is also integrated to provide addressing capability for active-matrix driving of 65K pixels. The implemented biochip demonstrated high-throughput synthesis of high-quality oligonucleotides, ranging from 20 to 150 nucleotides in length at femtomole quantities, with optimization of electrode design and driving schemes. The platform maintains compatibility with standard TFT manufacturing processes while eliminating the need for complex photochemical control systems inherent to conventional approaches. This work paves the way for cost effective, large-scale biochip fabrication, with potential applications in synthetic biology.

Session 15: EDT | Emerging Spintronic and Cryogenic Devices

9:00 AM, Continental 5

Co-Chairs: David J. Michalak, TNO and Ching-Tzu Chen, IBM

This session includes 6 papers in new areas of spintronics and cryogenic devices. The first two papers focus on spintronic devices. The first paper is from Maxwel Gama Monteiro from imec and presents field-free switching perpendicular SOT-MRAM 4kb array using out-of-plane polarized spin current. The second paper, by Shuaiyu Gong from the Chinese Academy of Sciences, presents BEOL-compatible HZO/Pt-nanocluster-based magnetoelectric devices that enable voltage-controlled magnetization switching with ultra-low energy. The last four papers focus on cryogenic devices spanning superconducting digital logic circuits, cryogenic control devices, and fabrication and high-volume characterization of quantum computing devices. The third paper, by Ankit Pokhrel from imec, presents advancements in superconducting digital logic made from NbTiN materials. The fourth paper, by Erbing Hua from TU Delft, reports on low-noise cryogenic memristors for spin qubit voltage control. The fifth paper, by Sofie Beyne at imec, reports on characterization of silicon-based spin qubit devices fabricated using 0.33NA EUV lithography. The last paper, by Marco Candido from University of New South Wales, demonstrates chip-level automated cryogenic screening of spin qubit devices fabricated at a 300mm fabrication facility.

9:05 AM

15-1 First Demonstration of Field Free Switching Perpendicular SOT-MRAM 4kb Array Using Z-Spin

Maxwel Gama Monteiro Junior[1], Ankit Kumar[1], Vaishnavi Kateel[1], Bob Vermeulen[1,2], Birte Coester[1], Jyotirmoy Chatterjee[1], Giacomo Talmelli[1], Alvaro Palomino Lopez[1], Joseba Urrestarazu-Larranaga[1], Simon Van Beek[1], Kurt Wostyn[1], Siddharth Rao[1], Van Dai Nguyen[1], Gouri Sankar Kar[1]
[1]imec, [2]KU Leuven

In a fully integrated 300 mm platform for perpendicular Spin-Orbit Torque Magnetoresistive RAM (SOT-MRAM), we demonstrate field-free functionality in a 4kb characterization array. Our results show that a stack design of a ferromagnet (FM) / Ta bilayer track enables Z-Spin based switching on SOT-MRAM devices in the <10ns regime. Furthermore, our devices show tunneling magnetoresistance (TMR) > 90% and retention >10 years, which is comparable to a standard field assisted SOT device. The devices integrated in 4kb sub-bank array configurations show excellent distributions and low variability between dies, paving the way for advanced memory applications of SOT-MRAM.

9:30 AM

15-2 Voltage-Driven HZO/Pt-nanocluster-based Magnetoelectric Devices with Low Thermal-Budget (400℃) and Ultra-Low Energy Consumption (0.46pJ/μm^2)

Shuaiyu Gong[1,2,3], Yanru Li[4,4], Xiaopeng Li[5], Shuo Xu[4,4,6], Xiaoyu Dou[5], Ying Li[7], Jianfeng Gao[4,4,6], Weibing Liu[4,4], Peiyue Yu[4,4], Jingsong Huang[4,4,6], Jixuan Wu[5], Meiyin Yang[4,4,6], Dashan Shang[4,4,6], Jiezhi Chen[5], Jun Luo[4,4,6]
[1]Key Laboratory of Fabrication Technologies for Integrated Circuits, Chinese Academy of Sciences, [2]Institute of

Microelectronics, Chinese Academy of Sciences, [3]School of Integrated Circuits, University of Chinese Academy of Sciences, [4]Chinese Academy of Sciences, [5]Shandong University, [6]University of Chinese Academy of Sciences, [7]Hebei University of Technology

Conventional current-driven spin torque devices suffer from high energy consumption, though voltage-controlled magnetoelectric (ME) alternatives enable ultralow-energy operation, their CMOS integration remains constrained by special single-crystal substrates (ABO_3) and high processing temperatures (~700 °C). Here, we developed a CMOS-compatible ME device with $Hf_{0.5}Zr_{0.5}O_2$ (HZO)/Pt-nanocluster heterostructure, which is fabricated at 400°C on silicon wafer. The Pt-nanocluster enhances the ME coupling by reducing the dissociation energy barrier of Co-O bonds at lateral Pt/Co interfaces. This mechanism enhances oxygen mobility during ferroelectric polarization switching in HZO, enabling voltage-controlled magnetization reversal with ultra-low energy consumption (0.46 $pJ/\mu m^2$) – significantly below spin torque devices (100-1000 $pJ/\mu m^2$). This establishes a critical pathway for CMOS-compatible, low-power magnetoelectric logic and memory integration.

9:55 AM

15-3 Advancements in NbTiN Based Circuits for Superconducting Digital Logic

Ankit Pokhrel[1], Benjamin Huet[1], Blake Hodges[1], Seifallah Ibrahim[1], Ruiheng Bai[2], Steven Brebels[1], Trent Josephsen[1], Daan Buseyne[3], Gilles Delie[1], Diziana Vangoidsenhoven[1], Adham Elshaer[1], Daniel Perez Lozano[1], Ravi Tiwari[1], Sujan K. Sarkar[1], Luke Zhang[1], Steven Deckers[1], Hasantha Malavipathirana[1], Julian G. Pinzon[1], Peter Koufalis[1], Rishi Mishra[2], Cequn Li[2], Sabine O'Neal[1], Katja Nowack[2], Margriet Van Bael[3,3], Anne-Marie Valente-Feliciano[4], Manu Perumkunnil[1], Zsolt Tokei[1]

[1]imec, [2]Cornell Univeristy, [3]KU Leuven, [4]Thomas Jefferson National Accelerator Facility

Superconducting (SC) electronics have emerged as a promising platform for high-speed, energy-efficient computing and quantum information processing. This work, centered on NbTiN, presents recent advances in material science and fabrication methods leading to significant improvements in performance, scalability and vertical integration. We specifically report on fabrication and characterization of key components, including Josephson junctions (JJs), flux trapping structures and SC interconnects. Together, these efforts represent critical steps towards realizing practical, complex, dense and large-scale SC integrated circuits.

10:45 AM

15-4 Ultra-Low-Noise Cryogenic Multilevel Memristors for Sub-100 µV Spin Qubit Biasing

Erbing Hua[1], Stijn Heemskerk[2], Nikolaj Nitzsche[2], Wagno Bragança[2,2], Hanzhi Xun[2], Ali Kaichouhi[2], Ryoichi Ishihara[2,2]

[1]Quantum and Computer Engineering, Delft University of Technology, [2]Delft University of Technology

We demonstrate interface-enhanced memristors (OxReRAM) tailored for cryogenic spin-qubit control. By engineering a sparse filament network, our devices achieve eight nonvolatile resistance levels with an ultra-low read noise rate of around 0.3 %. When embedded in a cryogenic gain stage with R_L = 30 kΩ and V_{in} = 0.3 V, it will deliver a ±1 V output range and sub-100-µV resolution using only six memristors per channel. This single-line biasing architecture will reduce wires, paving the way for large-scale quantum processors.

11:10 AM

15-5 300mm fabrication of silicon quantum dot spin qubits using 0.33NA EUV lithography

Sofie Beyne[1], Stefan Kubicek[1], Clement Godfrin[1], Bart Raes[1], Thomas Van Caekenberghe[1,2], Johan De Backer[1], Yannick Hermans[1], Fergo Treska[1], Shuchi Kaushik[1], Yuchao Jiang[1], Bianca Chou[1], Tanushree Sarkar[1], Yanan Li[1], Sugandha Sharma[1], Sylvain Baudot[1], Luis Francisco Pinotti[1], Yosuke Shimura[1], Roger Loo[1,3], Bart Kenens[1], Jelle Vande Weeghde[1], Evi Vrancken[1], Kevin Vandersmissen[1], Jerome Mitard[1], Katia Devriendt[1], Danny Wan[1], Kristiaan De Greve[1,2]

[1]imec, [2]KU Leuven, [3]Ghent University

We demonstrate a 300mm process for 2 types of SiMOS quantum dot spin qubit device architectures enabled by EUV lithography. First, we show an overlapping gates process for spin qubits, where the gates are fabricated for the first time using 0.33NA EUV lithography. We report excellent reproducibility, full wafer room temperature functionality and good quantum dot and qubit metrics at 10mK. Second, to enable qubit scalability, we

demonstrate a single-layer gate device architecture fabricated with 0.33NA EUV patterning of the gate, and 2 damascene EUV BEOL layers.

11:35 AM

15-6 Investigation of 300mm Process SiMOS Spin Qubit Device Uniformity with Automated Cryogenic Probing

Marco Candido[1,2], Ensar Vahapoglu[2,3], Juan Pablo Dehollain[4,5], Kevin Simoes[4,5], Hyma Harish Vallabhapurapu[2], Kok Wai Chan[2,3], Tuomo Tanttu[2,3], Paul Steinacker[2,3], Will Gilbert[2,3], Santiago Serrano[2,3], Fay E. Hudson[2,3], Thomas Ohki[4,5], David Reilly[4,5], Chih Hwan Yang[2,3], Andre Saraiva[2], Christopher C. Escott[2], Arne Laucht[2,3], Wee Han Lim[2,3], Andrew S. Dzurak[2,3], Nard Dumoulin Stuyck[2,3]
[1]Sydney, University of New South Wales, [2]Diraq, [3]University of New South Wales, [4]Emergence Quantum, [5]University of Sydney

Industrial foundry processing of semiconductor spin qubits has revolutionized the scale and yield of quantum processors fabrication, which in turn enables the use of data-driven process optimization and device screening. This requires high-throughput electrical characterization at cryogenic temperatures with complex measurement protocols. Here, we present an investigation of the performance of SiMOS based spin qubit devices made in an optimized 300mm process flow. We showcase how the qubit device uniformity allows efficient characterization using a die-scale cryogenic prober at an effective electron temperature ($T_{e,eff}$) of 0.8 ± 0.2 K, which is below the temperature threshold for high-fidelity qubit operation. We leverage machine learning for autotuning the devices and extract key metrics for charge sensor and quantum dot performance for 16 double-dot devices. Our results enable further process optimization as well as variability-aware design of larger SiMOS spin qubit processors.

Session 16: PMA | High Performance RF FET and HBT Technologies

9:00 AM, Continental 6

Co-Chairs: Vibhor Jain, GlobalFoundries and Colombo Bolognesi, ETH Zurich

The session includes 7 papers focusing on high speed and high performance RF FETs, HEMTs and HBTs based on III-V and Si/SiGe material system. The papers collectively showcase significant advancements in SiGe BiCMOS, InP HBTs, GaN HEMTs and 3D RF integration technologies, each addressing critical aspects of high-speed communication systems. The first paper by Uppili Raghunathan from GlobalFoundries presents a breakthrough in complementary bipolar CMOS integration, reporting the highest-ever PNP fT/fMAX of 210/240 GHz alongside NPN fT/fMAX of 420/430 GHz. The second paper is an invited talk by Dr. Rucker introducing IHP's 130 nm BiCMOS platform, achieving record-breaking SiGe HBT performance with peak fT/fMAX of 500/610 GHz. The third paper by Thibaud Fache from CEA-Leti demonstrates a novel 3D sequential integration approach combining standalone SiGe HBTs with low-temperature RFSOI switches and local trap-rich substrates paving the way for fully monolithic RF front-end modules. The next paper by Thanh Viet Dinh from NXP presents a novel 3.3V FinFET-based RF-LDMOS with a record fT/fMAX of 60/240 GHz and robust reliability, achieved through innovative dummy gate field-plate configurations and P-well engineering. Moving to III-V devices and InP HBTs, the next paper by Filippo Ciabattini from ETH Zürich showcases transferred-substrate InP/GaAsSb DHBTs on silicon, delivering over 2.2 W/mm output power and >38% PAE at 140 GHz. The next paper by Jaeyong Jeong from KAIST, Korea introduces a heterogeneous 3D (H3D) RF platform that monolithically integrates InGaAs and GaN HEMTs and stacks them onto CMOS using low-temperature die-to-wafer bonding achieving exceptional RF performance. The final paper on RF GaN HEMTs will be presented by Chandrashekhar Savant from Cornell University reporting the first demonstration of ultrathin epitaxial AlBN barrier GaN FerroHEMTs with ferroelectric gating achieving RF performance of fT/fMAX=76/139 GHz.

9:05 AM

16-1 A High Performance SiGe Complementary Bipolar-CMOS Technology for Efficient Communication

Uppili S Raghunathan, Steven Shank, Sarah McTaggart, Megan Lydon-Nuhfer, Aaron Vallett, Cameron Luce, Ramsey Hazbun, Daisy Dillon, Stephen Bosley, Saurabh Sirohi, Jui-Chu Lee, Adam Divergilio, Kim Newton, Vasudeva Kuppireddy, Hari Kishore Kakara, Jeffrey B Johnson, Alexander Derrickson, Vibhor Jain, John Jack Pekarik
GlobalFoundries

High-Performance SiGe HBTs with PNP fT/fmax of 210/240 GHz and NPN fT/fmax of 420/430 GHz are reported here. This is the highest fT measured for a PNP HBT, surpassing the previous reported result by 20%. Together with the

NPN and CMOS, these are the highest values simultaneously reported for co-integrated HBTs in a Complementary-Bipolar-CMOS(CBiCMOS) process, with fT/fmax over twice the previously-published results. The active-load PNP HBT reduces 40% power in TIAs and improves 32% bandwidth in linear drivers compared to NPN-only circuits. Additionally, the improved NPN HBT has high beta (5000 points) for better noise figure (0.1-0.15dB) over previous generations.

9:30 AM

16-2 SiGe-BiCMOS Technology with 500 GHz f_T and 610 GHz f_{MAX} (Invited)

Holger Rücker, Bernd Heinemann, Florian Bärwolf, Walter Batista Pessoa, Martin Drost, Matteo Elviretti, Fabian Fidorra, Gerhard G. Fischer, Alexander Fox, Oksana Fursenko, Dirk J. Hagen, Elena Hasler, Anne Hesse, Falk Korndörfer, Andreas Krüger, Philipp Kulse, Thomas Lenke, Marco Lisker, Rasuole Lukose, Steffen Marschmeyer, M. Andreas Schubert, Sebastian Schulze, Florian Teply, Andreas Trusch, Christian Wipf, Dirk Wolansky, Yuji Yamamoto
IHP - Leibniz Institute for High Performance Microelectronics

This paper reports on process development and advances in the high-frequency performance of SiGe BiCMOS technologies. It presents a 130 nm BiCMOS technology with high-speed SiGe HBTs that offer peak f_T/f_{MAX} values of 500 GHz/610 GHz, a BV_{CBO} of 3.6 V, and CML ring oscillator gate delays of 1.38 ps. These are the fastest HBTs available in any SiGe BiCMOS technology. In addition to high-speed HBTs, CMOS, and RF passive components, the technology integrates HBTs for higher voltages with f_T/BV_{CBO} of 280 GHz/5.6 V and 185 GHz/7.1 V, respectively.

9:55 AM

16-3 Unlocking High-Performance Si RF Platforms with SiGe HBT and RFSOI Switch Technologies

Thibaud Fache[1], Thomas Bordignon[2], Martin Perrose[1], Luca Lucci[1], Daniel Gloria[2], Obaid Adami[1], Frederic Monsieur[2], Pascal Gourod[2], Sebastien Kerdiles[1], Francois-Xavier Darras[1], Pablo Acosta-Alba[1], Magali gregoire[2], Adrien Estellon[2], Fabrizzio Morris Anak[2], S Boughaleb[2], Jerome Prouvée[1], Olivier Richard[1], Nicolas Guitard[2], Blend Mohamad[1], A Pavlovski[1], Mickal Ribotta[1], Laurent lachal[1], Xavier Garros[1], Perrine Batude[1], Pascal Chevalier[2]
[1]CEA-Leti, [2]STMicroelectronics

This work demonstrates key solutions for the ultimate cointegration of SiGe HBT, PDSOI switches and high-quality passives with 3D sequential integration. A Local Trap Rich (LTR) with RF isolation comparable to the state-of the-art TR substrate, and a bulk SiGe HBT withstanding the top tier thermal budget are shown. A high-performance RFSOI switch processed at 600°C featuring $R_{ON}\times C_{OFF}$=96fs & RF_{Vmax}=2.7V is also demonstrated, which is a major breakthrough. Together, these results pave the way towards a fully monolithic Silicon Front-End Module (FEM) for advanced efficient and cost-effective RF systems.

10:20 AM

16-4 Novel 3.3V RF-LDMOS in 16nm FinFET with the Best RF and Reliability Trade-off Enabling Integrated Watt-level Power Amplifiers in Connectivity SoC

Thanh Viet Dinh[1], Nick Pronin[1], Alessandro RUGGIERI[2], Thiago Hanna Both[1], Asanga Perera[1], Oliver Dieball[1], Ralf Pijper[1], Maarten Swanenberg[1], Juan Xie[1], Lisa Tondelli[2], Luca Selmi[2]
[1]NXP Semiconductors, [2]UniMORE

A novel FinFET-based LDMOS device with a breakdown voltage of ~11V and a record RF performance, with f_{MAX}~240 GHz and f_T~60 GHz, is presented. The device is composed of a thick gate oxide in the channel and uniquely-connected dummy gates in the drain extension to optimize the electric field for reducing device degradation due to hot carrier injection. The device reliability is further improved by using a baseline mask to partly block the P-well implant for optimizing channel-drift junction. The device concept has been intensively assessed by TCAD and fabricated in a 16nm FinFET node. A 5 GHz power-amplifier utilizing this novel LDMOS has been designed and taped-out. On-wafer measurements were done to characterize DC, RF and reliability of these new devices. The measured characteristics, both on device and circuit level, confirmed the breakthrough RF performance and robustness of this new LDMOS compared to state-of-the-art LDMOS in FinFET processes

11:10 AM

16-5 InP/GaAsSb Double Heterojunction Bipolar Transistors on Silicon Delivering $P_{OUT,SAT}$ > 2.2 W/mm and PAE > 38% at 140 GHz

Filippo Ciabattini[1], Sara Hamzeloui[2], Akshay Mahadev Arabhavi[2], Giorgio Bonomo[2], Mojtaba Ebrahimi[2], Andrea Cercaci[2], Olivier Ostinelli[2], Colombo Bolognesi[2]
[1]Department of Information Technology and Electrical Engineering, Eidgenössische Technische Hochschule Zürich, [2]Eidgenössische Technische Hochschule Zürich

We present the development of InP/GaAsSb DHBTs transferred onto silicon using Au-Au wafer bonding. Compared to conventional DHBTs on InP, these transferred devices maintain similar DC current gain but achieve a significantly wider DC safe-operating-area, increasing from 27 to 48 mW/μm^2 due to ~30–40 ℃ lower junction temperatures. Small-signal measurements confirm comparable fT/fMAX performance. G-band load-pull tests show enhanced large-signal capabilities: single-finger DHBTs reach 15.2 dBm output (2.2 W/mm, 38.4% PAE) at 140 GHz, while multi-finger designs deliver 18.7 dBm (2.47 W/mm). These results demonstrate DHBTs-on-Si can rival or surpass traditional power amplifiers, offering high efficiency with silicon integration benefits.

11:35 AM

16-6 Heterogeneous 3D Integrated RF Platform: III-V/III-N RF Blocks transfer onto CMOS via Die-to-Wafer Bonding Technology for High-Frequency and High- Density Wireless Communication Systems

Jaeyong Jeong, Chan Jik Lee, Yoon-Je Suh, Nahyun Rheem, Bong Ho Kim, Joon Pyo Kim, Juhyuk Park, Sanghyeon Kim
KAIST

In this work, we report the first demonstration of a heterogeneous 3D (H3D) integrated RF platform that monolithically integrates InGaAs and GaN HEMTs as a single RF block, further stacking this RF block onto a CMOS wafer with die-to-wafer bonding at temperatures as low as 150°C. This approach preserves the intrinsic characteristics of both the III-V/III-N RF block and CMOS. The resulting H3D RF platform enables compact, high-performance RF front-end modules capable of both high-power and low-noise operation. The fabricated InGaAs HEMT achieves fT/fMAX of 320/293 GHz, while GaN HEMT exhibits fT/fMAX of 108/115 GHz. These results confirm the feasibility of monolithic III-V/III-N integration and H3D stacking with CMOS at low temperatures, highlighting the integration potential and scalability of the H3D RF platform for advanced, high-density RF solutions.

12:00 PM

16-7 Epitaxial Ultrathin Hi-K AlBN/GaN FerroHEMTs

Chandrashekhar Prakash Savant[1], Kazuki Nomoto[2], Thai-Son Nguyen[2], Wenwen Zhao[2], Stephen Daniel Funni[2], Siyuan Ma[2], Madhav Ramesh[2], Judy Jeeyong Cha[2], Huili Grace Xing[2], Debdeep Jena[2]
[1]Department of Materials Science and Engineering, Cornell University, [2]Cornell University

We report the first observation of ferroelectric gating in ultrawide bandgap, high-K, crystalline AlBN barrier GaN high electron mobility transistors (FerroHEMTs) using the thinnest ferroelectric nitride barrier to date. We observe a strong effect of gate length scaling from L_g = 1.5 μm to 44 nm on the ferroelectric gating performance of AlBN/GaN FerroHEMTs. While the short L_g = 44 nm gate length devices exhibit ferroelectric gating with large repeatable hysteretic I_d - V_{gs} loops and 40 mV/decade sub-Boltzmann characteristic, the devices with longer L_g = 800 nm to 1.5 μm exhibit little to no hysteresis. We report the first RF performance of AlBN/GaN HEMTs with f_T/f_{MAX} = 76/139 GHz, introducing first epitaxial ultrawide bandgap, high-K, and ferroelectric AlBN barrier technology to RF and mm-wave electronics.

Session 17: MS | Thermal Analyses for Novel Transistors, Stacked Designs and Power Devices

9:00 AM, Continental 7-9

Co-Chairs: Divya Prasad, AMD and Mohammad Hasanuzzaman, IBM Research

In this session, the first two papers discuss innovations in the arena of wide-band-gap transistors. In particular, the first paper discusses detailed gate-stack engineering techniques, and the second paper covers deep-learning accelerated electro-thermal modeling framework for wide-band-gap power devices. The next two papers highlight thermal challenges and excellent modeling techniques in advanced packaging schemes in both 2D and 3D

systems. The next paper covers thermal evaluation of modern device architectures like CAA and GAA vertical channel transistors. Finally, the last paper extends to reliability, defect dynamics and aging in TFTs.

9:05 AM

17-1 Gate Stack Engineering and Carrier Dynamics of Optically Triggered GaN Transistors on 650V GaN-on-Si Power Platform

Rafid Hassan Palash[1], Jung-Han Hsia[2], Soumik Saha[3], Nittya Ananda Biswas[3], Qingyun Xie[2], Victor Moroz[4], Tomás Palacios[2], Nadim Chowdhury[3]
[1]Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, [2]Massachusetts Institute of Technology, [3]Bangladesh University of Engineering and Technology, [4]Synopsys, Inc.

This work investigates the physics and engineering of the gate stack in GaN-based optically-triggered-transistors (OTTs) on 650V GaN-on-Si power platform to enable, for the first time, optical triggering at longer wavelengths (≥395nm), thereby overcoming the intensity and efficiency limitations of traditional triggering wavelengths (≤365nm) of GaN OTTs. An experimentally calibrated TCAD model, combined with a process–device–system level simulation framework, is employed to analyze the influence of gate stack design on OTT performance. The device physics of conventional p-GaN/AlGaN/GaN gate-stack in lateral OTTs are systematically explored focusing on optical illumination patterns, metal/p-GaN contact type (Ohmic/Schottky), gate connection topology (floating/source-connected), and carrier dynamics. A novel p-InGaN/p-GaN/AlGaN/GaN gate stack for lateral GaN OTTs is proposed and analyzed, addressing the triggering wavelength and response time limitations of conventional p-GaN/AlGaN/GaN gate-stack. The proposed gate stack demonstrates a high responsivity of 6.7×10^{3}A/W while reducing the fall time by 5.75× under the illumination of 395nm UV source.

9:30 AM

17-2 A Deep Learning-Accelerated Unified Electro-Thermal Modeling Framework for Wide-Bandgap Power Devices

Cheng Zhang, Wenyu Lu, Hang Su, Xuetong Zhou, Liujun Long, Yunheng Hu, Jiayu Zeng, Li Zheng, Xinhong Cheng
Shanghai Institute of Microsystem and Information Technology

This paper presents a Physics-Informed Graph Neural Network (PIGNN) for multi-scale electro-thermal modeling of WBG devices. The framework achieves >5000× speedup over Sentaurus while adapting to new geometries/biases without retraining. Validated through GaN field-limiting ring optimization and SiC/GaN SPICE parameter extraction, it demonstrates superior stability and efficiency over conventional methods.

9:55 AM

17-3 Breaking Thermal Bottleneck in 3D HBM-on-GPU Integration via System-Technology Co-Optimization

Yukai Chen, Melina Lofrano, Diksha Moolchandani, Herman Oprins, Geert Van der Plas, Julien Ryckaert, Dwaipayan Biswas, James Myers
IMEC

Stacking High Bandwidth Memory directly atop GPUs offers greater bandwidth and package density but introduces thermal challenges. This work presents a comprehensive thermal analysis and co-optimization strategy for realistic 3D HBM-on-GPU integration. Using a unified simulation framework, we evaluate mitigation techniques across technology and system levels, including HBM base die removal, adjacent stack merging, frequency scaling, and double-sided cooling. Results show that GPU peak temperature can be reduced from 141.7°C to 70.8°C, with potential for improved throughput density (performance per unit package area) over 2.5D. This study highlights the importance of STCO and demonstrates the viability of HBM-on-GPU 3D stacking.

10:20 AM

17-4 Transistor-to-Package Thermal Simulation with Hotspot Mitigation by Decoupling

Tao Chou, Hsin-Cheng Lin, Ching-Wang Yao, Yu-Sheng Lai, Fang-Yu Chang, Chia-Wei Tseng, Tzu-An Chang, He-Wen Shen, C. W. Liu
National Taiwan University

Physics-based thermal SPICE of NSs and CFET integrated with FEM enables a comprehensive thermal simulation from transistor level to package level. Self-heating in transistors contributes 2–9°C to the ΔT. The thermal impact of

a hotspot is decoupled and accounts for 27%–53% of max ΔT. The CFET has 27%-25% higher max ΔT than NSs. The sensitivity of ΔT w.r.t. layer thickness provides insights for thermal solutions. Four approaches for hotspot mitigations are investigated, including dummy via in BEOL, AlN bonding, thermal vias, and high-thermal-conductivity ILD for the BSPDN, which can reduce max ΔT by 5%, 5%, 7%, and 23%, respectively.

11:10 AM

17-5 Thermal Evaluation and Comparison of CAA and GAA Indium Tin Oxide Vertical Channel Transistors

Young Suh Song[1], Jimin Kang[2], Hyeonho Gu[3], Haotian Su[2], Yuan-Mau Lee[2], Shan X. Wang[2], Jimin Kwon[3], H.-S. Philip Wong[2], Eric Pop[2]
[1]Department of Electrical Engineering, Stanford University, [2]Stanford University, [3]Ulsan National Institute of Science and Technology

We carry out the first thermal simulations and comparison of channel-all-around (CAA) and gate-all-around (GAA) indium tin oxide (ITO) vertical channel transistors (VCTs), rooted in the experimental measurements. Despite potential electrical advantages and scalability, such CAA or GAA ITO device may face thermal challenges due to the low thermal conduc-tivity of ITO and their structure, with ITO channel deeply buried in a vertically-etched trench. Our quantitative thermal analysis reveals that CAA architecture can achieve ~48% reduction in the maximum device temperature rise (ΔTmax) compared to the GAA counterpart. Moreover, by adopting molybdenum (Mo) as source/drain material, ΔTmax further decreases, resulting in ~56% total reduction of both ΔTmax and thermal resistance (Rth). This thermal analysis of advanced device architectures lays the groundwork for future vertically-stacked technologies.

11:35 AM

17-6 Experiments and Modeling of Defect Dynamics and BTI in Doped In_2O_3 TFTs Undergoing Densification during 400 °C Post-BEOL Forming Gas Annealing (FGA)

Yu-Hsin Kuo[1], Chengyang Zhang[2], Priyankka Gundlapudi Ravikumar[2], Sanghyun Kang[2], Taeyoung Song[2], Marco A. Villena[3], Luca Larcher[4], Hwan Kim[5], Minji Hong[5], Pilsang Yun[5], Gaurav Thareja[4], Shimeng Yu[2], Daewon Ha[5], Suman Datta[2,2], Julia E. Medvedeva[6], Asif I. Khan[2,2]
[1]ECE, Georgia Tech, [2]Georgia Tech, [3]University of Granada, [4]Applied Materials, [5]Samsung Electronics Co. Ltd., [6]Missouri University of Science & Technology

We present the first systematic study of how mandatory 400 °C FGA affects oxide TFTs, combining electrical tests with DFT-based MD simulation. Our model device is a W-doped In_2O_3 channel on 5 nm HfO_2, capped by a hybrid stack that blocks external hydrogen. Key insights follow. (1) Early FGA (10 min) compacts the amorphous matrix, spawning shallow In–H–In donors (~0.65 eV below EF), which deepen PBTI (ΔVTH = –122 mV at 85 °C, 2 ks) and create an IDS–VGS “kink.” (2) Extended FGA (>40 min) partially crystallizes the film, locking hydrogen into deep In–OH/In–In complexes, suppressing mobility and recovering PBTI (ΔVTH = –10.4 mV) while erasing the kink. (3) After 60 min, subthreshold swing tightens by ~37 %, and both PBTI and NBTI improve with a polarity reversal. These results clarify hydrogen–defect dynamics during post-BEOL FGA and highlight the need for robust H-blocking encapsulation to ensure reliable oxide-TFT integration.

Session L: FETs Lunch

12:30 PM, Plaza

A Century of Miracles: From the FET’s Inception to the Horizons Ahead, Hiroshi Iwai, NYCU/Inst. Sci. Tokyo
From the invention of the FET exactly a century ago to today’s ultra-high-performance CMOS, the journey has been propelled by a highly improbable chain of coincidences—a true "mission impossible" that has led to a series of technological miracles. In this talk, we will revisit those extraordinary milestones and explore the next potential miracle that could once again redefine the very limits of performance.

Semiconductor Market Outlook: Speedbumps on the Road to a Terrabuck?, Bob Johnson, Gartner
As the semiconductor market enters a period of strong growth mode driven by demand from AI, the next few years will not be business as usual. External forces are gathering on the horizon which will create opportunities for those who can anticipate and manage a challenging environment.

Session 18: MT | NAND Technology

2:15 PM, Grand Ballroom A
Co-Chairs: Hongmei Wang, Micron and Yanli Zhang, Infineon
The NAND Technology session features eight papers highlighting the latest industry advancements in 3D NAND, FeNAND, and innovations in process and structural design. The session opens with a presentation on the integration and characterization of 321-layer 3D NAND, offering insights into SK hynix's most advanced technology to date. Following that, Samsung presents its work on Mo-based wordline metal integration for its 9th-generation V-NAND. SK hynix will then share progress on integrating an isolated charge trap layer into its 176-layer 3D NAND, supporting continued scaling. KAIST will present research on a BON tunneling layer aimed at enhancing the memory window. SK hynix will also introduce a novel split-cell integration approach to enable PLC (passive logic cell) functionality. The second half of the session focuses on FeNAND. Micron will present a low-voltage vertical FeNAND featuring a wide memory window and strong retention, addressing read disturb and interference through gate stack optimization. KAIST will follow with a FeNAND cell incorporating a 2 nm IGZO trap layer, achieving high endurance and retention via trap profile engineering. Finally, KAIST will showcase a scalable NC-NAND design using ultrathin Mo and sub-3.5 nm HZO, enabling a wide memory window and efficient triple-level cell operation.

2:20 PM
18-1 Integration and Characteristics of 321 Layer 3D-NAND Utilizing Vertical Stack-up and Lateral Scaling
Jongho Jung, Sungjae Chung, Haechan Park, Jihyun Seo, Yongsik An, Youngrok Kim, Eunmi Kim, Jinho Bin, Deayoung Moon, Dukeui Lee, Iksoo Choi, Jungryul Ahn, Eunjung Ko, Woopyo Jung, Sunyong Cha, Hyun Ahn, Sungyong Chung
SK hynix
This paper presents our most advanced 3D NAND technology. Compared to our previous 238-layer node, the new V9 generation demonstrates a 44% increase in net die per wafer along with 14%/15% improvements in read/write speeds, respectively. A key enabler of this advancement is the implementation of a 321-layer (321L) stack—currently the industry's highest—realized through a state-of-the-art triple-plug (3-Plug) integration scheme. In addition to vertical scaling, this work also addresses lateral scaling, achieving a 11% improvement in cell density through innovations in plug and slit integration.

2:45 PM
18-2 Molybdenum-integrated High-Speed 3D NAND: Breakthrough in Scalibility and Reliability for Mass Production

Seok Jung Yun, Hee Seung Wang, Hyun Jun Ahn, Keun Lee, Kwangmoon Kim, Jungsook Kim, Dongwook Lee, Giuk Jeong, Raeyoung Lee, Dong-sik Lee, Jonghyun Park, Hyunsoo Woo, Sejun Park, Byoungil Lee, Kyungyoon Noh, Seungwan Hong, Sung-Hoi Hur
Samsung Electronics Co.
Beyond extreme scaling in 3-dimensional NAND (3D NAND) flash memory has been limited in Tungsten (W) word line (WL) due to insufficient resistivity and reliability. We report the first mass production of low resistive and fluorine (F)-free Mo for high-speed 9th generation samsung 3D NAND. Industrial-scale high-volume manufacturing of Mo gate was successfully demonstrated by mass verifications of the cell characteristics and reliability, therefore, our innovative breakthrough opens the path toward further scalability for next-generation flash memory.

3:10 PM
18-3 A Highly Scalable Isolated Charge Trap Nitride Layer Implemented in a 176-layer 3D NAND Flash with Superior Threshold Voltage Distribution and Charge Retention
Sang-wan Jin, Doyoung Kim, Seonghee Kim, Eunmi Kwon, Hyungjun Yang, Keonsoo Shim, Jongung Baek, Jeongah Hong, Yousang Yoon, Seokmin Jeon, Seongman Jeon, Sangseob Lee, Joon Hyuk Park, Ju Yeab Lee, Sanghyun Oh, Jum Soo Kim, Youngmann Cho, Seonyong Cha
Sk hynix
In this work, we present the successful implementation of a novel charge trap nitride isolation (CTI) technology in a 176-layer production-scale 3D NAND device. In this approach, the charge trap nitride (CTN) layer is confined by a

pocket with sidewalls formed through a deposition process, a first-time implementation that improves CTI scalability. We achieved sufficient process maturity to produce fully functional chips across a substantial wafer area, enabling chip-level evaluation of threshold voltage (Vth) distribution and retention. This technology achieves an 11% reduction in read time (tR), a 6.9% narrower Vth distribution width, and a 45% improvement in high-temperature retention after cycling compared to conventional continuous CTN cells. These results suggest the current tier pitch limit could be reduced by 4 nm in terms of Vth distribution and by more than 10 nm from a retention perspective, supporting our CTI technology as a compelling solution for continued 3D NAND scaling.

3:35 PM

18-4 Bandgap-Engineered Boron Oxynitride Tunneling Layer for Reliable PLC operation of 3D V-NAND Flash Memory Devices

Dae Hyun Kang[1], Jaejoong Jeong[2], Youngkeun Park[2], Yungjun Kim[2], Dong Hun Sin[3], Han Mei Choi[3], Byung Jin Cho[2,2]
[1]School of Electrical Engineering, KAIST, [2]KAIST, [3]Samsung Electronics

We introduce boron oxynitride(BON) as a novel tunneling layer to replace SiON in the conventional SiO_2/SiON/SiO_2 stack. BON exhibits lower valence band offset and higher conduction band offset, improving hole injection during erase and reducing electron loss during retention. These characteristics of BON help overcome the tradeoff between erase efficiency and the reliability. Devices with BON show a 25.4% steeper ISPE slope, 9 times and 23 times faster erase before and after cycling, and improved retention and endurance. These advantages enable reliable PLC operation, achieving 3 times lower standard deviation and 4 times larger state-to-state margin compared to SiON-based devices.

4:25 PM

18-5 Mass Producible Multi-Site NAND Flash for Fast, Reliable and Power-Efficient 5-Bits/Cell Operation

Jung-Shik Jang, In Su Park, Hyunseung Yoo, Hokyun Jang, Younghun Yun, Seongman Jeon, Rhogyu Kwak, Nayeong Yang, Seokmin Choi, Wongeun Choi, Hyunwoo Kim, Keonsoo Shim, Hyunmin Lee, Changheon Park, Dongyean Oh, Juyeab Lee, Sang Hyun Oh, Jumsoo Kim, Sunghoon Lee, Woopyo Jeong, Seonyong Cha, Hyun Ahn
SK hynix

A new 5-bits NAND flash memory has been demonstrated with oval shaped cell and channel cut structure. By reducing the number of Vt states required to store the bits using physically divided cell space, much faster, more reliable and power-efficient characteristics are achieved even in the 5-bits per cell operation.

4:50 PM

18-6 Low VPGM Operation, Read Disturb, and Cell-to-Cell Interference in Scaled Vertical FeNAND

Pankaj Sharma, Albert Fayrushin, Davide Resnati, Karine Florent, Aurelio Mauri, Jin Yue, Jay Ahn, Emilio Camerlenghi, Akira Goda, Terry Kim, Richard Hill
Micron Technology

This study demonstrates GIL engineered vertical GAA 3D ferroelectric NAND device that achieves wide MW of ~12V at sub-15V programming voltage. This marks 4.4V reduction compared to previous results, while achieving record retention and endurance. We present first experimental evaluation of read disturb (RD) and cell-to-cell interference (C2C) in FeNAND at scaled WL pitch, with benchmarking against NAND. The observed increases in RD (3×) and C2C (7×) are attributed to design trade-offs made to achieve wider MW, including thinner channel interlayers and gate-side charge injection. These challenges can be mitigated through advanced gatestack engineering, optimized geometries and alternative biasing-schemes.

5:15 PM

18-7 A Highly Reliable Ferroelectric NAND Cell with Ultra-thin IGZO Charge Trap Layer: Trap Profile Engineering for Endurance and Retention Improvement

Hyunjun Kang[1], Hongrae Joh[2], Junhyeok Kwak[2], Giuk Kim[2], Hyojun Choi[2], Hoon Kim[2], Sanghyun Park[3], Kwangyou Seo[3], Kwangsoo Kim[3], Wanki Kim[3], Daewon Ha[3], Jinho Ahn[4], Sanghun Jeon[2]
[1]School of Electrical Engineering, Korea Advanced Institute of Science and Technology, [2]Korea Advanced Institute of Science and Technology, [3]Samsung Electronics Co., Ltd, [4]Hanyang University

We demonstrate a ferroelectric NAND(FeNAND) cell integrating a 2 nm-thick IGZO charge trap layer(CTL) within a metal-gate interlayer-oxide semiconductor-ferroelectric-channel interlayer-Si(MISFIS) gate stack to enhance reliability. The IGZO CTL suppresses oxygen vacancy formation in the ferroelectric layer and introduces an energy band offset to reduce charge loss. In-situ N_2 doping at 2 sccm tailors deep-level trap profiles, yielding a 9.4 V memory window at an operation voltage below 17 V. The device achieves stable triple-level-cell retention over 10 years and robust endurance beyond 80k program/erase cycles, validating MISFIS FeNAND as a promising architecture for 3D ferroelectric memories.

5:40 PM

18-8 Ultrathin-Mo-Enabled NC-NAND with Sub-3.5 nm HZO for Scalable High-MW Operation

Sangho Lee[1], Yongsu Kim[2], Yangjin Jung[2], Seungyeon Chang[2], Seokjoong Shin[2], Giuk Kim[2], Hongrae Joh[2], Woongin Kim[2], Sanghyun Park[3], Kwangyou Seo[3], Kwangsoo Kim[3], Wanki Kim[3], Daewon Ha[3], Jinho Ahn[4], Sanghun Jeon[2]
[1]School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), [2]Korea Advanced Institute of Science and Technology (KAIST), [3]Samsung Electronics, [4]Hanyang University

We propose a device-level integration strategy for NC-CTF memory enabling low programming voltage and high reliability. Sub-3.5 nm HZO achieves robust NC behavior via ultrathin Mo electrode engineering, minimizing hole CD scaling penalties. An oxide semiconductor ensures thermal stability for 3D integration, while a source-tied covering metal enhances ERS performance. TiO_2 band engineering at the BO/CTL interface suppresses electron injection, improving endurance. The resulting NC-NAND demonstrates a wide memory window (~13 V) and stable TLC operation, highlighting its scalability and energy efficiency for next-generation 3D flash memory.

Session 19: ALT | Transistors with 2D, Oxide Semiconductor, Novel Materials

2:15 PM, Grand Ballroom B

Co-Chairs: Gabriela Dilliway, SemiTech Insights and Koji Watanabe, Tokyo Electron Limited

The integration of two-dimensional (2D) materials and semiconductor oxides into CMOS technology is reshaping the landscape of modern electronics. With their atomically thin structures and tunable electronic properties, 2D materials offer promising avenues for scaling beyond silicon, enabling high-performance logic with reduced power consumption and enhanced device flexibility. In parallel, semiconductor oxides—valued for their wide bandgaps, chemical robustness, and thermal stability—are gaining traction as viable channel materials in advanced CMOS architectures. This session presents eight cutting-edge papers from both academia and industry, highlighting recent progress in the use of 2D and oxide semiconductors for novel CMOS channel designs. We begin with a comparative thermal analysis of 2D semiconductors versus silicon in nanosheet transistors, followed by a demonstration of 2D channel-last integration in stacked-channel FETs. Subsequent talks introduce innovative approaches to side contacts in monolayer MoS_2 FETs and a vertical pFET featuring symmetrical source and drain configurations. The session then shifts focus to three papers exploring InOx-based channel technologies, showcasing their potential for low-leakage, high-mobility CMOS applications. Finally, we conclude with a presentation on a tellurium-based Schottky barrier FET (SBFET), offering a glimpse into emerging material systems for future logic devices.

2:20 PM

19-1 First Comparative Thermal Evaluation of 2D Semiconductor vs. Silicon Nanosheet Transistors

Young Suh Song[1], Haotian Su[2], Tara Pena[2], Anton E. O. Persson[2], Kyungwhan Yang[3], Eilam Yalon[4], Robert K.A. Bennett[2], Zherui Han[2], Kathryn Neilson[2], Jimin Kang[2], Jerry A. Yang[2], H.-S. Philip Wong[2], Shan X. Wang[2], Eric Pop[2]
[1]Department of Electrical Engineering, Stanford University, [2]Stanford University, [3]Samsung Electronics, [4]Israel Institute of Technology

We present the first comparative thermal analysis of gate-all-around (GAA) nanosheet (NS) transistors with 2D semiconductors vs. Si channels. Despite potential electrical advantages, replacing Si- with 2D-NS may introduce thermal challenges due to the low thermal boundary conductance (TBC) between the 2D material and its environment. Building on our thermal measurements of monolayer MoS2 and ultrathin HfO2, we use simulations to compare the thermal behavior of 2D- vs. Si-NS transistors at 6–12 nm gate lengths. In this regime, the thermal role of the contacts dominates, and 2D-NS transistors are at a disadvantage due to their small contact area. We find 2D wrap-around contacts (WAC) are thermally superior to edge contacts (EC), and an even better choice are

metallized multilayer 2D contacts. These bring the thermal resistance of 2D-NS within 50% that of Si-NS transistors.

2:45 PM

19-2 Novel Channel-last Integration of ALD MoS_2 Into Stacked Channel FETs on 300mm Wafers

Sylvain Barraud[1], Michael Rodriguez-Fano[2], Jean-Michel Pedini[2], Stephane Cadot[2], Rihab Chouk[2], Benjamin Dey[2], Jean-Michel Hartmann[2], Ahmed Gharbi[2], Corinne Comboroure[2], Aurelien Sarrazin[2], François Boulard[2], Lucie Laraignou[2], Alain Campo[2], Helen Grampeix[2], Clement Castan[2], Julian Sturm[2], Aurelie Souhaite[2], Andrea Lassenberger[2], Ludovic Couture[2], Denis Mariolle[2], Pauline Hauchecorne[2], Virginie Loup[2], Erwan Gapihan[2], Kevin P. O'BRIEN[3], Uygar E. AVCI[3], François Andrieu[2]
[1]Minatec Campus, CEA-Leti, [2]CEA-Leti, [3]Intel Corporation

We report, for the first time, the integration of ALD MoS_2 into stacked channel transistors via a novel channel-last integration scheme. Our approach provides many advantages concerning channel formation, enhanced delamination control, contact optimization, reduced contamination risks, and the ability to reuse process modules of existing Si GAA routes. First, we developed and optimized ALD MoS_2 films. We optimized S/D contacts using back-gated MoS_2 transistors. Niobium contacts delivered an I_{ON}/I_{OFF} ratio up to 10^4 and outperformed titanium counterparts. Then, we validated key modules of our channel-last integration, overcoming most of technological roadblocks limiting the 2D TMD integration into advanced CMOS.

3:10 PM

19-3 Investigation into Side Contact Behavior and Transfer Length of Monolayer MoS_2 FETs

Daniel Walter Eppler[1], Hao-Yu Lan[2], Zhihong Chen[2], Joerg Appenzeller[2]
[1]West Lafayette, School of Electrical and Computer Engineering, Purdue University, [2]Purdue University

For the first time, FETs with contacts only to the edge of monolayer MoS_2 channels – in the following called “side contacts” – are demonstrated, exhibiting surprisingly low source/drain contact resistances for Ni contacts if compared to top contacted devices. Channel lengths have been scaled down to 20nm, utilizing gate stacks with sub-1nm EOT. Side contacted transistors are compared to top contacted devices of the same dimensions. These side contacted FETs exhibit up to 200 µA/µm current despite having very small contact areas of only~160nm^2. The large current density observed in case of side contacted devices is attributed to a 150x improved side contact interfacial resistance density of 5.3 Ω*µm^2. compared to conventional top contacted transistors. Indirect evidence of sub 2nm transfer length (L_{Tt}) in top contacted transistors, far smaller than previously reported values, is provided.

3:35 PM

19-4 First Demonstration of 2D Semiconductor Vertical pFET with Symmetrical Source/Drain Transport of I_{on}>100 µA/µm with Extremely High thermal Stability up to 550 ℃ by Scalable Direct Channel Growth

Xin Wang, Xiong Xiong, Yuhang Chen, Tianyue Fu, Zhiyu Wang, Yanqing Wu
Peking University

Vertical field-effect transistors (VFET) offer unique structural advantages for 3D integrated circuits. However, scalable implementations of two-dimensional (2D) semiconductor vertical transistors remain largely unexplored. In this work, we present the first demonstration of a self-aligned 2D semiconductor vertical pFET with a 20 nm physical gate length, enabled by direct source/drain seed-layer growth of the 2D channel for scalable integration. The device features a 2 nm-thick (tri-layer) p-type WSe_2 channel with excellent source/drain symmetry transport characteristics, exhibiting an on/off current ratio >10^7 and an on-state current >100 µA/µm. The WSe_2 VFET also demonstrates thermal process tolerance at 550 °C and stable operation at 100 °C, attributed to the intrinsic single-crystalline quality of the WSe_2 channel and the high melting point of the tungsten (W) source/drain contacts. This work highlights the great potential of 2D semiconductors for thermal stability vertical transistors for future CFET logic and 4F^2 DRAM applications.

4:25 PM

19-5 First Demonstration of Fluorine Supercritical Fluid Passivated 2-floor GAA In_2O_3 Nanosheets Achieving Positive 0.7 V V_T, 67 mV/dec SS, 1309 μA/μm $I_{D,max}$, Record 47 mV PBS, and Negligible NBS

Yu-Shan Wu[1], Rong-Wei Ma[2], Yuan-Ming Liu[2], Johannes Gracia[2], Chuan-Wei Kuo[3], Hsien-Ming Sung[2], Hidenari Fujiwara[2], Ting-Chang Chang[3], C. W. Liu[2]

[1]Graduate Institute of Electronics Engineering, National Taiwan University, [2]National Taiwan University, [3]National Sun Yat-Sen University

2-floor gate-all-around (GAA) In_2O_3 nanosheets is passivated by fluorine supercritical fluid (F-SCF) at 200°C. The device achieves a subthreshold swing (SS) of 67 mV/dec, a threshold voltage (V_T) of 0.7 V, and a high $I_{D,max}$ of 1309 μA/μm at V_{DS}= 3 V/ Vov=2.3 V. F-SCF enables conformal and non-destructive defect passivation, improving electrical characteristics and reliability. Positive bias stress (PBS) reliability is improved, with a reduced ΔV_T of 47 mV at V_{OV}=3 V after 8000 s. With minimal hydrogen content in gate dielectrics, no ΔV_T is observed after negative bias stress (NBS). Structural and chemical analyses confirm F incorporation and intact channel morphology. The compatibility of F-SCF with 3D stacked oxide semiconductor (OS) devices has great potential for BEOL transistor applications.

4:50 PM

19-6 Overcoming Threshold Voltage -Performance-Stability Tradeoff with Double-Gate Gallium Doped In_2O_3 MOSFETs using Tri-Layer HfO_2-ZrO_2-HfO_2 Gate Stack

Eknath Sarkar[1], Hyeonwoo Park[2], Md Abdullah Al Mamun[3], Hyun Jae Lee[2], Dyutimoy Chakraborty[2], Emmanuel Quezada[2], Chengyang Zhang[2], Sharadindugopal Kirtania[2], Mengkun Tian[2], Asif I Khan[2], Shimeng Yu[2], Kyeongjae Cho[3], Hwan Kim[4], Changik Im[4], M.J Hong[4], Daewon Ha[4], Suman Datta[2]

[1]School of Electrical and Computer Engineering, Georgia Institute of Technology, [2]Georgia Institute of Technology, [3]University of Texas at Dallas, [4]Samsung Electronics Co., Ltd

We demonstrate double-gate (DG) amorphous oxide semiconductor (AOS) MOSFETs by employing a laminated HfO_2/ZrO_2/HfO_2 (HZH) gate stack, achieving a holistic enhancement in threshold voltage (V_T) modulation, drive current, and reliability. Key accomplishments include attaining a ~400 mV positive V_T shift, significantly lower equivalent oxide thickness and near-ideal subthreshold swing, leading to superior drive current. Enhanced mobility and reduced trap densities, confirmed through tail state analysis, contribute to improved transistor performance, while BTI modeling demonstrates effective degradation suppression. Notably, the device achieves an I_{ON} of 470 μA/μm at V_{OV}=V_{DS}=1V, with a field-effect mobility of 75 cm^2/V·s, and maintains an exceptionally low PBTI-induced ΔV_T shift of only 10 mV at 2 MV/cm E-field stress and 95 °C. This HZH gate stack offers a unique balance between high drive current, mobility, and stability, outperforming all reported AOS transistors and establishing a promising platform for future high-performance, reliable BEOL AOS transistors

5:15 PM

19-7 Enhancement-mode BEOL In_2O_3 FETs with Record Logic Performance: Experiments and Compact Modeling

Yanjie Shao[1], Dylan Ma[2], Dimitri A. Antoniadis[3], Lan Wei[2], Jesús A. del Alamo[3]

[1]Microsystems Technology Laboratories, Massachusetts Institute of Technology, [2]University of Waterloo, [3]Massachusetts Institute of Technology

We demonstrate record logic performance in BEOL-compatible enhancement-mode amorphous oxide semiconductor (AOS) field-effect transistors (FETs). These devices exhibit near-ideal scalability down to 40 nm in channel length (L_{ch}). Using an In_2O_3 channel by plasma-enhanced atomic-layer deposition, we achieve enhancement-mode operation in L_{ch} = 40 nm devices with a maximum drive current of 1.35 mA/μm, a peak transconductance of 490 μS/μm, and a close-to-thermal-limit average room-temperature subthreshold swing of 63 mV/dec, all at V_{ds} = 0.5 V. A contact resistance (R_c) of 84.5 Ω.μm, record-low among E-mode AOS-FETs, is demonstrated. Capacitance-voltage characteristics reveal different threshold voltages in the intrinsic channel and in the source/drain to gate overlap regions. We develop a physics-based MVS-AOS model which accurately captures the essential physics at play in our experimental devices, including capacitance with frequency dispersion in the contact region. This work advances state-of-the-art BEOL AOS-FET technology for BEOL monolithic 3D integration.

5:40 PM

19-8 Reinterpreting p-type Te FETs as SBFETs Enables Reduction of Hysteresis and Contact Resistance

Minjae Kim[1], Useok Choi[2], Junho Ban[2], Uigwan Shin[2], Hae-Won Lee[2], Gyeongsu Lee[2], Byoung Hun Lee[2,2]
[1]Department of Electrical Engineering, Pohang University of Science and Technology (POSTECH), [2]Pohang University of Science and Technology (POSTECH)

Te is a promising p-type semiconductor, but its narrow bandgap (~0.7 eV), high off-state current, electrical hysteresis, DIBL-like behavior, and difficulty in achieving enhancement-mode operation hinder practical use. We argue that these issues arise from misinterpreting Te Schottky barrier (SB) FETs as conventional MOSFETs and relying on room-temperature measurements. We propose three key strategies: (1) suppressing oxygen-related defects from channel, (2) engineering Schottky barriers at metal/Te interface to reduce contact resistance (R_C), and (3) performing low-temperature characterization. Using these approaches, we demonstrate enhancement-mode Te FETs with negligible hysteresis and DIBL, and low R_C (1.5 kΩ·μm) and high mobility (54 cm^2/V·s).

Session 20: NC | Probabilistic In-Memory Computing

2:15 PM, Continental 1-3

Co-Chairs: Tuo-Hung (Alex) Hou, National Yang Ming Chiao Tung University and Daniele Ielmini, Politecnico di Milano

This session presents six papers about probabilistic in-memory computing leveraging device uncertainty and randomness through probabilistic models and algorithms to process data and make decisions, excelling in applications such as data analysis, decision-making under uncertainty, and data security. The first paper by Y. Hou, et al., from HKUST, presents a fully analog VCMA-MTJ Ising machine integrating weight and p-bit functonalities without DAC, ADC, or TRNG. The second paper by Y. Su, et al., from Tsinghua University, presents IGZO-based 2T gain-cell PUF for secure in-memory computing. The third paper by V. B. Adishesha et al., from NUS, presents a 28nm STT-MRAM-based Ising Machine for solving large scale 8000-node p-bit problems. The fourth paper by G. Zhang, et al., from Zhejiang University, introduces a self-selective memristor array for light-weight adversarial steganography. The fifth paper by K. Selcuk, et al., from UCSB, presents a DAC-free, asynchronous stochastic MTJ p-bit design with on-chip annealing for enhanced scalability and efficiency. The sixth paper by C. Xia, et al., from Peking University, presents an IGZO TFT-based optical and electrical p-bit with capabilities for Deep Boltzmann Machines in-sensor image recognition, reconstruction and generation.

2:20 PM

20-1 Fully Analog In-Memory Annealing Ising Machine with Unified Deterministic and Probabilistic Computing Array Using CMOS-integrated VCMA-MTJ

Yaoru Hou[1], Zhihua Xiao[1], Xuezhao Wu[1], Zihan Tong[1], Shuyu Wang[2], Mingtao Chen[2], Chang Ho Chan[1], Yunyang Lu[1,1], Yiyang Ho Zhang[1], Chin-Chung Chen[3], Zheng Zhu[3], Albert Lee[3], Di Wu[3], Hao Cai[2], Qiming Shao[1]
[1]Hong Kong University of Science and Technology, [2]Southeast University, [3]InstonTech

This work presents a fully analog in-memory annealing Ising machine based on CMOS-integrated VCMA-MTJ, developed through a device-circuit-system co-design approach to address bottlenecks in traditional digital CMOS and mixed-signal CMOS+X Ising machines. By unifying classical bit and probabilistic bit functionalities into the VCMA-MTJ device, the proposed chip, fabricated using 28-nm CMOS+VCMA-MTJ technology, enables simultaneous deterministic calculations and probabilistic spin updates, thus achieving a fully analog in-memory Ising machine. This work demonstrates spin density enhancement of 8.9× to 73.7×, while achieving annealing power per spin of 28μW (power reduction of 2.46× to 147×) compared to prior CMOS and CMOS+X-based Ising machines.

2:45 PM

20-2 Embracing OS Gain Cell into Hardware Security: M3D Integrated 1Kb IGZO 2T0C Gain Cell PUF Array In-Situ Encrypting RRAM CiM for Secure Edge AI

Yanbo Su[1], Jianshi Tang[1], Zhidong Tang[1], Yijia Fan[1], Xueqi Li[1], Zhixing Jiang[1], Junyang Zhang[1], Yibei Zhang[1], Ruofei Hu[1], Kexin Wang[1], Yide Li[1], Gangping Yan[2], Guilei Wang[2], Chao Zhao[2], Peng Yao[1], Bin Gao[1], He Qian[1], Huaqiang Wu[1]
[1]Tsinghua University, [2]SAMT

We present the first monolithically 3D-integrated PUF-CiM architecture for secure edge AI. It vertically stacks an IGZO-based 2T GC PUF array (for cryptographic keys), an RRAM CiM core, and Si CMOS logic. The fabricated 1Kb IGZO PUF exploits inherent randomness, demonstrating high entropy. Stacking the PUF atop the CiM core enables in-situ encryption via XOR-based weight programming while preserving computational accuracy. System evaluation confirms only authorized keys decrypt and use the network. This M3D-PUF chip achieves 39× lower energy and 15.8× higher speed versus 2D designs, providing hardware-intrinsic security through unique key generation.

3:10 PM

20-3 PRIME: Probabilistic Reconfigurable Ising Memory Engine an 8000 p-bit Ising Machine with 28 nm Multifunction STT-MRAM

Vivek Bettadapura Adishesha[1], Taehwan Kim[2], Jongsun Park[2], Hyunsoo Yang[3]
[1]Electrical and Computer Engineering, National University of Singapore, [2]Korea University, [3]National University of Singapore

Probabilistic computing is rising in significance for solving complex combinatorial problems. Scaling Ising machines demands dense p-bit arrays. In this work, we present PRIME (Probabilistic Reconfigurable Ising Memory Engine), one of the first multifunctional MRAM architectures based on perpendicular STT-MTJ using standard 1T1R bitcells. PRIME unifies four modes: true random number generation, non-volatile memory, probabilistic bits, and stochastic neurons, through digital pulse-width modulation. Over 8,000 MTJ-based p-bits are implemented on-chip, demonstrating the feasibility of a superscalar Ising processor and solving an 8,000-node optimization problem. PRIME's pulse-controlled design enables a low-power, reconfigurable, and scalable platform for probabilistic computing and AI.

4:00 PM

20-4 First Demonstration of Adversarial Steganography Using Self-Selective Array for Efficient and Secure Cloud-Edge Computing

Guobin Zhang[1,2,3], Xuemeng Fan[1,3], Zijian Wang[1,3], Pengtao Li[1,2,3], Yiqi Zhou[4], Daying Sun[4], Kun Ren[1,3], Yunlong Li[1,3], Bin Yu[1], Qing Wan[2], Dawei Gao[1,3], Yishu Zhang[1,2,3]
[1]Zhejiang University, [2]Yongjiang Laboratory, [3]Zhejiang ICsprout Semiconductor Co., Ltd, [4]Nanjing University of Science and Technology

We demonstrate dynamic adversarial steganography using a 256×256 self-selective memristor array for efficient, lightweight data security. The SSM achieves record metrics: selective ratio > 10^6, nonlinearity > 10^6, on/off ratio > 10^3, and 512 conductance states (9-bit) with scalability exceeding 1.568 Tb. Integrating an adversarial optimization strategy and conductance self-correction scheme, the model enables high-precision embedding with 36.1 dB peak signal-to-noise ratio for images and 28.3 dB signal-to-noise for audio. Hardware benchmarks show 59.44× lower energy, 15.60× faster computation, and 2.30× greater area efficiency versus graphics processing unit implementation. This work enables scalable, undetectable data hiding with cloud compatibility.

4:25 PM

20-5 DAC-Free p-bits: Asynchronous Self-Coloring and On-Chip Annealing

Kemal Selcuk[1], Navid Anjum Aadit[2], Corentin Delacour[2], Jared Quintana Silva[2], Nihal Sanjay Singh[2], Haruna Kaneko[3,4], Shun Kanai[3,4], Yu-Jui Wu[5], Yi-Hsuan Chen[5], Yu-Sheng Chen[5], Yi Ching Ong[5], Kuo-Ching Huang[5], Harry Chuang[5], Hideo Ohno[3], Shunsuke Fukami[3,4], Kerem Yunus Camsari[2]
[1]Department of Electrical and Computer Engineering, University of California, Santa Barbara, [2]University of California, Santa Barbara, [3]Research Institute of Electrical Communication, Tohoku University, [4]Tohoku University, [5]Taiwan Semiconductor Manufacturing Company

Nanodevice implementations of probabilistic bits (p-bits) traditionally rely on analog-input methods, hindering scalability due to the necessity of digital-to-analog converters or analog crossbar connectivity. We introduce a DAC-free, digital-input digital-output p-bit architecture leveraging stochastic logic and intrinsic entropy from nanodevices, demonstrated here with stochastic magnetic tunnel junctions (sMTJs). Our approach achieves sigmoidal tuning curves without analog overhead and addresses variability using calibrated delay lines and XOR whitening. This design enables asynchronous operations, naturally resolving conflicts without explicit graph coloring, supporting scalable parallelism. Critically, our tunability circuit facilitates on-chip annealing by simply

adjusting delay lines, eliminating the need to recalculate or shuttle new weights during annealing, as experimentally demonstrated with interacting five sMTJs with FPGA. Benchmarking reveals significant area and power reductions compared to existing DAC-free designs. Our architecture represents a significant advancement toward scalable probabilistic computing in integrated CMOS and nanodevice systems.

4:50 PM

20-6 Achieving in-situ Optical Perception Probabilistic Computing with Optoelectronic IGZO TFTs

Chenhao Xia[1], Haoxuan Wang[2], Weican Chen[2], Guanwen Yao[2], Yuze Lu[2], Hongjie Zeng[2], Lei Sun[2,3], Dedong Han[2,3], Xiaoyan Liu[2,3], Yunyi Fu[2,3], Fei Liu[2,3]

[1]School of Integrated Circuits, Peking University, [2]Peking University, [3]Beijing Advanced Innovation Center for Integrated Circuits

We experimentally demonstrate, for the first time, a 4096-bit hardware probabilistic computing (p-computing) system capable of in-situ optical perception, based on integrated optoelectronic InGaZnO thin-film transistors (OE-IGZO TFTs). The proposed OE-IGZO TFT exhibits superior optoelectronic characteristics compared to other devices. A self-calibrating architecture stabilizes probabilistic bit (p-bit) probability distributions, enabling stable operation >1000 s. The p-bit units in optical mode exhibits characteristics of hardware-level activation function. Accordingly, the optical probability distribution sampling augmentation (OPDSA) technique is proposed. In situ, p-bit units in electrical mode act as hardware neurons within the deep Boltzmann machine. Under optical and electrical modulation, respectively, the output distributions of the hardware circuit achieve sigmoidal fitting rates > 99%. The proposed p-bit network demonstrates high-accuracy image recognition, reconstruction and generation, highlighting its potential for advanced in-hardware visual intelligence applications.

Session 21: ODI | Displays and Integrated Photonics

2:15 PM, Continental 4

Co-Chairs: Michael Cich, Google and Haiding Sun, University of Science and Technology of China

This session includes seven papers on topics related to integrating photonic devices, especially using dissimilar materials, to achieve new functionality with high performance. It starts with a pair of papers demonstrating compute-in-memory and an ultra-fast on-chip memory by the National Univ. of Singapore and the Univ. of Wisconsin-Madison. This is followed by a paper showing a monolithically-integrated emitter and detector acting as a quantum entropy source from KAUST. Another set of papers, by Univ. of Singapore and Univ. of Tokyo, demonstrate integration of dissimilar materials (lithium niobate and InP/InGaAs) on CMOS circuitry. A paper by Harvard presents improvements in power consumption for energy-efficient digital-to-optical converters in lithium niobate. Finally, an FeMTFT has been used to drive micro-led displays, by Southeast University.

2:20 PM

21-1 First Demonstration of Monolithic 3D Integrated Si PICs and BEOL Oxide Semiconductor EICs for Electronic-Photonic Compute-in-Memory (CiM)

Rui Shao[1], Yue Chen[2], Yuxuan Wang[2], Xuanqi Chen[2], Xudong Jia[2], Jiawei Sun[2], Chen Sun[2], Wei Wang[3], Xiao Gong[2,4]

[1]Department of Electrical and Computer Engineering (ECE), National University of Singapore, [2]National University of Singapore, [3]Shanghai Industrial μTechnology Research Institute, [4]Agency for Science, Technology and Research (A*STAR)

This work demonstrates a monolithic 3D integrated electronic-photonic compute-in-memory architecture, enabled by foundry-compatible oxide semiconductor electronic and photonic integrated circuits. The proposed two-transistor-and-one-diode (2T1D) EPCiM was first fabricated in a SiPh foundry and subsequently processed in university cleanroom, showcasing a practical and scalable integration pathway. This architecture features a simultaneous yet decoupled electrical-write and optical-read scheme, experimentally achieving key highlights: ultra-low leakage in OS EICs, fast electrical write time of 9.6 ns, a high-speed CiM throughput of 2.4 GS/s with an intrinsic optical read latency < 500 fs, and distinct 3-bit write states, enabling a high computing density of 8.46 TOPS/mm2. A low CiM error σ of 0.075 was experimentally demonstrated at a 2.4 GS/s rate with a projected 4.46 TOPS/W power efficiency. We further propose and evaluate a 2T1D EPCiM array as a hardware tensor core in a pretrained VGG-19 CNN, achieving 90.83% accuracy on the CIFAR-10 classification task.

2:45 PM

21-2 Demonstration of a Cross-Coupled Differential Regenerative Photonic Latch for Integrated Ultra-fast On-Chip Memory

Md Abdullah-Al Kaiser[1], Sugeet Sunder[2], Akhilesh Jaiswal[3], Ajey Jacob[2]
[1]Electrical and Computer Engineering, University of Wisconsin-Madison, [2]USC Information Sciences Institute, [3]University of Wisconsin-Madison

We experimentally demonstrate, for the first time, a cross-coupled differential regenerative photonic latch (pLatch) fabricated on GlobalFoundries (GF) Fotonix Silicon Photonics platform. We showcase robust read and write operations using differential photonic signaling, paving the way to scalable photonic SRAM - critical for advancing photonic in-memory computing, AI accelerators, tensor cores and active routing on photonic interposers for data center chips. Additionally, we report 1-bit photonic in-memory computing scalable to multi-bit operations as well as multi-bit first-in-first-out (FIFO) structure. The photonic latch reliably holds data for several minutes while supporting concurrent photonic read/write and compute operations. Simulations performed on the GF Fotonix platform validate ultra-fast operations up to 20 GHz, with resilience to 0.8 V single-ended electrical noises at 1.8 V VDD and ±25 pm tolerance to input laser wavelength drift. This work represents the first demonstration of a foundational building block for integrated photonic memory and memory-driven compute systems.

3:10 PM

21-3 On-Chip Tamper-Resistant Quantum Entropy Source Enabled by Monolithic Emitter–Detector Integration

Hang Lu[1], Omar Alkhazragi[2], Heming Lin[2], Tien Khee Ng[2], Boon Ooi[2]
[1]Electrical and Computer Engineering (ECE), King Abdullah University of Science and Technology, [2]King Abdullah University of Science and Technology

We demonstrate the first monolithic integration of a photodetector (PD) directly onto a vertical-cavity surface-emitting laser (VCSEL) on a single chip to form a compact optical entropy source for quantum random number generation (QRNG) aimed at embedded hardware security applications. In contrast to conventional hardware RNG configurations that rely on discrete components, our system exploits intensity fluctuations from a VCSEL captured in situ by the co-integrated PD without any external optics, amplification or careful alignment. This monolithic approach minimizes system operation, reduces footprint, and enables scalable, no-leakage, parallel, and alignment-free RNG, making it ideally suited for secure edge devices. The extracted random bitstream achieves an RNG rate of 10 Gb/s from every single channel. The random numbers successfully pass the full NIST statistical randomness test suite. This work establishes a new class of on-chip photonic entropy sources and lays the foundation for embedded secure hardware-based cryptographic systems.

4:00 PM

21-4 Heterogeneous Integration of Performant Lithium Niobate-On-Silicon Micro-Ring Modulator by High-Precision Micro-Transfer Printing

Jie Yang[1], Ao Yan[2], Jin Feng Leong[2], Chenxuan Zeng[2], Jieming Pan[2], Shi Zhao[2], Gan Zhang[2], Zefeng Xu[2], Baochang Xu[2], Tiancheng Yu[2], Xu Cheng[2], Sonu Hooda[2], Evgeny Zamburg[2], Benjamin Chee-Keong Tee[2], Aaron Voon-Yew Thean[2]
[1]Department of Electrical and Computer Engineering, National University of Singapore, [2]National University of Singapore

We demonstrate the first lithium niobate (LN) Micro-Ring Modulator (MRM) heterogeneously integrated on a silicon photonic platform via a high-precision, back-end-of-line (BEOL) compatible micro-transfer printing (MTP) method. Fully patterned, low-loss LN devices are aligned and transferred onto CMOS-compatible Si photonic chips with high yield and sub-150 nm accuracy. The resulting hybrid MRM achieves ultra-low insertion loss (<0.6 dB), and a low half-wave voltage–length product ($V_\pi L$ = 1.5 V·cm), enabled by optimized waveguide and electrode design. The approach supports both in-plane and vertical stacking on Si waveguides, offering flexible integration pathways. Verified through array-level transfer on foundry-fabricated chips with finished redistribution layers (RDL) interconnects, this scalable, non-invasive method overcomes key limitations of prior LN-on-Si techniques and enables monolithic integration of high-speed photonic I/Os for next-generation optical interconnects.

4:25 PM

21-5 Ultralow-power, High-speed Integrated Si Photonic Switch Enabled by III–V/Si Hybrid MOS Optical Phase Shifters

Tomohiro Akazawa[1], Rui Tang[2], Hanzhi Tang[2], Makoto Okano[3], Nobuyuki Matsuda[4], Kasidit Toprasertpong[2], Shinichi Takagi[2], Mitsuru Takenaka[2]
[1]Department of Electrical Engineering and Information Systems, The University of Tokyo, [2]The University of Tokyo, [3]National Institute of Advanced Industrial Science and Technology, [4]Tohoku University

Energy-efficient and high-speed Si photonic switches are key enablers for handling ever-increasing data traffic in data centers. However, conventional Si photonic switches have failed to simultaneously achieve low-power and high-speed switching. In this work, we demonstrate an integrated Si photonic switch based on III–V/Si hybrid MOS optical phase shifters, with ultralow static power consumption on the order of several hundred femtowatts for entire switch and nanosecond-order fast switching while achieving high modulation efficiency of 0.13 Vcm. Proposed Si photonic switch paves the way for energy-efficient, high-speed optical switching architectures suitable for AI data centers.

4:50 PM

21-6 Simultaneously Improving Power Consumption and Linearity of Digital-to-Optical Converters (DOCs) for Energy-Efficient Optical Data Transmission

John Davis[1], Georgios Kyriazidis[1], Jalil Morris[1], David Kong[1], Dylan Meyer-O'Connor[1], Melinda Modisette[1], Shahabeddin Mohin[2], Gage Hills[1]
[1]Harvard University, [2]Massachusetts Institute of Technology

Digital-to-Optical Converters (DOCs) convert digital electrical signals directly to analog optical signals, eliminating the need for Digital-to-Analog Converters in Electro-Optic modulators, for more energy-efficient communication. DOCs suffer from either: High non-linearity in transmitted optical power vs. input digital code (quantified by Integral Non- Linearity & Differential Non-Linearity). High INL & DNL degrade trade-offs in power, bandwidth, and Bit Error Rate. High signal count: previous “Engineered Segment Length” (ESL) DOCs improve INL & DNL, but require thermometer-coded digital inputs: N −1 inputs for N -level Pulse-Amplitude Modulation (PAM-N). N − 1 inputs requires N − 1 electronic drivers, increasing power. Here, we present ESL-star, a design technique enabling DOCs to achieve near-zero INL & DNL with only log2(N) inputs. Compared to thermometer-coded ESL-DOCs, ESL*-DOCs reduce power of electronic drivers by 3.82×; We fabricate ESL*-DOCs in Lithium Niobate, implementing PAM-16. Compared to non-ESL*-DOCs with log2(N) inputs, we measure 3.7× better INL & 2.1× better DNL.

5:15 PM

21-7 Flexible 1T-1FeMTFT Active-matrix Micro-LED Display Utilizing ITO FeMTFT with Record Memory Window of 0.63 V/nm and Low Power Consumption

Tingrui Huang[1], Guangan Yang[2], Yimeng Sang[3], Mingming Liu[4], Gengyu Li[4], Heng Cheng[4], Zuoxu Yu[4], Zhiyuan Fu[4], Siyang Liu[4], Haoliang Shen[5], Wangran Wu[4], Runxiao Shi[4], Zhe Zhuang[3], Weifeng Sun[4]
[1]School of Integrated Circuits, Southeast University, [2]Nanjing University of Posts and Telecommunications, [3]Nanjing University, [4]Southeast University, [5]Suzhou Laboratory

This work presents a flexible active-matrix μ-LED display using ITO-based ferroelectric metal thin-film transistors (FeMTFTs) on polyimide. The FeMTFT integrates a 12-nm $Hf_{0.5}Zr_{0.5}O$ layer crystallized at 400°C, achieving a record memory window of 0.63 V/nm (7.5 V), field-effect mobility >30 $cm^2V^{-1}s^{-1}$, on/off ratio of 4×10^8, and endurance $>5\times10^7$ cycles. Excellent flexibility is demonstrated through 10^5 bending cycles at a 4-mm bending radius. A novel 1T-1FeMTFT pixel circuit by leveraging ferroelectric non-volatility, enabling ultralow dynamic power consumption, dual driving schemes (PAM/PWM), and a 200 kHz pixel refresh rate. The flexible display achieves a high resolution of 447-PPI and low power consumption.

Session 22: EDT | Advances in Oxide Semiconductor Devices: Physics, Integration, and Reliability

2:15 PM, Continental 5

Co-Chairs: Asif Khan, Georgia Tech and Soo Gil Kim, SK Hynix

Oxide semiconductors are rapidly emerging as key enablers for future logic and memory technologies, offering unique opportunities in performance scaling, low-temperature processing, and integration with advanced 3D architectures. This session brings together breakthroughs spanning device physics, materials engineering, and reliability, highlighting how new design strategies and material innovations are driving oxide devices closer to manufacturable solutions. The session opens with a paper showing that quantum-well carrier confinement in oxide transistors can overcome the traditional mobility–threshold voltage trade-off, achieving high mobility with minimal VT penalty and strong bias stress reliability. The next paper reports the demonstration of positive-threshold dual-gated ITO FETs enabled by ZrO_2 dielectrics, addressing a long-standing challenge in oxide electronics and enabling stable enhancement-mode operation for advanced 3D integration. The third paper highlights amorphous Ge-doped In_2O_3 grown by ALD, which achieves an improved balance of mobility, stability, and crystallization resistance for integration across front- and back-end processes. Building on device integration, the fourth paper demonstrates the BEOL-compatible p-type nanosheet FETs, fabricated under low thermal budget conditions and exhibiting strong electrostatics and high drive current. In the fifth paper, contact length scaling is shown to suppress contact-induced doping, effectively mitigating threshold roll-off and bias stress instability in short-channel oxide devices. Reliability aspects are then explored in the sixth contribution, which examines whether high-temperature PBTI testing can serve as a predictor of long-term oxide FET lifetime, revealing key limitations of accelerated testing methods. Finally, the session concludes investigating statistical variability in crystalline nanosheet oxide FETs, systematically comparing them with amorphous counterparts and highlighting the impact of grain size on device performance.

2:20 PM

22-1 Quantum-Well Carrier Confinement Overcomes Mobility–V_T Trade-off of Oxide Semiconductor Transistors

Jimin Kang[1], Koustav Jana[2], Nathaniel Safron[3], Donglai Zhong[3], Shengman Li[2], H.-S. Philip Wong[2]
[1]Department of Electrical Engineering, Stanford University, [2]Stanford University, [3]Taiwan Semiconductor Manufacturing Company

We present H-shaped quantum-well channel to overcome the mobility-threshold voltage (μFE–VT) trade-off associated with channel thickness (tCH) modulation in oxide semiconductor FETs (OSFETs). Compared to nano-laminated, Ω-shaped well, and single-material channel with equal tCH, the H-shaped quantum-well achieves the highest μFE, minimal negative VT shift, and the lowest transition region voltage (VTR). A dual-gated $In_2O_3/SnO_2/In_2O_3$ (ISI) H-shaped quantum-well channel demonstrates μFE = 33.8 $cm^2/V{\cdot}s$, VT = 0.15 V, and VTR = 0.47 V. Physics-based analytical model confirms the quantum confinement effect and explores μFE–VT design space via barrier height/thickness tuning. This concept is further validated in $WO_x/In_2O_3/WO_x$ and doping-profile-engineered In2O3. The H-shaped quantum-well exhibits minimal VT roll-off (ΔVT = 36 mV from 1 μm to 60 nm channel lengths), and excellent positive bias stress reliability (ΔVT = 10 mV at 2 MV/cm for 1000 s).

2:45 PM

22-2 First Demonstration of Positive Threshold in Dual-Gated ITO FETs with ZrO_2 Dielectric

Dylan James Matthews, Md Sazzadur Rahman, Jiazheng Chen, Arijit Sarkar, Avantika Chopra, Sanjeev Chauhan, Anthony Fernando, Tania Roy
Electrical and Computer Engineering(ECE), Duke University

We demonstrate the first enhancement-mode dual-gated (DG) indium tin oxide (ITO) transistors with positive threshold (V_{Th}), addressing the negative V_{Th} challenge in amorphous oxide semiconductor (AOS) devices for monolithic 3D integration. Our novel strategy – optimized ITO stoichiometry, ZrO_2 dielectrics and interface dipole engineering – overcomes top-gate-induced doping, achieving positive V_{Th} in four gate stacks. The ZrO_2 DG stack excels with record-high I_{ON} (projected 1.25 mA/μm at 20 nm), low SS (<100 mV/dec), minimal hysteresis (<11 mV), and robust stability under voltage stress and 1000 DC sweep cycles, retaining positive V_{Th} up to 125°C. Dipole-engineered bilayers, while effective, show higher hysteresis and lower reliability. These pioneering results establish single-dielectric ZrO_2 stacks as optimal for high-performance, reliable AOS logic in 3D-integrated systems.

3:10 PM

22-3 Atomic Layer Deposited Amorphous Ge-Doped Indium Oxide with Enhanced Mobility–Stability–Crystallization Resistance Tradeoff

Yushan Lee[1], Xingtian Wang[1], Jiahui Duan[1], Guanyu Zhou[1], Md Abdullah Al Mamun[2], Yiyan Yu[1], Pengyu Ren[1], Sizhe Ma[1], Shan Deng[1], Zijian Zhao[1], Yixin Qin[1], Alessandra Leonhardt[3], Serges Zambou[3], Aditya Chauhan[3], Ranjith Ramachandran[3], Glen Wilk[4], Xiao Gong[5], Kyeongjae Cho[2], Christopher Hinkle[1], Kai Ni[1]
[1]University of Notre Dame, [2]University of Texas at Dallas, [3]ASM Helsinki, Finland, [4]ASM Phoenix, USA, [5]National University of Singapore

This work presents Ge-doped In_2O_3 (Ge:In_2O_3) as a high-performance oxide semiconductor with a favorable tradeoff among mobility, stability, and crystallization resistance, enabling both <400 °C BEOL 3D integration and >500 °C 1T1C DRAM processes. Compared to Si:In_2O_3, Ge:In_2O_3 offers: (i) higher mobility (42 cm^2/V·s at 1:5 doping) due to delocalized Ge 4s orbitals, supported by DFT and experiments; (ii) better thermal and bias stability from higher oxygen vacancy formation energies, maintaining amorphous structure post-500 °C anneal (Si stable to 800 °C); (iii) excellent device performance (~60 mV/dec, ~10^{10} ON/OFF, and ON currents of 150 μA/μm for Si, 330 μA/μm for Ge); and (iv) high reliability with |ΔVTH| < 100 mV under ±2 V stress for 1000 s using 5–10 nm dielectrics. These results position ALD-grown Ge:In_2O_3 as a unified solution for next-generation 3D integration.

3:35 PM

22-4 First Demonstration of BEOL-Compatible P-Type Gate-All-Around (GAA) Nanosheet FETs Featuring $Te_{0.8}Se_{0.2}$ Channel (L_{ch} down to 100 nm), In_2O_3/ITO Gate, and Thermal Budget ≤ 150 °C

Yiyuan Sun[1], Ying Xu[2], Zijie Zheng[2], Zhilun Zhang[2], Gengchiau Liang[3], Kai Ni[4], Xiao Gong[2,5]
[1]Department of Electrical and Computer Engineering, National University of Singapore, [2]National University of Singapore, [3]National Yang-Ming Chiao Tung University, [4]University of Notre Dame, [5]Technology and Research (A*STAR)

We demonstrate the first BEOL-compatible p-type gate-all-around (GAA) nanosheet (NS) FET, fully fabricated at temperatures below 150 °C, marking a key milestone toward BEOL p-type FETs for monolithic 3D integration (M3DI). This achievement is enabled by a PVD-deposited $Tellurium_{0.8}Selenium_{0.2}$ ($Te_{0.8}Se_{0.2}$) channel, an In_2O_3/ITO gate formed by ALD and sputtering, and a novel channel release process employing an ITO sacrificial layer, all of which are large-scale manufacturable. The fabricated $Te_{0.8}Se_{0.2}$ NS FET with a channel length (L_{ch}) of 100 nm demonstrates the record-high on-current of 73 μA/μm at $V_{OV} = V_{DS} = -2$ V, the highest reported among BEOL-compatible p-type FETs. The benefits of the GAA NS architecture are demonstrated by a 71% reduction in DIBL at L_{ch} = 100 nm, along with notable improvements in *SS* and I_{on}/I_{off} ratio. Furthermore, negative bias temperature instability (NBTI) assessments confirm robust bias stress reliability, underscoring the technology's promise for reliable and scalable M3DI.

4:25 PM

22-5 Contact Length Scaling Enables Suppression of Contact Doping in Oxide Semiconductor Transistors

Chi-Hsin Huang[1], Koustav Jana[2], Samantha Afra van Rijs[2], H.-S. Philip Wong[2]
[1]Electrical Engineering, Stanford University, [2]Stanford University

We show that scaling contact length together with the channel length of oxide semiconductor transistors (OSFETs) suppresses threshold voltage roll-off, subthreshold swing degradation, and bias stress instability commonly seen in short-channel oxide transistors. Reducing the contact length decreases the generation of excess carriers induced by contact doping. Enhancement-mode ITO transistors with a 60 nm channel and 40 nm contact length achieve an on-current of 120.7 μA/μm, an off-current of 10^{-8} μA/μm, and a V_T of 0.66 V at 0–2 V gate bias. Enhancement-mode operation persists at an 80 nm contacted gate pitch. These results demonstrate that "short-channel effects" in OSFETs can be mitigated by scaling both channel and contact lengths simultaneously, highlighting their potential for low-power, three-dimensional integrated circuits.

4:50 PM

22-6 Can High-Temperature PBTI Testing Predict Dual-Gated ITO FET Long-term Reliability?

Md Sazzadur Rahman[1], Dylan Matthews[2], Bingyu Zhang[2], Hintsa Fisseha Hishe[2], Marco A. Villena[3], Jiazheng Chen[2], Avantika Chopra[2], Sanjeev Chauhan[2], Anthony Fernando[2], Gaurav Thareja[4], Tania Roy[2]
[1]Department of Electrical and Computer Engineering, Duke University, [2]Duke University, [3]University of Granada, [4]Applied Materials Inc.

We study positive bias temperature instability (PBTI) in a dual-gated (DG) indium tin oxide (ITO) FET at room temperature (RT), 85 °C and 125 °C. We report record low threshold voltage shift, ΔV_{th} ~ -37.8 mV for top gate stress (2 V) at 125 °C. At higher temperatures, ΔV_{th} is reduced and V_{th} recovery is slower than at RT under identical stress. Simulations show elevated temperatures anneal shallow traps and activate deep-level traps, but these effects are temporary. These observations question the validity of elevated temperature accelerated reliability testing for ITO transistors, as it may not reflect long-term AOS device instability.

5:15 PM

22-7 On the Statistical Variability of High-Performance Crystalline InGaOx Nanosheet Oxide Semiconductor FETs

Xingyu Huang[1], Kota Sakai[1], Anlan Chen[1], Takanori Takahashi[2], Mutsunori Uenuma[3], Yukiharu Uraoka[2], Takuya Saraya[1], Toshiro Hiramoto[1], Masaharu Kobayashi[1]
[1]The University of Tokyo, [2]Nara Institute of Science and Technology, [3]National Institute of Advanced Industrial Science and Technology

Crystalline oxide semiconductor nanosheet FETs (OS NSFETs) are promising devices for monolithic 3D integration and memory applications. However, its characteristics variability has not been fully understood yet. In this work, we have fabricated poly-InGaOx (IGO) FET array with L_g=60nm and systematically investigated characteristics variability. While poly-IGO FETs outperform amorphous IGO (a-IGO) FETs and compete with Si FETs in terms of on-current and transconductance, electrostatics (V_{th} and subthreshold swing) are prone to high variability compared to a-IGO FETs, which is attributed to the finite grain size and grain boundaries. This work provides the foundation for R&D of high-performance crystalline OS NSFETs for manufacturing.

Session 23: RSD | Reliability Across Materials and Mechanisms

2:15 PM, Continental 6

Co-Chairs: Huimei Zhou, IBM and Sumy Jose, NXP

In this session, different reliability mechanisms related to TDDB, thermal, ESD, endurance etc. are studied together with various defect-dominated mechanisms on ovonic threshold switching devices, Si, GaN, GeAsSe, ZrO etc. -based devices. First, advanced low-percentile MoL TDDB lifetime is studied, followed by a paper on three-dimensional thermal conductivity distribution measurements of interconnect stacks down to 3nm process nodes. The third paper explores full-Chip ESD with ultra-low leakage current on GaN-based devices. Fourth, GeAsSe ovonic threshold switching devices are studied based on nanosecond-scale transient defect spectroscopy. Next, localized ferroelectric phase variation of scaled FeFETs are studied and discussed from both experiment and modeling. The sixth paper presents decoupling capacitors with ultra-high-k dielectrics. The final two papers explore robust endurance and stable Vt at 125C in Si-channel FeFETs and time-frequency polarization modeling of high-density 3D MIM capacitors.

2:20 PM

23-1 Advanced Low-Percentile MoL TDDB Lifetime by Non-parametric Reliability Modeling

An Shun Teng, Ching Jing Wu, Pei-Jean Liao, Jen-Hao Lee, Ryan Lu, Y.K. Hwang
Taiwan Semiconductor Manufacturing Company

Process-induced variability in reduced MoL poly control gate (PC) and diffusion contact (CA) dimensions poses challenges for TDDB analysis, particularly in fitting low-percentile distributions. This work introduces a spacing-adjusted constant electric-field stress (CES) method, benchmarked against the conventional constant-voltage stress (CVS) model, validating intrinsic acceleration factors. To improve low-percentile reliability assessments, a non-parametric statistical sampling approach combined with localized Physical Failure Analysis (PFA) extends

lifetime projections. The streamlined model simplifies TDDB evaluation while offering robust lifetime projections for Figure of Merit (FOM) assessments.

2:45 PM

23-2 First Demonstration of Three-Dimensional Thermal Conductivity Distribution Measurements of Interconnect Stacks Down to 3nm Process Nodes

Zifeng Huang, Kaiqi Yang, Yiyun He, Zuoyuan Dong, Zixuan Sun, Xiyuan Tang, Ming Li, Runsheng Wang, Zhe Cheng
Peking University

We measured the three-dimensional thermal conductivity (TC) distribution of interconnect stacks via layer-by-layer time-domain thermoreflectance mapping. Additionally, steady-state thermoreflectance was employed to measure the cumulative thermal resistance across layers. The effective TCs of most interconnect layers of chips with different process nodes (28 nm - 3 nm) are only 1-2 W/m-K, significantly lower than the TC of bulk copper, resulting from high thermal boundary resistance at interfaces, dense grain boundaries, and amorphous metal components. Thermal models were developed to predict the temperature distribution in double-side integrated transistors with a novel cooling strategy. This work establishes quantitative thermal metrology for device/packaging design.

3:10 PM

23-3 Full-Chip ESD for the Monolithic GaN-Based Gate Driver With Ultra-Low Leakage Current

Sheng-Hsi Hung[1], Po-Jui Chiu[1], Chi-Yu Chen[1], Xiao-Quan Wu[1], Yu-Ting Huang[1], Chien-Wei Cho[1], Ke-Horng Chen[1], Ying-Hsi Lin[2], Tsung-Yen Tsai[2], Shian-Ru Lin[2], Hann-Huei Tsai[3]
[1]National Yang Ming Chiao Tung University, [2]Realtek Semiconductor, [3]Taiwan Semiconductor Research Institute

This paper proposes a normally-off ESD clamp with zero transient leakage for monolithic GaN ICs. Area is minimized by integrating the timing and voltage detection functions. A fast gate-charging path improves discharge capability. The risk of overvoltage during ESD events is mitigated. Field plate structures are optimized to reduce intrinsic capacitance and leakage, enhancing response speed. Experimental results demonstrate that the proposed circuit withstands 6 kV under HBM and 100 V under MM tests. When applied in an ACF converter, the ESD clamp enables compliance with ENERGY STAR 9.0 standards, achieving a power consumption of 98 mW in no-load mode.

3:35 PM

23-4 Nanosecond-scale Transient Defect Spectroscopy of Switching and Relaxation Dynamics in GeAsSe Ovonic Threshold Switching Devices

Nishant Saxena[1], Chengcheng Wang[2], Zeyu Hu[2], Guosheng Wang[2], Zheng Chai[3], Wei Zhang[2], Daniele Garbin[4], Robin Degraeve[4], Sergiu Clima[4], Taras Ravsher[4], Andrea Fantini[4], Jian Zhang[2], Attilio Belmonte[4], Gouri Sankar Kar[4]
[1]School of Engineering, Liverpool John Moores University, [2]Liverpool John Moores University, [3]Xi'an Jiaotong University, [4]IMEC

Ovonic Threshold Switch (OTS) devices are crucial selectors in high-density memory, suppressing sneak-path leakage. We present a nanosecond-scale transient defect spectroscopy study on GeAsSe OTS devices, revealing a two-stage switching process governed by two defect groups. Group D1 defects delocalize rapidly (<0.4 ns) triggering fast switch-on at high fields, but relax slowly (>µs). Group D2 defects activate later at lower fields and higher currents, relaxing rapidly (~ns) on switch-off. The study shows how operating conditions modulate these dynamics, offering crucial temporal insights and validating prior models to guide high-speed, reliable OTS selector design.

4:25 PM

23-5 On the Localized Ferroelectric Phase Variation in Scaled FeFET: Experiments and Modeling

Omkar Phadke[1], Minji Shon[1], Stuart Wodzro[1], Halid Mulaosmanovic[2], Stefan Dünkel[2], Sven Beyer[2], Asif Khan[1], Suman Datta[1], Shimeng Yu[1]
[1]Georgia Institute of Technology, [2]GlobalFoundries

In this work, the impact of "localized" ferroelectric/dielectric (FE/DE) phase variation on FeFET is evaluated. A new phenomenon, namely, Localized Phase Distribution Induced Threshold Voltage Shift (LPTS), is experimentally

characterized. Here, a scaled FeFET exhibits an asymmetric current readout that leads to a threshold voltage (V_{TH}) shift that is dependent on the direction of current flow (when source and drain are flipped). The device characterization with device-to-device and cycle-to-cycle variations is performed on 28 nm bulk Si FeFET technology platform. The supporting TCAD simulations point towards the presence of an uneven number of FE/DE phases near Source-Channel vs Drain-Channel boundary to be the main cause of asymmetric current readout. Further, the simulation indicates that the novel non-destructive "capacitive" readout of a scaled FeFET to be more robust towards phase variation compared to the traditional current readout as the LPTS effect is amortized

4:50 PM

23-6 Next-generation Embedded Decoupling Capacitors for Advanced CMOS Technology: from Ferroelectrics to Ultra High-k Dielectrics

Sou-Chi Chang, Christopher Neumann, Bharath Bangalore Rajeeva, Marc Jaikissoon, Katie Neilson, Kaan Oguz, I-Cheng Tung, Sudarat Lee, Christopher Jezewski, Gary Allen, Thomas Hoff, Dominique Adams, Rob Jordan, Emmannuel Khora, Felipe Bedoya, Jeff Bielefeld, Brian Barley, Charles Winslow, Ryan Pearce, Haydee Kim, Ece Alat, Quinlan Smith, Che-yun Lin, Wangzhou Wu, Ibrahim Azad, Juichin Alcantara, Bernal Granados Alpizar, Nicolas Butzen, Adedapo Oni, Scott Clendenning, Tristan Tronic, Matthew Metz, Uygar Avci
Intel Corporation

In this paper, promising MIM options ranging from ferroelectric HZO to ultra high-k dielectric such as TiO and STO are demonstrated for DCAP applications, offering a multi-generational capacitance boost over state-of-the-art without compromising reliability metrics such as leakage, capacitance drift, and breakdown. This work demonstrates the potential of a range of MIM capacitance density enhancements, which are both stable and low leakage, for next-generation advanced CMOS technologies.

5:15 PM

23-7 Experimental Demonstration of Robust High Temperature Operation with 10^4 Endurance and Stable V_{Th} at 125 °C Enabled by IL Scavenging in Si Channel FEFETs

Priyankka Ravikumar[1], Andrea Padovani[2], Chinsung Park[1], Prasanna Venkatesan Ravindran[1], Malak Desouky[1], Shimeng Yu[1], Luca Larcher[3], Gaurav Thareja[3], Kijoon Kim[4], Kwangyou Seo[4], Kwangsoo Kim[4], Wanki Kim[4], Daewon Ha[4], Asif Khan[1]
[1]Georgia Institute of Technology, [2]University of Modena and Reggio Emilia, [3]Applied Materials, [4]Samsung Electronics

Endurance degradation in conventional Si-channel ferroelectric FETs (FEFETs) is predominantly governed by trap generation within the interfacial SiO_2 (IL) layer, where elevated electric fields accelerate defect generation, leading to polarization screening and progressive memory window (MW) collapse. In this work, we demonstrate that using scavenging to thin the IL layer induces a sub-stoichiometric SiOx, with higher capacitance. This IL engineering results in a new mode of FEFET failure driven by the underlying FE rather than the IL. This FE-driven degradation paradigm yields three critical advantages: (1) suppression of subthreshold slope(SS) (2)MW stability over extended cycling, and (3)reduction in read-after-write latency. We also demonstrate robust MW at 125°C with 10^4 cycles endurance and low SS. Further, we quantify trap generation and electric fields in stack using Ginestra™ simulations. These findings establish a pathway to high endurance, stable V_{Th} FEFETs via the implementation of an IL engineering-driven control of ferroelectric degradation.

5:40 PM

23-8 Unified Time-Frequency Polarization Modeling in High-Density 3D MIM Capacitor Arrays

Abygaël Viey[1], Julien Delprato[1], Marios Barlas[2], Bruno Reig[1], Alexis Krakovinsky[1], Charles Gresset[1], Xavier Federspiel[2], David Roy[2], Xavier Garros[1]
[1]CEA-Leti, [2]STMicroelectronics

This paper introduces a unified physical model for high-density 3D MIM capacitors in the BEOL. Capacitance *C(f)* data (1 kHz - 20 GHz) and RC simulations yield a fast charging time (~10 ps), excluding RC delay as the source of long current transients $I_D(t)$ (1-10 s). Instead, a polarization model with a single exponent n=0.725 accurately predicts $I_D(t)$, *C(f, V)*, and *G(f, V)* over voltage, temperature, and area. The dielectric response and

thermally activated $I_D(t)$ support a hopping relaxation mechanism. This framework also bridges with recent AC-TDDB models, offering a compact, physically consistent approach to advanced capacitor design.

Session 24: MS | Beyond CMOS and Advanced Transport Model

2:15 PM, Continental 7-9
Co-Chairs: Raseong Kim, Intel and Aryan Afzalian, imec
This session includes 6 papers that describe advances in the area of beyond CMOS and advanced transport models. The first paper, by Vandwalle from X-FAB, describes an enhanced drift-diffusion Monte Carlo framework for predictive modeling of Photon Detection Probability in SPADs. The second paper, by Duan from Beihang University, reports a real time sensing storage computation paradigm for Ising computing. The third paper by Datta (invited) from Purdue University, will cover building block for P-Circuits. The fourth paper, by Shin from KAIST, presents a supercell mode-space method (SuperMS) to overcome the computational barrier for large scale quantum mechanical NEGF simulations. The fifth paper, by Absar from University of Waterloo, introduces a compact model for silicon quantum-dots spin qubits designed for integration with standard CMOS EDA tools. The last paper of this session by Zhang from Beihang University, presents a AI-driven full-chain simulation and full-parameter-space optimization framework for spin-CMOS hybrid circuits, spanning material-to-system flows.

2:20 PM
24-1 Drift-diffusion Monte Carlo Simulation for Predictive Photon Detection Probability in FSI SPADs Across UV–VIS–IR
Mathieu Vandwalle[1], Anton Greve[2], Daniel Gabler[2], Mirjam Skadell[2], Hannes Schmidt[2], Alexander Zimmer[2], Vinoothini AP Jeya Kumar[3], Yi Ling Ong[3]
[1]Technology-ProcDev, X-FAB Global Services GmbH, [2]X-FAB Global Services GmbH, [3]X-FAB Sarawak Sdn. Bhd.
An enhanced drift-diffusion Monte Carlo framework is presented for predictive modeling of Photon Detection Probability (PDP) in Front-Side Illuminated (FSI) Single-Photon Avalanche Diodes (SPADs) across the Ultra-Violet-Visible-Infra-Red (UV–VIS–IR) spectrum. This approach combines wavelength-dependent absorption modelling with measured quantum efficiency (QE) to estimate optical losses. Carrier trajectories capturing photo-generated carriers outside depletion region are used to compute avalanche breakdown probabilities, in contrast to conventional static electric-field-based models. For the first time, absorption and triggering profiles are integrated to predict PDP across various device areas and processes at various temperatures and voltages achieving relative errors below 5% across the entire UV–VIS–NIR spectrum. Beside this, PDP values exceeding 60% in the UV–VIS range were achieved for smallest area at room temperature and 2V excess bias. This framework provides a robust and scalable simulation tool for predictive SPAD design optimization without parameter fitting.

2:45 PM
24-2 Ultrafast Dual-Mode Ising Machine with Real Time Sensing-Storage-Computation Paradigm
Xinrui Duan, Yijiao Wang, Jiayao Wu, Wen Shi, Yutian Wang, Weisheng Zhao
Beihang University
This work presents the first real time sensing-storage-computation paradigm for Ising computing based on FDSOI and MRAM devices. Our key innovations include: (i) a monolithically integrated system that enables image sensing, localized acquisition and storage of coefficients, and in-situ spin-state updates—eliminating off-chip data conversion and transfer; (ii) the write current direction of magnetic tunnel junction (MTJ) is modulated by adjacent coupling coefficients, enabling highly parallel local Hamiltonian optimization; (iii) the intrinsic thermal stochastic switching behavior of on-chip MTJ devices is used for further Hamiltonian optimization, without complex peripheral circuits or control logic. The 28×28 prototype completes the full process from in-sensor coefficient generation to image segmentation efficiently within 180 ns, achieving higher-quality solutions in only two steps. The proposed approach delivers a 266× speedup and 1.56× area savings over state-of-the-art methods, demonstrating outstanding area efficiency and fast, robust Ising solving capabilities.

3:10 PM

24-3 Building Block For P-Circuits (Invited)

Supriyo Datta[1], Risi Jaiswal[2], Ming-Che Li[2], Lakshmi Anirudh Ghantasala[2,3], Behtash Behin-Aein[3], Joseph Makin[2], Shreyas Sen[2]

[1]ECE, Purdue University, [2]Purdue University, [3]Ludwig Computing

Device simulation has traditionally focused on simple building blocks. In the past, these were transistors or inverters or perhaps an arithmetic unit. Today, the biggest workloads are machine-learning algorithms, and ASICs are being designed specifically to accelerate such applications. While such accelerators commonly use full precision weights and activations, it is becoming common to replace them with low-bit quantized versions, supported by appropriate retraining when necessary. Here we focus on circuits that implement such low-bit quantized models using binary stochastic neurons (BSN) or p-bits. A suitable building block for such p-circuits is a unit that operates on each p-bit and updates it in accordance with a computation performed on n neighboring p-bits that it interacts with. The reciprocal of the energy needed to perform this elementary operation (ELOP) provides a metric, ELOP/S/W, that can be used to compare different implementations. We illustrate our methodology using a simple example of MNIST classification.

4:00 PM

24-4 Large-scale Quantum Simulations Enabled By Real-space-equivalent SuperMS Method

Mincheol Shin

Korea Advanced Institute of Science and Technology

In this work, we present a real-space-equivalent supercell mode-space method (SuperMS), a break-through method to overcome the computational barrier to enable large-scale quantum mechanical simulations employing the state-of-the art non-equilibrium Green's function (NEGF) method. SuperMS eliminates the limitations of conventional mode-space approaches and works as versatile as real-space method, while offering efficiency gains of several orders of magnitude and maintaining accuracy within a few percent. We showcase the power of SuperMS by considering two computationally demanding, large-scale problems of gate-leakage current of three-dimensional nanowire transistors and five-stack nanosheet field effect transistors with top contacts.

4:25 PM

24-5 Spin Qubit Compact Model Enabling Co-Simulation of Quantum Dynamics with Classical Control Electronics

Rubaya Absar, Zack D. Merino, Bingjian Du, Paul Jonghyuk Park, Chenao Yuan, Dylan Ma, Johan Alant, Shiyu Su, Jonathan Baugh, Lan Wei

University of Waterloo

Performance of Silicon quantum-dot (QD) spin qubits ultimately depends on cryogenic CMOS control electronics whose non-ideal behavior feeds back onto the qubits. Predicting system-level fidelity therefore demands a joint quantum–classical simulation flow that mainstream SPICE tools do not yet provide. We introduce a compact model for silicon QD spin qubits designed for integration with standard CMOS EDA tools. This model enables co-simulation of qubits and CMOS control circuits, capturing key non-idealities such as decoherence, and waveform distortion and quantization effects introduced by the CMOS control circuit. Simulations of single- and two-qubit gate sequences, performed within a commercial SPICE simulator, reproduce the expected quantum dynamics with high fidelity, confirming the validity of the model and QD/CMOS co-simulation framework. By addressing a critical gap in hybrid quantum–classical co-design and cosimulation, this work provides a practical, EDA-ready framework for modeling, designing, optimizing, and scaling silicon-based quantum processors.

4:50 PM

24-6 Confidence-Aware AI-Driven Full-Chain Simulation and Material-to-System Co-Optimization of Spin-CMOS Hybrid Circuits

Zhizhong Zhang[1,2], Xiaohan Sun[3,3], Jian Sun[3], Yu Sun[3,3], Maotong Cui[3,3], Xinyu Wang[3], Jiyu Xie[3], He Zhang[3], Xuan Li[4], Kaihua Cao[3], Gefei Wang[4], Hongxi Liu[4], Weisheng Zhao[3,3], Yue Zhang[3,3]

[1]Fert Beijing Institute, MIIT Key Laboratory of Spintronics, School of Integrated Circuit Science and Engineering,

Beihang University, [2]National Key Laboratory of Spintronics, Hangzhou International Innovation Institute, Beihang University, [3]Beihang University, [4]Truth Memory Tech. Corporation
We present the first AI-driven full-chain simulation and full-parameter-space optimization framework for spin-CMOS hybrid circuits, spanning material-to-system flows. We develop device-level AI-PDE solvers enhanced with Monte Carlo Dropout and Extended Fiducial Inference to quantify 95% confidence intervals for each physics module of circuit unit, validated via rigorous leave-one-out testing. A hierarchical graph neural network (GNN) with hypernetwork augmentation interconnects these solvers into array-level plug-and-play paradigm. We combine reinforcement learning (RL) with generative adversarial network (GAN) to explore the high-dimensional parameter space in polynomial time and demonstrate robust convergence behavior on industrial circuit-level benchmark tasks. Validation on the SMIC180-nm PDK and MRAM design demonstrates 10^3× acceleration, <500 SPICE runs to convergence, and optimized designs delivered in minutes, outperforming state-of-the-art conventional strategies.

Session 25: SMB | MEMS and Sensors for Touch, Flow, Vision, and Gas Detection

2:15 PM, Imperial A
Co-Chairs: Sophie Giroud, CEA-Leti and Mario Kiuchi, Sumitomo Precision Products
This session includes 7 papers highlighting recent advances in MEMS and sensors for tactile perception, flow control, vision processing, and gas detection. The first paper, by Yingjie et al., reports on a CMOS-compatible tactile sensor array that achieves an unprecedented density of 444,000 pixels/cm^2, enabling highly precise touch detection. The second is an invited paper, presented by Eun Sok Kim from the University of Southern California. This paper explores acoustic droplet ejection as a novel approach for semiconductor chip handling. The third paper, by Wenhua et al., demonstrates a monolithically integrated architecture that couples a supercapacitor with an optoelectronic memristor array, paving the way for energy-autonomous in-sensor image processing. The fourth paper by Shin et al. is proposing a mixed gas discrimination sensor with a self-cancelation circuit offering compact and low-power consumption. The three next papers are presenting H_2 gas sensors held with different technologies and material-based devices. Feng et al. reports a Self-biased Pd/Ag-HEMT hydrogen sensor integrating an ultrathin AlGaN layer. The next presentation by Chiang et al. explores the use of TiO_2 all-RRAM as Gas sensor. Finally, Doris Ng et al. presents two systems for H_2 monitoring, a purity sensor based on ScAlN pyroelectric detection and a leakage sensor using RRAM and Pd as sensitive layer.

2:20 PM
25-1 First Demonstration of a CMOS-Compatible Tactile Sensor Array with Record-High Density of 444,000 pixels/cm^2 for Precise Tactile Perception
Bowen Zhu[1], Yingjie Tang[2], Lang Wang[2], Siyu Zhang[2], Hui Zhang[3], Yan Wang[2], Chenbo Wu[2], Yitong Chen[2], Guolei Liu[2], Chen Liu[2], Fanfan Li[2], Dingwei Li[2], Huihui Ren[2], Qi Huang[3], Liaoyong Wen[2]
[1]Electronic and Information Engineering, Westlake University, [2]Westlake University, [3]Westlake Institute for Optoelectronics
This work presents the first integration of complementary metal-oxide-semiconductor (CMOS) technology with resistive pressure-sensitive films to achieve ultra-high-density tactile sensing array (640 × 512). The sensor array achieves a record high density of 4.44 × 10^5 pixels/cm^2, a 15-μm pixel resolution, and a 97.9 kPa^{-1} tactile sensitivity, enabling discrimination of fine textural features in miniaturized objects. It enables real-time monitoring of force contact trajectory with optical flow algorithms through inter-frame image comparison. Moreover, the ultra-high pixel density—and thus the large collected sensory data—facilitates a 100% accuracy within only 2 training epochs in font classification tasks through a simple Convolutional Neural Networks (CNN). This CMOS-compatible sensing approach enables precise and high-fidelity tactile perception.

2:45 PM
25-2 Acoustic Droplet Ejectors for Semiconductor Chip Ejection (Invited)
Eun Sok Kim
Electrical and Computer Engineering, University of Southern California
This paper presents self-focusing acoustic droplet ejectors for *on-demand* picking and placing of semiconductor chips. The ejectors, based on Fresnel Half Wavelength Bands for near field focusing, are capable of ejecting liquid droplets whose diameter is from 0.01 to 3.3 mm with 1 – 300 MHz focused ultrasound (FUS). With proper designs

and operational conditions with 1 – 2 MHz FUS, the ejected droplets can carry 0.4-mm-thick silicon chips with square sizes of 0.6x0.6 - 3.1x3.1 mm^2 as well as a rectangular size of 12x4 mm^2. Semi-automatic loading of silicon chips to the ejection site is obtained with a laser-machined plastic plate (with a guiding channel), which is placed at the top of the liquid over the ejector.

3:10 PM

25-3 Monolithically Integrated One-Supercapacitor-One-Optoelectronic Memristor Array for Energy-Autonomous In-Sensor Image Processing

Wenhua Wang[1], Zhihan Zhang[2], Dong Wang[2], Zhengxiao Wu[2], Ruizhe Zhang[2], Kai Liu[2], Lindong Wu[2], Xiaodong Huang[2]

[1]School of Integrated Circuits, Southeast University, [2]Southeast University

For the first time, we experimentally demonstrate a homogeneous 1-supercapacitor-1-optoelectronic memristor (1SC-1OM) pixel array for self-powered in-sensor image enhancement and encryption. The V_2O_5/LiPON/ZnO SCs exhibit remarkable 6-level energy storage capability and the V_2O_5/ZnO OMs show tunable violet light responses. By monolithically integrating the SC with OM through a shared bottom electrode, the 1SC-1OM array enables fully energy-autonomous in-situ image processing, powered solely by SCs. The image enhancement experiment demonstrates that 1SC-1OM array can significantly promote image quality while suppressing environmental noises. Lossless image encryption is achieved by generating encryption keys via OM resistance variations and storing the keys as energy in SCs. Validation using license plate numbers and fingerprint images confirms excellent encryption performance of the array, reducing pixel correlation in encrypted images to 0.0027 while maintaining 100% recognition accuracy in decrypted images.

4:00 PM

25-4 Innovative Mixed Gas Discrimination System Using Integrated Self-Cancellation Circuit

Hunhee Shin[1], Jinwoo Park[2], Donghee Kim[2], Kangwook Choi[2], Joon Hwang[2], Min-Kyu Park[3], Gyuweon Jung[2], Jong-Ho Lee[2]

[1]Department of Electrical and Computer Engineering, Seoul National University, [2]Seoul National University, [3]Gachon University

We propose a novel mixed gas discrimination system using dual FET-type gas sensors integrated with a self-cancellation circuit. The proposed system accommodates diverse mixed gas environments through adjustment of annealing and operating conditions, while achieving real-time readouts without additional signal processing. The fabricated system has a small area (110×90 μm^2) and consumes low power (1.2 mW @ 160 °C). The system also provides high sensitivity (0.56 V/ppm) and high linearity (R^2 = 0.988) based on application-specific priorities. As a practical demonstration, the fabricated system successfully monitors egg freshness by detecting an increase in H_2S levels selectively.

4:25 PM

25-5 Self-Biased Ultrahigh-Response Hydrogen Sensor Utilizing Subthreshold Operation of Pd/Ag-HEMT Featuring thin AlGaN Barrier

Cheng Feng, Shengjie Zhang, Hui Guo, Rong Zhang, Youdou Zheng, Dunjun Chen

Nanjing University

This work presents a transformative, self-biased H_2 sensor using a Pd/Ag-alloy (24.7 at% Ag)-gated HEMT with an ultrathin 4-nm AlGaN barrier. Near-zero threshold voltage (~0.5 V) enables subthreshold operation without external bias. Strategic d-band engineering downshifts the Pd d-band center by 0.15 eV, significantly weakening Pd-H bonds and suppressing β-phase hydride formation. This enhances H_2 adsorption energy by 54% and desorption capacity by 12.5× versus pure Pd. Operating optimally at 200°C, the sensor achieves an unprecedented 16,269% response to 200 ppm H_2—surpassing state-of-the-art self-biased HEMT sensors by over 1000×. It sets a new benchmark for zero-power, high-performance hydrogen sensing.

4:50 PM

25-6 Direct Atomic Layer Processing TiO2 RRAM Gas Sensor for Real-time All-RRAM Intelligent Gas Perception

Tung-Lin Chiang[1], Ming-Chun Hong[1], Shang-Rong Huang[1], Bo-Cheng Chiou[2], Chih-Sheng Lin[2], Yu-Hao Cheng[1], Mohammad Dehghan[3], Tomáš Ščepka[3], Matej Horský[3], Peter Švec[3], Micha Patrnciak[4], Tomáš Plecenik[4], Matthias Carnoy[5], Karolis Parfeniukas[5], Maksym Plakhotnyuk[5], Benjamin Borie[5], Boris Hudec[3], Tuo-Hung Hou[1]
[1]National Yang Ming Chiao Tung University, [2]Industrial Technology Research Institute, [3]Slovak Academy of Sciences, [4]Comenius University Bratislava, [5]ATLANT 3D

We propose a real-time, all-RRAM intelligent gas perception(RIGP) system, featuring three cross-layer innovations. At the process level, a direct atomic layer processing(DALP) technique eliminates lithography and etching, enabling high-quality TiO_2 deposition with variable thicknesses at designated locations. At the device level, a vertical TiO_2-based gas sensor, built using a standard RRAM structure, exhibits high H_2 responsivity. At the architecture level, the RRAM gas sensor enables seamless integration with an RRAM-based in-memory computing(IMC) platform. Variable TiO_2 thicknesses facilitate local feature extraction within the sensor array. The RIGP system, incorporating hardware-encrypted acceleration of reservoir computing(RC), achieves accurate real-time detection and inherent security.

5:15 PM

25-7 Miniature Hydrogen Purity (0.1% resolution) and Room Temperature Hydrogen Leakage (20 ppm sensitivity) Sensor Systems for Quality and Leakage Control towards Safe and Sustainable Adoption of Hydrogen

Doris Ng[1], Subhranu Samanta[1], Zekun Wang[2], Linfang Xu[1], Weida Toh[1], Murugan Arulchozhan[1], Hazwani Khairy[1], Zhixian Chen[1], Yuxi Cui[3], Lemuel Joel Kuehsamy[3], Song Jie Seah[3], Qixing Liu[2], Zhonghua Gu[1], Huanhuan Wang[1], Anwesha Mukherjee[2], Rachel Ang[1], Anmin Kong[1], Kristel Wee[1], Chong Pei Ho[1], Weiguo Chen[1], Qingxin Zhang[1], Fuu Ming Kai[3], Dan Zhao[2]
[1]Institute for Microelectronics (IME), Agency for Science, Technology and Research (A*STAR), [2]National University of Singapore, [3]National Metrology Centre (A*STAR)

We develop 2 novel H_2 gas sensor systems:(1)H_2 purity sensor with high resolution ~0.1%, allowing direct measurement of H_2 purity and differentiating H_2 between 99.8% and 99.9% and (2)back-end of line–compatible cross-point resistive random access memory-based H_2 leakage sensor achieving H_2 sensing at record low 20ppm level at room temperature for any kind of metal-insulator (semiconductor)-metal MI(S)M-based H_2 leakage sensors. The thumb-size H2 purity sensor developed with a novel design configuration using 20% scandium aluminum nitride-based pyroelectric detector allows for in-line real-time monitoring of H2 purity concentration. This H2 leakage sensor is highly selective, reliable and resistant to 95% RH.

Session 26: PMA | Advanced GaN Power Devices: Integration, Reliability and High Voltage

2:15 PM, Imperial B

Co-Chairs: Mengyuan Hua, Southern University of Science and Technology and Julien Buckley, CEA-Leti

This session includes work covering integration and reliability of GaN-based power devices. The first paper by Haohao Chen from Southern University of Science and Technology achieves hot electron hardening using selective p-GaN gate epitaxy. The second paper by Zheng Wu from the Hong Kong University of Science and Technology addresses crosstalk issues. The third paper by Hongchang Cui from the University of Hong Kong studies the short-circuit robustness of bidirectional GaN power devices. The fourth paper by Jingjing Yu from Peking University introduces an e-mode p-GaN gate HEMT with non-destructive drain breakdown. The fifth paper by Jiawei Cui from Peking University demonstrates a multi- to-single-channel structure with p-GaN gate and normally-off behavior. The sixth paper by Zhiqing Yang from Xidian University demonstrates a GaN fast avalanche device. The seventh paper by Yuan Qin from Virginia Tech demonstrates a multi-channel GaN monolithic bidirectional switch with 5 kV breakdown voltage. The eighth paper by Chenyue Chu from the University of Science and Technology of China demonstrates a complementary half-bridge with normally-off diamond p-FET and GaN n-FET.

2:20 PM

26-1 Hot-Electron Hardening in GaN-based Power Devices by Selective p-GaN Gate Epitaxy

Haohao Chen[1], Jinjin Tang[2], Junting Chen[2], Jianxun Liu[3], Siyuan Ye[2], Junlei Zhao[2], Zhaojun Liu[2], Jun Ma[2], Qian Sun[3], Mengyuan Hua[2]

[1]Department of Electronic and Electrical Engineering, Southern University of Science and Technology, [2]Southern University of Science and Technology, [3]Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences

A selective epitaxial p-GaN gate fabrication process is developed for GaN-based power HEMTs, which avoids p-GaN growth and etching processes in the access region, thereby achieving enhanced robustness against hot-electron effects. The SEG-HEMTs exhibit higher on-state breakdown voltage and enhanced semi-on stress reliability. Further reduction of R_{ON} can also be realized by incorporating gate recess. The SEG method provides a viable solution for high-power GaN devices requiring high reliability.

2:45 PM

26-2 A Crosstalk-Free GaN-on-Si Power Integration Platform for Integrated Half-Bridge Circuits and Bi-directional Switches

Zheng Wu[1], Zongjie Zhou[2], Longge Deng[2], Tao Chen[2], Yat Hon Ng[2], Yan Cheng[2], Yutao Geng[2], Kevin Jing Chen[2]

[1]Department of Electronic and Computer Engineering (ECE), The Hong Kong University of Science and Technology, [2]The Hong Kong University of Science and Technology

This work presents a holistic GaN-on-Si p-GaN gate double-channel HEMT power integration platform that supports monolithically integrated bridge circuits and bi-directional switch (BDS) that are free of substrate crosstalk and back-gating effects. The proposed platform solely involves planar HEMT fabrication process and only requires standard low-cost low-resistivity Si substrates. The double-channel structure creates a critical hole-spreading channel sandwiched between two closely coupled 2DEG channels. In the on-state, the hole-spreading channel is able to store and distribute holes injected from the *p*-GaN gate, effectively compensating buffer trapping and shielding the 2DEG channels from substrate-induced crosstalk. In the off-state, the holes quickly recombine with the 2DEG in the un-depleted access regions, preserving the voltage blocking capability.

3:10 PM

26-3 Intrinsically Exceptional Short-Circuit Robustness (t_{SC} > 30 μs, I_{SC} > 230 A) in Monolithic Bidirectional GaN-on-Si Power HEMTs

Hongchang Cui[1], Xiaoming Liu[2], Xin Yang[3], Hehe Gong[3], Rong Yang[2], Liang Song[2], Fu Chen[2], Jianping Wang[2], Pengju Kong[2], Yi Sun[2], Yuhao Zhang[3]

[1]Department of Electrical and Electronic Engineering, The University of Hong Kong, [2]InnoScience (Suzhou) Technology, [3]The University of Hong Kong

GaN monolithic bidirectional switch (MBDS) has gained broad traction recently. However, its short-circuit robustness has not yet been reported. This work fills this gap by evaluating short-circuit performance of a 650 V, 15 mΩ GaN-on-Si MBDS under a bus voltage of 400 V. The device demonstrates an exceptionally long short-circuit withstand time (t_{SC}) exceeding 50 µs, with a peak short-circuit current over 230 A, and withstands repetitive 30 µs short-circuit without degradations. This t_{SC} far surpasses typical system requirement (10 µs). To elucidate underlying mechanism, we construct two hybrid GaN BDS using discrete GaN HEMTs from a different vendor under two substrate biasing schemes. The hybrid BDS with *separate substrates* shows a t_{SC} < 1 µs, while the one with a *common substrate* exceeds 100 µs. This difference suggests that superior short-circuit robustness is inherent to MBDS. These results suggest promise of GaN MBDS for demanding automotive and industrial applications.

3:35 PM

26-4 E-Mode p-GaN Gate Punch-Through HEMT with Robust Non-Destructive Drain Breakdown

Jingjing Yu[1], Junjie Yang[2], Yunhong Lao[2], Sihang Liu[2], Han Yang[2], Xuelin Yang[2], Maojun Wang[2], Kevin Jing Chen[3], Bo Shen[2], Jin Wei[2]

[1]School of Integrated Circuits, Peking University, [2]Peking University, [3]The Hong Kong University of Science and Technology

This work presents an E-mode p-GaN gate punch-through HEMT (PT-HEMT) which could achieve repetitive non-destructive breakdown, addressing the long-standing issue in conventional GaN HEMTs—the lack of avalanche robustness. In the OFF-state, high drain bias induces full depletion of the punch-through gate, so the energy band along the 2DEG channel is punched through, resulting in a non-destructive breakdown. For 650-V rating, PT-HEMT with L_{GD} = 12 μm achieves a non-destructive breakdown voltage of 874 V. The punch-through breakdown exhibits a positive temperature coefficient. Under current-driven breakdown tests, the PT-HEMT survives over 100 repetitive breakdown cycles, with negligible degradation of threshold voltage, R_{ON}, and dynamic R_{ON}. Thus, drain voltage is clamped at punch-through voltage lower than that for destructive breakdown to dissipate transient overvoltage spikes. Therefore, the proposed PT-HEMTs are empowered with an inherent capability to handle transient overvoltage spikes in high-power switching applications.

4:25 PM

26-5 Multi-to-Single Channel HEMT (M2S-HEMT): An E-Mode p-GaN Gate Power Transistor Decoupling Threshold Voltage and ON-Resistance

Jiawei Cui[1], Xingyu Fu[2], Yuanyang Xu[2], Zhenghao Chen[2], Sihang Liu[2], Hengyuan Qi[2], Junjie Yang[2], Maojun Wang[2], Xuelin Yang[2], Bo Shen[2], Jin Wei[2]

[1]School of Integrated Circuits, Peking University, [2]Peking University

The development of high-voltage GaN HEMT with large L_{GD} is challenged by the trade-off between V_{th} and R_{ON}. This work proposes a GaN multi-to-single channel HEMT (M2S-HEMT) to combine the merits of E-mode p-GaN gate structure and the low-resistance multi-channel drift region. A fluorine-based plasma treatment was conducted before loading the sample into MOCVD chamber, which was found to effectively suppress Si contamination at the regrowth interface. The device exhibits an E-mode operation with V_{th} of +1.3 V. The blocking capability is over 3 kV. For a L_{GD} of 77 μm, the M2S-HEMT achieves a low R_{ON} of 28.2 Ω·mm and a $R_{ON,SP}$ of 24.82 mΩ·cm^2. **T**he total R_{ON} value of the device is among the lowest for GaN power HEMTs with similar L_{GD}. These results highlight the GaN M2S-HEMT structure as a competitive solution for developing next-generation, high-voltage power switches with significantly reduced conduction losses.

4:50 PM

26-6 First Demonstration of GaN Fast Avalanche Device

Zhiqing Yang[1], Ming Xiao[1], Zihan Lin[1], Yu Wen[1], Kui Dang[1], Xiufeng Song[1], Shenglei Zhao[1], Kai Su[1], Zeyang Ren[1], Yajie Wang[2], Xiaopeng Jing[2], Jinfeng Zhang[1], Yue Hao[1], Jincheng Zhang[1]

[1]Xidian University, [2]Xi'an Electronic Engineering Research Institute (XEERI)

We report the first experimental demonstration of a high-performance, high-voltage fast avalanche pulse device (APD) based on GaN. In this work, we fabricate fast APDs on native GaN substrates, incorporating slanted field plates and small-angle junction termination extensions. The device exhibits an ultra-high peak output power density of 1.68 GW/cm^2, a minimum rise time of 93 ps, a peak output voltage of 2.57 kV, and a steepened voltage range of 2.41 kV. The achieved power density surpasses that of Si and GaAs APDs by over two orders of magnitude and is tens of times higher than that of SiC counterparts. We developed an integrated GaN APD module featuring a diamond heat sink and GaN substrate thinning. The co-packaged energy storage capacitor and APD structure yields low thermal resistances of 0.73 and 26.1 K/W for devices with radii of 500 and 50 μm, respectively.

5:15 PM

26-7 5 kV, 20 mΩ·cm^2 Enhancement-Mode Multi-Channel GaN Monolithic Bidirectional Switch

Yuan Qin[1], Ming Xiao[2], Hongchang Cui[3], Xin Yang[3], Zineng Yang[3], Hehe Gong[3], Kai Liu[4], Ivan Kravchenko[5], Florin Udrea[6], Kai Cheng[4], Han Wang[3], Yuhao Zhang[3]

[1]CPES, Virginia Tech, [2]School of Microelectronics, Xidian University, [3]University of Hong Kong, [4]Enkris Semiconductor Inc., [5]Oak Ridge National Laboratory, [6]University of Cambridge

We demonstrate, for the first time, a multi-channel GaN monolithic bidirectional switch (MBDS), achieving symmetric breakdown voltage (*BV*) up to 5 kV, specific on-resistance of 20 mΩ · cm^2, and threshold voltage of 0.8 V in both voltage/current polarities. This *BV* is the highest reported in MBDS devices to date. This device is fabricated on a 150-mm 5-channel GaN-on-sapphire wafer and features a dual junction termination extension for electric field management. Enhancement-mode operation is achieved by a monolithic cascode architecture, which

integrates two low-voltage, single-channel HEMTs on either side of the multi-channel region. Furthermore, we evaluate the HTRB reliability and dynamic performance of the multi-channel GaN MBDS in comparison to a single-channel counterpart. The multi-channel MBDS exhibits reduced dynamic on-resistance (R_{on}) and less than 5% R_{on} shift after 1,000 s HTRB stress at 1.7 kV and 150 °C. These results suggest the promise of multi-channel GaN MBDS for medium-voltage power electronics.

5:40 PM
26-8 Demonstration of Complementary Power Circuit with Normally-Off Diamond P-FET and GaN N-FET
Chenyue Chu[1], Jiahong Du[2], Xinxin Yu[3], Longjie Jiang[2], Ziang Wang[2], Xuan Xie[2], Kun Tan[4], Changhui Zhao[4], Xi Tang[4], Shibing Long[2], Shu Yang[2]
[1]School of Microelectronics, University of Science and Technology of China, [2]University of Science and Technology of China, [3]Nanjing Electronic Devices Institute, [4]Anhui University
In this work, we demonstrate a normally-off HfO_2/H-diamond p-FET featuring a high ON/OFF current ratio (I_{ON}/I_{OFF}), a high Baliga's figure-of-merit (BFOM) and enhanced threshold voltage stability, which is verified by μs-level high-speed pulsed *I-V* characterizations. Based on the band alignment between the H-diamond and HfO_2 extracting from XPS, it is revealed that the diamond p-FET can deliver normally-off operation with the low-work-function gate metal depleting the 2DHG channel, whereas a low R_{ON} can be simultaneously realized with well-accumulated 2DHG in the access region. For the first time, a complementary power circuit featuring a normally-off diamond p-FET and GaN n-FET with matched conduction capability has been demonstrated, which can effectively stabilize the gate driving voltage compared with the counterpart with series connection of n-FETs, and show great potential for next-generation high-frequency and high-power electronics applications.

Session 27: Evening Panel | FETs at 100: Yesterday's Sparks, Tomorrow's Arcs -- Human Wisdom Meets AI Insight

8:00 PM, Continental 4-5
Moderator: John Wuu, Advanced Micro Devices
Panelists: Serge Biesemans, imec; Daewon Ha, Samsung Electronics; Chenming Hu, UC Berkeley; H.-S. Philip Wong, TSMC and Stanford University

- What were the most important breakthroughs in the 100-year journey of FETs?
- Which emerging innovations are most likely to propel FETs forward?
- What role will AI play in advancing FET design, materials, and manufacturing?
- Or, as heterogeneous integration and advanced packaging gain prominence, will the importance of FETs diminish…or become even more central?

When Julius E. Lilienfeld filed the patent for his Field Effect Transistor (FET) concept in 1925, decades before a working device could be built, few could have predicted the profound impact FETs would have a century later. Over the years, through countless innovations, FETs have demonstrated remarkable resilience and versatility. They have evolved in their topologies and materials and adapted to a wide array of applications from computing, memory, and storage to sensors, RF, and power management. To reflect on and celebrate the extraordinary journey of FETs through the last 100 years and to anticipate the exciting developments that lie ahead – this panel brings together experts from industry and academia, augmented by AI-driven insights, for a lively, thought-provoking discussion and debate

Wednesday, December 10, 2025

Session 28: EDT | Monolithic 3D and Novel Device Architectures for Post-Moore Computing

9:00 AM, Grand Ballroom A
Co-Chairs: Saptarshi Das, Penn State and Tania Roy, Duke University
This session highlights cutting-edge device and system innovations for computing beyond Moore's law, focusing on 2D semiconductors, monolithic 3D integration, novel memories, and neuromorphic/Ising accelerators. The presented works span from wafer-scale growth and CFET integration of 2D materials to BEOL-integrated memories (ferroelectric, capacitive, RRAM, PCM, and Amp Cell). Together, they demonstrate how heterogeneous integration,

low-temperature processes, and device–circuit co-design is converging to enable energy-efficient, scalable, and application-driven architectures. Beyond conventional CMOS scaling, these technologies offer transformative pathways for high-throughput computing, AI accelerators, and combinatorial optimization solvers. The session opens with an invited talk by Prof. Zhou Peng, highlighting wafer-scale 12-inch MoS_2 growth, integration of a high-yield (99.9%) RISC-V processor, and graphene-based flash memory advances that chart a clear path toward monolithic 3D Si–2D heterogeneous VLSI. This is followed by a presentation on collaborative work demonstrating wafer-scale vertical CFETs with single-crystal TiO2 dielectrics, achieving high-gain inverters, SRAM cells, and record-low standby power. The next paper reports the first BEOL integration of dual-gated W-doped In_2O_3 non-volatile capacitive memory on TSMC's 40 nm CMOS platform, setting records in Con/Coff (~64), endurance (>10^9 cycles), and enabling digital capacitive compute-in-memory (Cap-DCIM) with 146× performance-per-watt-per-area gains over analog CIMs. Building on this, IBM researchers present a novel disc-type phase-change memory (PCM) architecture that delivers ultra-low programming currents (~20 μA), over 100× higher resistance, and wide resistance windows spanning more than three orders of magnitude—ideal for analog neural inference. The session then features the first foundry-demonstrated monolithic 3D integration of Si CMOS, RRAM, and CNFETs on 200 mm wafers, achieving tucked sense amplifier arrays with 4× higher bandwidth and throughput compared to CMOS baselines. Shifting focus toward sensor-edge applications, the next talk introduces a CMOS-compatible IGZO TFT + HfO_2 capacitor 2T-2C compute cell, which eliminates ADC/DAC interfaces for near-pixel analog processing, reducing power by up to 93% while preserving CNN accuracy at 99.6%. Following this, KAIST researchers present a compact monolithic 3D oscillatory Ising machine based on biristor oscillators and ferroelectric capacitor couplers, achieving 4-bit tunable coupling, 14× density improvement, and 33× lower standby power in solving MaxCUT problems. The session concludes with the introduction of the "Amp Cell" memory architecture, a 3D-scalable design with fully decoupled read, write, and store paths (Si FET + OS FET + MIM capacitor), experimentally demonstrated on a 40 nm CMOS+X chip with >5× retention improvement and seamless scalability beyond SRAM and DRAM.

9:05 AM

28-1 Electronic Devices and System Integration in the Post-Moore Era (Invited)

Peng Zhou, Mingrui Ao, Yutong Xiang, Chong Wang, Shaohua Fan, Qicheng Sun, Chunsen Liu, Jing Wan, Wenzhong Bao

Fudan University

As modern VLSI approaches scaling and power limits, two-dimensional (2D) materials offer new opportunities for advanced logic and memory integration. This work showcases major advances in 2D materials for next-generation VLSI, including 12-inch monolayer MoS_2 growth via CVD, a MoS_2 RISC-V WUJI processor with 99.92% transistor yield and inverter gain >550, and graphene-based flash PoX memory with sub-1 ns programming, 1.8 V window, and >10^5 cycle endurance. These breakthroughs highlight the potential of 2D materials in logic, memory, and 2D–Si heterogeneous integration, enabling monolithic 3D architectures for energy-efficient, high-performance computing beyond Moore's Law.

9:30 AM

28-2 Wafer-Scale Low-Power 2D CFET Logic Enabled by Single-Crystal Dielectrics: Experimental Demonstration and 1-nm-Node Projection

Jing-Kai Huang[1], Ni Yang[1,2], Yu-Cheng Lu[3], Ji Zhang[4], Zeheng Wang[5], Yu-Ming Chang[1,6], Fangyuan Zheng[2], Lingqi Li[6], Haomin Liu[6], Lin-Yun Huang[1], Chenyang Li[1,6], Chengdong Yao[6], Pengfei Yang[2], Wanying Li[2], Wanqing Meng[6], Pak San Yip[2], Dong-Keun Ki[2], Kah-Wee Ang[6], Geng-Chiau Liang[7], Yi Wan[6], Sean Li[4], Vita Pi-Ho Hu[3], Lain-Jong Li[1,6]

[1]Nexstrom Pte. Ltd., [2]University of Hong Kong, [3]National Taiwan University, [4]University of New South Wales, [5]CSIRO, [6]National University of Singapore, [7]National Yang-Ming Chiao Tung University

We demonstrate a CFET technology that combines wafer-scale single-crystal TiO_2 dielectrics with vertically stacked n-MoS_2/p-WSe_2 channels. A release–transfer process yields atomically flat TiO_2 films with sub-nanometre CET. The resulting CFET inverters switch rail-to-rail across a broad voltage range (V_{dd} = 0.1 V to 1 V) with record-low standby power, high gain, and robust noise margins, validated by wafer-scale NAND/NOR gates. TCAD extrapolation at the 1-nm node indicates that 4N2P six-transistor SRAM arrays on this platform surpass projections

based on Si-GAA in terms of speed and energy, underscoring the viability of 2D semiconductors for next-generation logic.

9:55 AM

28-3 Monolithic 3D Integration of Dual-Gated ALD Oxide-Channel Non-Volatile Capacitive Memory on 40nm Si CMOS for Digital Compute-in-Memory

Junmo Lee[1], Chengyang Zhang[2], Siheon Park[2], Hyeonwoo Park[2], Tae-Hyeon Kim[2], Leo Jen-Chieh Liu[3], Elia Ambrosi[3], Mingyuan Song[3], Xinyu Bao[3], Meng-Fan Chang[3], Suman Datta[2], Shimeng Yu[2]
[1]Electrical and Computer Engineering, Georgia Institute of Technology, [2]Georgia Institute of Technology, [3]Taiwan Semiconductor Manufacturing Company

We experimentally demonstrate monolithic 3D (M3D) integration of dual-gated (DG) non-volatile capacitive memory (nvCAP) featuring an ALD W-doped In2O3 (IWO) channel on a TSMC N40 CMOS foundry chip. Key technical advancements include: i)Novel dual-gate engineering of nvCAP, effectively addressing the weak erase and poor retention challenges with oxide-channel ferroelectrics; ii)Successful BEOL integration of DG nvCAP on a foundry CMOS chip, achieving a non-destructive capacitive on/off ratio of ~64.4 at V_{read}=0 V—a record among reported BEOL-compatible nvCAPs; iii)Non-destructive read endurance of >10^9 cycles at |V_{read}|=1 V; iv)Introduction of the M3D non-destructive "capacitive" digital-compute-in-memory (Cap-DCIM) paradigm, achieving 146× higher figure-of-merit (TOPS/W×TOPS/mm^2) compared to leading analog CIMs and 111× lower static power versus SRAM-based DCIMs—enabled by highly scalable digital logic and BEOL-integrated DG nvCAPs. Finally, we experimentally validate the operational principle of M3D Cap-DCIM, by demonstrating BEOL capacitance-modulated FEOL transistor current amplification through the monolithically integrated DG nvCAPs on a foundry CMOS chip.

10:20 AM

28-4 Disc-type Phase-change Memory Devices for Analog In-memory Computing

Ghazi Sarwat Syed[1], Timothy Philicelli[2], Jesse Luchtenveld[2], Vara Prasad Jonnalagadda[2], Siddharth Gautam[2], Kevin Brew[2], Bart J Kooi[3], Vijay Narayanan[2], Nicole Saulnier[2], Abu Sebastian[2]
[1]Europe, IBM Research, [2]IBM Research, [3]University of Groningen

We introduce a phase-change memory (PCM) device architecture tailored for analog in-memory computing (AIMC). By confining the active, switchable volume to a compact disc-shaped geometry, this design enables a continuum of non-volatile states programmable at ultra-low currents and power (~40 μA/μW), even in standard undoped $Ge_2Sb_2Te_5$ (GST). Furthermore, the device exhibits a resistance window spanning over three orders of magnitude, while achieving higher overall resistance.

11:10 AM

28-5 Foundry Monolithic 3D Unlocks Large Throughput Benefits: 3D Memory with Tucked Sense Amplifiers + Logic using Heterogeneous Silicon CMOS + Resistive RAM + Carbon Nanotube FETs

Suhyeong Choi[1], Atharva Raut[2], Tong Wu[2], Samuel A. Dayo[1], Andrew D. Bechdolt[1], Gairik Dutta[1], Shengman Li[1], Dennis Rich[1], Raymond H. Yang[1], Andrew C. Yu[3], Mark Nelson[4], Max M. Shulaker[3], Robert M. Radway[1,5], Subhasish Mitra[1], Tathagata Srimani[2]
[1]Stanford University, [2]Carnegie Mellon University, [3]Massachusetts Institute of Technology, [4]SkyWater Technology, [5]University of Pennsylvania

We present the first foundry monolithic 3D (M3D) memory + logic circuits: full M3D memory arrays with tucked sense amplifiers (TSAs) driving digital logic. Achieved at SkyWater Foundry (90/130 nm, 200 mm wafers), our M3D process heterogeneously integrates Si CMOS, resistive RAM (RRAM), and carbon nanotube FETs (CNFETs) in the BEOL at ≤415 °C. The M3D design achieves 4× read bandwidth and 4× compute throughput at iso-latency and iso-footprint compared to a Si CMOS + RRAM baseline. This is enabled by (a) a robust BEOL-compatible process with CNFET + RRAM offering comparable performance and endurance to Si FET + RRAM, (b) ultra-dense vias at RRAM metal pitch, and (c) novel TSA-based 3D memory architectures. Our platform includes a complete M3D PDK, enabling further design innovations. This demonstration marks the first foundry-supported M3D memory + logic platform with practical integration of advanced device technologies.

11:35 AM

28-6 CMOS-Compatible Oxide TFT-Based 2T-2C Near-Pixel Analog Compute Cell with ADC/DAC-Free Sensor Interface for Energy-Efficient Feature Extraction

Seonjae Park[1], Seungyeob Kim[2], Seungmin Cheon[2], Taeseung Jung[2], Jinho Ahn[3], Sanghun Jeon[2]
[1]Graduate School of Semiconductor Technology, Korea Advanced Institute of Science and Technology, [2]Korea Advanced Institute of Science and Technology, [3]Hanyang University

This work presents a CMOS-compatible 2T-2C near-pixel compute cell based on oxide TFTs with an ADC/DAC-free Sensor interface, enabling energy-efficient analog-domain feature extraction for vision systems. The cell integrates IGZO TFTs and HfO_2 capacitors to support signed-weight MAC operations without requiring ADC/DAC. A comparator circuit allows direct analog-to-pulse conversion from a photodiode, enabling one-to-one sensor-to-cell mapping. The fabricated device demonstrates linear weight updates, strong retention, and calibrated MAC accuracy (R^2 = 0.992). Compared to 2T-0C, it achieves up to 93% lower power and up to 43% smaller footprint in peripheral circuits. CNN simulations using 2T-2C-based feature maps achieve 99.6% classification accuracy, highlighting its potential for low-power, high-accuracy near-sensor vision.

12:00 PM

28-7 A Compact and Energy-Efficient Monolithic-3D Oscillatory Ising Machine with Novel Ferroelectric Multi-Weight Coupling

Joon Pyo Kim[1], Song-hyeon Kuk[1], Jaeyong Jeong[1], Hyeon Jun Lee[1], Juhyuk Park[1], Bong Ho Kim[1], Jongmin Kim[2], Sanghyeon Kim[1]
[1]KAIST, [2]KANC

Ising machines have emerged as promising hardware accelerators for solving NP-hard combinatorial optimization problems (COPs) by minimizing the Ising Hamiltonian. In this work, we present a highly scalable monolithic 3D (M3D) oscillatory Ising machine (OIM), composed of bi-stable resistor (biristor)-based oscillators and a novel ferroelectric capacitor (FeCap) serving as a reconfigurable, multi-state coupling router. The FeCap enables 4-bit analog tuning of inter-oscillator coupling strength using a single device, allowing direct embedding of weighted Ising edges. The functionality of the M3D OIM is experimentally validated through solving MaxCUT problems, with performance further assessed via circuit-level simulations. Leveraging the intrinsic parallelism of oscillators and the vertical integration of M3D architecture, the proposed OIM achieves ×14 higher bitcell area density and over ×33 lower stand-by power consumption compared to CMOS ring oscillator-based implementations. This work offers a compact, low-power, and reconfigurable OIM platform with significant potential for solving large-scale and densely connected COPs.

12:25 PM

28-8 Amp Cell: 3D-Scalable Gain Cell Memory with Decoupled Read, Write, and Store

Shuhan Liu[1], Koustav Jana[1], Leo Jen-Chieh Liu[2], Elia Ambrosi[2], Mingyuan Song[2], Louis Liu[3], Jing Wang[3], Subhasish Mitra[1], Xinyu Bao[2], Meng-Fan Chang[2], John R. Hu[3], John Y. Chen[3], Jian Chen[1], H.-S. Philip Wong[1]
[1]Stanford University, [2]TSMC, [3]Nvidia Corporation

Scaling challenges of SRAM and DRAM stem from their inherently coupled read and write paths, creating conflicting design requirements that progressively degrade noise margins at advanced technology nodes. While gain cell memory decouples the read and write paths, its storage mechanism remains coupled with read and write. This work presents the design and experimental demonstration of a 3D-scalable amplified memory (Amp Cell) featuring complete decoupling of read, write, and store: Si FET for read, oxide semiconductor FET for write, and compact MIM capacitor for store. Through systematic device-array-circuit co-design, the Amp Cell requires no additional process steps, control signals, or significant circuit modifications compared to gain cell memory, while improving retention by >5× at VDD = 1 V and enabling both self-refresh capability and 3D scalability. The nT(nC) Amp Cell establishes a continuum bridging 1T1C DRAM and 2T gain cell, combining the advantages of both while eliminating their fundamental limitations.

Session 29: MT | Oxide Semiconductors for Memory

9:00 AM, Grand Ballroom B

Co-Chairs: Attilio Belmonte, imec and Kyunghwan Lee, Samsung

This session explores cutting-edge memory technologies leveraging oxide semiconductors. It opens with a presentation from Kioxia, showcasing a pioneering IGZO-based 3D DRAM test vehicle featuring eight stacked layers. Samsung follows with insights into another novel DRAM integration approach, highlighting a sub-10 nm vertical transistor architecture with an oxide semiconductor channel. Next, Changxin Memory Technology presents significant advancements in optimizing the performance and reliability of IGZO-based select transistors within multi-tiered memory structures. The session then shifts focus to alternative DRAM architectures, specifically 2T0C and 3T0C gain cells. The fourth talk, given by Semiconductor Energy Laboratory, introduces a 64-kbit In_2O_3-based 2T0C array implemented on Si CMOS, along with enhancements to the programming scheme. In the fifth presentation, Stanford explores hybrid gain cells that combine oxide semiconductors with silicon-based transistors, demonstrating superior performance compared to SRAM at technology nodes down to 5 nm. This is followed by a talk from Georgia Tech, offering an in-depth look at transistor architecture optimization in 3T0C gain cells, aimed at mitigating the impact of coupling capacitance on retention. Concluding the session, Samsung presents a novel FeFET structure based on oxide semiconductors, featuring a channel-all-around (CAA) MFMIS gate stack. This innovative design delivers excellent memory window and endurance, targeting 1T embedded memory applications.

9:05 AM

29-1 Highly Stackable Oxide-semiconductor Channel Transistor Technology for Future High-density and Low-power 3D DRAM

Mutsumi Okajima[1], Takamitsu Ochi[2], Tsuyoshi Sugisaki[2], Takanori Akita[2], Fumiya Kimura[2], Kiichi Sato[2], Isamu Ujiie[2], Sota Mochizuki[2], Masayoshi Iwayama[2], Atsuko Sakata[2], Shinji Miyano[2], Kazuki Eshiro[2], Takafumi Masuda[2], Tomomasa Ueda[2], Nobuyoshi Saito[2], Keiji Ikeda[2]

[1]Frontier Technology R&D Institute, Kioxia Corporation, [2]Kioxia Corporation

We demonstrate a highly stackable and ultra-low leakage oxide-semiconductor channel transistor technology for 3D Oxide-semiconductor Channel Transistor DRAM. It uses a replacement-type InGaZnO cell transistor with a novel Channel-Side WL configuration designed for Z-pitch scaling. This approach reduces process cost and improves multi-tier compatibility by utilizing mature Oxide/Nitride stacks and deposited channels. The prototype cell transistor demonstrates a high on-current of over 30μA/cell and an excellent on-off ratio >10^{13}. An 8-layer stacked cell transistor was successfully fabricated, confirming excellent transfer characteristics across all layers. We conclude that 3D OCTRAM is a promising candidate for future high-density, low-power 3D DRAM applications.

9:30 AM

29-2 Thermally Robust, Highly Reliable Amorphous Oxide Semiconductor Transistors for sub-10nm Cell-On-Peri Vertical Channel DRAM Transistors

Kyooho Jung[1], Ha-Jun Sung[2], Yunjin Jang[2], Kwang-Hee Lee[2], Daehyuk Son[2], Youngkwan Cha[2], Jung-kyun Kim[2], Na-Rae Han[2], Hyerim Hong[2], Moonil Jung[2], Jee-Eun Yang[2], Sangwook Kim[2]

[1]Samsung Advanced Institute of Technology, Samsung Electronics, [2]Samsung Electronics

We demonstrate, for the first time, thermally stable amorphous InGaO (a-IGO) vertical channel transistors with a 100 nm channel length that endure up to 550°C, enabling their integration into monolithic cell-on-periphery DRAM architectures. After 550°C N_2 annealing, the devices exhibit a threshold voltage shift below 0.1 eV and negligible drain current degradation. Excellent NBTI (ΔV_t = –8mV) supports robust operation beyond 10 years. MD and DFT simulations reveal that thermal stability stems from suppressed cation and anion out-diffusion at the channel–electrode interface.

9:55 AM

29-3 High Performance and Robust Oxide-semiconductor Channel Transistor DRAM with Multi-tiered Architecture

Hao Shi[1], Tongshang Su[1], Xiang Nie[1], Sijia Li[1], Xiangli Liu[1], Huan Liu[1], Peng Dai[1], Huihui Li[1], Jianqi Chen[2], Yuke Li[2], Xiangsheng Wang[2], Weihan Fan[2], Gangping Yan[2], Xueli Ma[2], Jinjuan Xiang[2], Hongbo Sun[2], Guilei Wang[2], Hao Meng[1],

Mingyuan Liu[1], Chao Zhao[2], Kanyu Cao[1]
[1]ChangXin Memory Technologies, Inc., [2]Beijing Superstring Academy of Memory Technology
We demonstrated the world's first backend-of-line (BEOL)-integrated, multi-tiered horizontal cell stacked DRAM architecture—featuring a IGZO FET—proving its feasibility through experimental validation on a 12-inch platform. The fabricated IGZO-based access transistor achieved an on-current of 60.9 μA/μm, a positive Vt=0.28V with SS=80 mV/dec. Additionally, the device demonstrated excellent reliability, a dynamic PBTI Vgmax of 1.66V ensuring a 10-year lifetime at 100℃. The projected TDDB Vgmax is 3.62 V for 10-years lifetime at 100℃. The breakthrough in this innovative architecture with addresses conventional scaling limitations while maintaining key electrical and reliability thresholds, offering a promising path for future DRAM innovation.

10:45 AM

29-4 Crystal Indium Oxide for 64-kbit 2T0C Memory with 5-ns Read Time, Monolithically Stacked on Si CMOS

Takanori Matsuzaki[1], Kazuma Furutani[1], Takanori Ito[1], Taichi Ebihara[1], Hiroki Inoue[1], Takeya Hirose[1], Haruki Katagiri[1], Chihiro Inagaki[1], Hiroki Komagata[1], Yoshinori Ando[1], Motomu Kurata[1], Tsutomu Murakawa[1], Fumito Isaka[1], Yuji Egi[1], Tatsuya Onuki[1], Yuto Yakubo[1], Toru Nakura[2], Shunpei Yamazaki[1]
[1]Semiconductor Energy Laboratory Co., Ltd., [2]Fukuoka University
A crystal indium oxide (IO) FET has an extremely low off-state current and a high on-state current, and is thus suitable for a low-power and high-speed device. The crystal IO FET is fabricated through BEOL process and accordingly shows promise as a memory device. We propose a 2T0C memory monolithically stacked on Si CMOS. The read-time of the 2T0C memory is shortened by canceling a voltage drop due to the capacitance coupling of a read word-line. The 64-kbit 2T0C memory including the crystal IO FET exhibits a read-time of 5ns, a write-time of 5ns, and a retention time exceeding 3,600s.

11:10 AM

29-5 Gain Cell Memory Scalability to 5-nm and Beyond

Shuhan Liu[1], Koustav Jana[1], Louis Liu[2], Jing Wang[2], Leo Jen-Chieh Liu[3], Elia Ambrosi[3], Mingyuan Song[3], Christian Kubicka[1], Yiming Tan[1], Haitong Li[4], Subhasish Mitra[1], Jian Chen[1], Thierry Tambe[1], Xinyu Bao[3], Meng-Fan Chang[3], John R. Hu[2], John Y. Chen[2], H.-S. Philip Wong[1]
[1]Stanford University, [2]Nvidia Corporation, [3]TSMC, [4]Purdue University
Memory architectures fundamentally comprise not only individual storage cells but complete systems integrating memory arrays with their essential peripheral circuits, which are equally critical for overall performance. This work examines gain cell memory array architecture and peripheral circuit design using fabricated N40 chip measurements and N5 node simulations. 3× density of HD SRAM was achieved through innovative folded and interleaved array layouts combined with 3D-stacked cell structures. The optimized 256×256 hybrid gain cell macro operates at 2 GHz array write/read frequency (95% of HD SRAM) with a cell size of 0.007 μm^2 at the 5-nm node. Importantly, we developed a comprehensive high-level array model for gain cell assessment and optimization, introducing the device metrics Qg×REFF that incorporates array-level effects to provide systematic guidance for gain cell device design.

11:35 AM

29-6 Demonstration of 3T0C Gain Cells with Self-Aligned Oxide Transistors for Parasitic-Aware Design at Array-level

Jay V Sonawane, Omkar S Phadke, Sunbin Deng, Faaiq G Waqar, Chengyang Zhang, Ming-Yen Lee, Shimeng Yu, Suman Datta
Georgia Institute of Technology
Capacitorless gain-cell (GC) memories realized with amorphous oxide semiconductor (AOS) devices promise cache-class speed with higher density. However, AOS transistors typically adopt gate-to-source/drain overlaps (OLs), introducing high parasitics that degrade speed, density, and energy efficiency at the array level. In this work, we performed a comparative study of parasitic effects in GCs using overlapped (OL) and self-aligned (SA) AOS transistors as testbeds. To eliminate WWL-SN capacitance coupling and other inherent issues in 2T0C configurations, we experimentally demonstrated 3T0C GCs with SA transistors, achieving negligible ΔV_{SN}, a 10× improvement in retention, and a 4.3× speed boost at iso-leakage compared to OL counterparts. Cycle-accurate

system analysis indicates that GPU caches using SA 3T0C GCs can reduce the energy-delay-area product by 75–80% while doubling capacity over SRAM within a single BEOL memory tier. These findings provide useful parasitic-aware design guidelines for GCs.

12:00 PM

29-7 Designing Oxide-Channel FeFETs for 1T Embedded Memory

Jung-Kyun Kim, Sijung Yoo, Seung-Geol Nam, Hyerim Hong, Kyooho Jung, Younjin Jang, Boram Kim, Sanghyun Jo, Na-Rae Han, Donghoon Kim, Kyung Mee Song, Jee-Eun Yang, Duk-Hyun Choe, Sangwook Kim, Jinseong Heo
Samsung Advanced Institute of Technology

We demonstrate IGZO-based MFMIS FeFETs optimized for monolithic 3D embedded memory. A novel structural parameter (AR2) and gate stack engineering with high-k IL and optimal-k FE improve erase performance and symmetry. Post-fabrication oxygen curing co-optimizes IGZO Vth and HZO Vc, suppressing oxygen vacancy-induced pinning and enhancing endurance. The resulting device achieves memory window >1.6 V, endurance >10^{12} cycles, and 20 ns operation under ±3 V pulses. This co-optimized structure offers scalable, low-power, and BEOL-compatible memory suitable for AI edge applications, bridging performance and integration requirements for future computing systems.

Session 30: MS | Advanced Memories and Ferroelectric Devices

9:00 AM, Continental 1-3

Co-Chairs: Benoit Sklénard, CEA-Leti and Pin Su, National Yang Ming Chiao Tung University

This session includes 6 papers that describe advances in the area of advanced memories and ferroelectric devices. The first paper, by Zhi Yao of Lawrence Berkeley National Lab (invited), introduces MicroEleX, a multiphysics simulation platform for emerging beyond CMOS devices. Magnetoelectric-magnonic and ferroelectric memories are used to demonstrate the simulation capabilities. The second paper, by Ha-Jun Sung from Samsung presents an Ab initio Informed Multi-scale Endurance modeling for Selector Only Memory devices. The third paper by Yichi Zhang from Tsinghua University reports a Monte Carlo ferroelectric switching model for AlScN ferroelectrics. The fourth paper, by Revanth Koduru from Purdue University, presents a 3D multi-grain phase-field framework for FeFETs and FeCap that captures the anisotropic structure of HfO2 and polycrystallinity. The fifth paper, by Prasanna Venkatesan from Georgia Tech, demonstrates a hybrid AI-accelerated methodology for endurance prediction in NVDRAM ferroelectric capacitors. The sixth paper, by Yi-Ming Tseng from National Taiwan University, reports a physics-aware Graph Neural Network (GNN) framework for FeFET variability analysis.

9:05 AM

30-1 Harnessing the Power of HPC and AI for Beyond-CMOS Microelectronics Device (Invited)

Zhi (Jackie) Yao[1], Rongqing Cong[2], Sajid Husain[2,3], Yingheng Tang[4], Thomas Lee[2], Andrew Nonaka[4], John Shalf[4], Sayeef Salahuddin[3], Lane Martin[2,4], Kaiyuan Yang[2], Ramamoorthy Ramesh[2,3,4]
[1]Applied Mathematics and Computational Research Division, Lawrence Berkeley National Lab, [2]Rice University, [3]UC Berkeley, [4]Lawrence Berkeley National Lab

Emerging beyond-CMOS devices demand multiphysics simulation that tightly couples electrostatics, quantum/charge-spin transport, and ferroic dynamics to predict performance and energy efficiency. Recent advances in GPU-enabled high performance computing and AI reduce full-3D, coupled simulations from weeks to hours, yet gaps remain in multiscale fidelity and seamless links to circuit, system, and materials design flows. We present MicroEleX, an open-source, GPU-accelerated framework that delivers near-linear scaling to thousands of GPUs, integrates AI surrogates, and provides plug-in physics modules to enable predictive, exascale co-design of beyond-CMOS technologies.

9:30 AM

30-2 Ab-initio Informed Multi-scale Endurance Modeling for Selector-Only Memory (AIMES)

Ha-Jun Sung[1], Hwasung Chae[2], Minwoo Choi[3], Jong-Bong Park[3], Bonwon Koo[3], Jabin Lee[2], Seoyeon Park[2], Semin Cheon[4], Yongsoo Yang[4], Youngjae Kang[3], Zhe Wu[2], Kiyeon Yang[3]
[1]Thin Film Technical Unit, Samsung Advanced Institute of Techonology, Samsung Electronics, [2]Semiconductor R&D

Center, Samsung Electronics, [3]Samsung Advanced Institute of Techonology, Samsung Electronics, [4]Korea Advanced Institute of Science and Technology (KAIST)
Ab-initio Informed Multi-scale Endurance Modeling for SOM (AIMES) is introduced to model endurance degradation in SOM devices. This framework combines electro-thermal model, atomistic simulations, drift-diffusion analysis, and TCAD simulations to connect atomistic processes with device-level V_{TH} shifts. According to TEM and AIMES analysis, the V_{TH} evolution mechanism consists of three stages: initial current-path relaxation, formation of band-gap gradients caused by migration, and late-stage degradation driven by out-diffusion. The multi-scale model reproduces measured V_{TH} and memory window cycling behavior, validating its accuracy. Furthermore, dopant engineering guided by AIMES significantly enhances endurance (suppressing migration and out-diffusion), thereby improving device lifetime and stability.

9:55 AM
30-3 First-Principles-Augmented KAI Model Bridging Nucleation to Domain-Wall in Ferroelectric AlScN
Yichi Zhang[1], Ye-Bin Dai[2], Haiming Qin[2,3], Hang Liu[4], Zijian Zhou[5], Suman Xia[2], Hongjing Lai[5], Heng Xiang[2], Chen Shen[6], Bowen Su[4], Dapeng Huang[4], Binbin Wu[2], Yeqing Zhu[4], Yujie Liu[4], Shufei Gu[4], Xueyan Zhang[4], Hao Zhang[2], Xinpeng Wang[2], Yu-Tao Li[4], Yi Tong[2], Kan-Hao Xue[5], Xiangshui Miao[5], Yi Yang[4], Tian-Ling Ren[4]
[1]School of Integrated Circuits, Tsinghua University, [2]Suzhou Laboratory, [3]University of Posts and Telecommunications, [4]Tsinghua University, [5]Huazhong University of Science and Technology, [6]PFTN Semicond Co. Ltd
A novel ferroelectric (FE) switching framework for wurtzite AlScN is proposed, fundamentally decoupling polarization reversal into two distinct physical stages: Sc-site-preferential nucleation succeeded by electric-field-driven domain wall (DW) propagation. Validated by μs-pulse electrical measurements, the first-principles-driven model quantifies the transition from nucleation-dominated to DW-dominated kinetics. The first phenomenological Monte Carlo framework for AlScN ferroelectrics is presented, experimentally verifying Sc-induced nucleation acceleration and enabling arbitrary-field switching waveform simulations. The advancement establishes the foundation for ultra-high-speed FeRAM and neuromorphic computing in AlScN.

10:45 AM
30-4 Impact of Anisotropic Crystal Structure on HfO_2 based Ferroelectric FETs and Capacitors: A 3D Multi-Grain Sublattice-Phase-field Model based Analysis
Revanth Koduru[1], Sumeet Kumar Gupta[2]
[1]Electrical and Computer Engineering, Purdue University, [2]Purdue University
We present a 3D phase-field simulation framework for doped hafnium oxide ferroelectric field effect transistors and capacitors. Our framework captures the structural anisotropy and polycrystallinity of HfO_2 by integrating the grain growth, sublattice-based TDGL equation, Poissons's and semiconductor drift-diffusion equations. We model the anisotropic crystal structure of HfO_2, characterized by continuous polar layer and alternate polar-spacer layers in orthogonal in-plane directions. We investigate its impact on multi-domain polarization configurations and polarization switching. Simulations demonstrate unit-cell-wide domains along APLS, consistent with first principles studies. We show the direction-dependent domain growth and nucleation mechanisms, leading to layer-wise P-switching along APSL. In FEFETs, we analyze the impact of APSL/CPL orientation relative to the channel length showing its effect on polarization switching and memory window.

11:10 AM
30-5 VLPred: Accelerated Endurance Prediction Platform Enabled by AI Surrogates
Prasanna Venkatesan Ravindran[1], Congrui Li[1], Andrea Padovani[2], Zekai Wang[1], Hari Jayasankar[1], Priyankka Ravikumar[1], Mohammad Nabian[3], Kaustubh Tangsali[3], Sheel Nidhan[3], Janaki Vamaraju[3], Ram Cherukuri[3], Prakhar Mathur[3], Enrico Piccinini[4], Riccardo Foiera[4], Roman Gafiteanu[4], Khandekar Akif Aabrar[1], Suman Datta[1], Surya Kalidindi[1], Luca Larcher[4], Asif Khan[1], Yiyi Wang[3], Gaurav Thareja[4]
[1]Georgia Institute of Technology, [2]University of Modena and Reggio Emilia, [3]NVIDIA, [4]Applied Materials
We present a fully integrated AI-driven framework for rapid endurance prediction in NVDRAM ferroelectric capacitors. Endurance testing is one of the most time- and resource-intensive steps in memory characterization. To overcome the scarcity of endurance training data, we propose an experimentally calibrated synthetic data

generation pipeline using kinetic Monte Carlo (kMC) simulations in Ginestra™. We train a transformer-based AI surrogate using this high-fidelity dataset, achieving an R^2 of 0.992 and enabling ~10^5x speedup in defect evolution prediction. The surrogate generates large-scale synthetic datasets by sampling initial defect profiles, which are then used to train a hybrid multi-layer perceptron (MLP)-attention model that maps early-life defect characteristics to Weibull endurance distributions. This final endurance prediction model achieves R^2 values of >0.98 for η and ~0.9 for β, demonstrating its reliability in capturing endurance distribution characteristics. Wafer-scale prediction of breakdown distributions is demonstrated in one-shot, reducing characterization time by over 10-orders of magnitude.

11:35 AM

30-6 Physics-Aware Graph Neural Network for FeFET Variability Analysis: Enabling 10^7× Speedup, R^2 > 0.98 Accuracy, and Efficient Transfer Learning to New Architectures

Yi-Ming Tseng[1], Wei-Tung Chen[2], Jia-Tai Wu[2], Cheng-Hung Wu[2], Vita Pi-Ho Hu[2,2]
[1]Graduate Institute of Electronics Engineering, National Taiwan University, [2]National Taiwan University

We propose and experimentally validate a Physics-aware Graph Neural Network (GNN) for FeFET variability analysis, addressing the limitations of traditional TCAD and DNN approaches. By representing the ferroelectric (FE) layer as a graph and incorporating multi-level physical features, our method achieves a 10^7× speedup compared to TCAD while maintaining high predictive accuracy (R^2 > 0.98). We further demonstrate superior data efficiency and adaptability over conventional DNNs through transfer learning, successfully extending the model to an UTBSOI architecture with only 500 samples (R^2 > 0.95) and, critically, from simulation to experimental data with 200 samples (V_{th} R^2 > 0.985). This work presents a robust methodology that bridges the simulation-to-experiment gap, enabling rapid and reliable variability analysis for advanced FeFET technologies.

Session 31: ALT/EDT/NC | Focus Session | Beyond Von Neumann, from Physics-inspired Solvers to Quantum Hardware Implementation

9:00 AM, Continental 4

Co-Chairs: Maud Vinet, Quobly and Changhwan Shin, Korea University

This session explores revolutionary computing paradigms that transcend traditional von Neumann architectures by harnessing physical phenomena and quantum mechanics for computational problem-solving. From CMOS-compatible physics-inspired optimization solvers to scalable quantum computing systems with cryogenic electronics, these emerging technologies demonstrate practical pathways toward solving computationally intractable problems while addressing critical engineering challenges in hardware implementation and system integration.

9:05 AM

31-1 Parallel Probabilistic Ising Architectures for Next Generation Computing (Invited)

Saavan Patel[1,2], Pratik Brahma[3], Brian Mao[2], Kevin Jordan[2], Sebastien Jomphe[2], Chirag Garg[3], Sayeef Salahuddin[3]
[1]Electrical Engineering and Computer Sciences, University of California, Berkeley, [2]InfinityQ Technology Inc., [3]University of California, Berkeley

Probabilistic computing offers a physics-inspired approach to solving problems that are intractable for conventional von Neumann architectures. These problems span diverse domains, including quantum simulation, neurobiology modeling, and combinatorial optimization. This paper presents two complementary implementations: (i) PASS, a mixed-signal asynchronous ASIC that enables true probabilistic computation using stochastic neurons, and (ii) TitanQ, a digital simulation platform for large-scale, parallel probabilistic computing. We demonstrate practical applications in statistical physics, finance, and energy systems, and show that Ising-based probabilistic approaches can achieve orders-of-magnitude performance improvements over traditional methods.

9:30 AM

31-2 CMOS Physics-Inspired Architectures and Circuits for Combinatorial Optimization Solvers (Invited)

Michael Flynn[1], Evangelos Dikopoulos[2], Luke Wormald[2,2], Ying-Tuan Hsu[2], Benjamin Datsko[2], Wei Tang[2], Zhengya Zhang[2]

[1]Electrical Engineering and Computer Science (EECS), University of Michigan, [2]University of Michigan

Recently, physics-inspired computing platforms have shown significant potential in accelerating the solution of combinatorial optimization problems (COPs) by leveraging continuous-time (CT) operation, massive parallelism, and direct mapping to coupled CMOS-based spins, often realized with oscillators. These systems leverage the natural dynamics of coupled oscillators to rapidly converge to COP solutions, resulting in substantial improvements in solution time and energy. This work introduces physics-inspired computing, outlines its advantages, and presents a framework for mapping complex COPs to coupled relaxation-oscillator spins. We demonstrate this approach with two custom solver systems, a direct 3-SAT engine and an LDPC decoder.

9:55 AM

31-3 Physical Computing with Coupled Oscillators - Current Progress and Future Directions (Invited)

Aida Todri-Sanial[1], Jelle N.G.W. Verest[2], Filip Sabo[2], Hyunwook Kim[2], Bram F. Haverkort[2], Thomas C.V. van Bodegraven[2], Simanta Lahkar[2], Aishani Mazumder[2], Wooseok Choi[3], Donato F. Falcone[3], Siegfried F. Karg[3], Bert J. Offrein[3], Valeria Bragaglia[3]

[1]NanoComputing Research Lab, Electrical Engineering Eindhoven University of Technology, [2]Eindhoven University of Technology, [3]IBM Research

Dynamical systems exhibit rich and intricate behaviors that can be exploited for physical computation. Physical computing draws inspiration from complex systems that continuously adapt, self-organize, and minimize energy to reach stable configurations, inherently enabling parallel processing. These features show promise in solving intractable scientific problems, such as NP-hard combinatorial optimization problems. Yet, engineering such dynamical systems for computing remains a challenge in terms of technology choice and scalable circuit design implementation. This paper presents an overview of recent advances and future directions in physical computing using coupled oscillatory neural networks (ONNs).

10:45 AM

31-4 High-fidelity Single-electron Shuttling in Industrially Fabricated Spin Qubit Devices (Invited)

Wolfram Langheinrich[1], Pascal Muster[1,2], Till Huckemann[2,3], Sebastian Pregl[1], Varvara Brackmann[4], Michael Friedrich[4], Felix Reichmann[5], Nikola David Komericki[6], Lars R. Schreiber[2,3], Hendrik Bluhm[2,3]

[1]Infineon Technologies Dresden GmbH & Co. KG, [2]RWTH Aachen University, [3]ARQUE Systems GmbH, [4]Fraunhofer Institute for Photonic Microsystems, [5]IHP - Leibniz Institute for High Performance Microelectronics, [6]Fraunhofer Institute for Applied Solid State Physics IAF

We have fabricated single-electron shuttling channels in Si/SiGe heterostructures within an industrial 200 mm production line as building block for a scalable quantum computing architecture with spin qubits. Shuttling is achieved via an adiabatically moved quantum dot which is formed by a series of only four individual, sinusoidal voltage signals, applied to a dense grid of gate electrodes, enhancing scalability. The gate module comprises a triple metal gate process, where one gate layer uses electron-beam lithography to achieve a pitch below 100 nm in the complex grid architecture. With the device presented in this study, we have been able to achieve an electron shuttling success rate higher than 99.9 % across a distance of 4.5 µm.

11:10 AM

31-5 Scalable Quantum Computing System with Integrated Millikelvin Cryogenic Electronics (Invited)

Shu-Jen Han

SEEQC

Current superconducting quantum computers rely on a brute-force scaling approach where quantum chips operating at the millikelvin stage inside dilution refrigerators are connected to room-temperature electronics via an ever-increasing number of signal wires. In order to drastically scale up such systems without facing refrigeration challenges, these quantum interface electronics need to be integrated at cryogenic temperatures, ideally at the same stage as qubits. In this paper, we show that, by employing energy-efficient digital superconducting

microchips based on single flux quantum (SFQ) technology, all essential functions to operate a full stack quantum computer can be integrated at 10 mK. They include digital qubit charge control and magnetic flux control for single- and two-qubit gates, as well as digital qubit readout.

11:35 AM

31-6 Using CMOS-compatible Optical Modulators to Solve I/O Challenges in Trapped-ion Quantum Computing (Invited)

Daniel Stick[1], Daniel Dominguez[1], Matthew Eichenfield[1,2], Michael Gehl[1], Craig Hogle[1], Andrew Leenheer[1], Hayden McGuinness[1]
[1]Sandia National Laboratories, [2]University of Colorado, Boulder

Most trapped-ion quantum computers confine ions using precise voltages applied to electrodes and manipulate their internal states and execute quantum logic operations using focused laser beams. The prevailing approach to scaling these systems from current sizes that use tens of ions to ones with several orders of magnitude more ions relies on the lithographic fabrication of larger electrode arrays as well as on-chip integrated waveguides for delivering light to each individual ion. This approach solves the crosstalk problem that is inherent to free-space light delivery, but it does not address the input/output problem associated with coupling many optical signals to the on-chip waveguides. Here we describe an approach that uses CMOS-compatible optical modulators that can substantially address this challenge, and analyze its architectural implications.

Session 32: RSD | Compute-in-Memory (CiM) Reliability

9:00 AM, Continental 5

Co-Chairs: Andrea Padovani, University of Modena and Reggio Emilia and Mario Lanza, National University of Singapore

This session includes two invited papers presenting techniques for non-destructive imaging of microchips and screening for manufacturing defects, alongside six contributions highlighting new techniques for enhancing the security and reliability of compute-in-memory (CiM) architectures. David Lerner (Intel, invited), underlines the importance of using a large set of diverse functional tests to screen defects that manifest silent data errors in data-center processors. Zhu (Peking University) demonstrates a BEOL-compatible FeFET-based multi-bit analog CiM achieving 93.8% inference accuracy, 10^{10} endurance cycles, and over 10 years of retention. Zhou (HKUST, Guangzhou), presents a new secure and energy-efficient CiM architecture that leverages the intrinsic nonlinearity of ferroelectric diodes. Lee (Seoul National University) introduces a novel MTJ-based physically unclonable function with error-free operation and concealability. In a second invited talk, Tomas Aidukas (Paul Scherrer Institute) demonstrates the capabilities of ptychographic X-ray computed tomography for 3D microchip imaging at 4 nm resolution. The final three papers explore ReRAM-based CiM solutions for security. Wang from the University of Hong Kong integrates computation, storage, and security primitives on a single chip. The paper by Gao from Peking University proposes a 1T1R RRAM-based CiM accelerator to mitigate security risks in edge applications. The last paper by Fu from Huazhong University presents a searchable encryption system relying on an adaptive modulation of the device's intrinsic randomness.

9:05 AM

32-1 Screening for Manufacturing Defects that Manifest as Silent Data Errors in Data Center Processors (Invited)

David Lerner, Ethan Hansen, Thiago Macieira, Sridharan Sakthivelu, Carlos Carrizo, Tzu-Lin Sun, Brad Kelly
Intel

Multiple industry studies have shown that Silent Data Errors (SDE) or other forms of Silent Data Corruption (SDC) caused by manufacturing defects may occur on System-On-Chip (SoC) devices deployed at-scale in Data Centers. While structural test techniques such as scan are an essential component of manufacturing tests, system-based functional testing is required to deliver SoCs that meet the quality demanded by cloud Data Centers. This paper will share results obtained from multiple generations of server SoCs that illustrate the importance of using a large set of diverse functional tests, such as those included in the Intel® Data Center Diagnostic Tool (DCDiag) test suite[1], to screen defects that manifest as SDE. The need to characterize new designs to create bespoke test

recipes will be discussed. Data that illustrates the limitations of relying on traditional test coverage metrics and architectural simulation to identify effective functional tests will also be reviewed.

9:30 AM

32-2 A BEOL FeFET based Multi-bit ACiM Macro with High Accuracy and Throughput via Device-Array-System Co-Optimization for Edge LM

Runteng Zhu[1], Zhiyuan Ning[2], Hanyong Shao[2], Guangnan Dai[3], Xiaojian Xu[2], Wenshuai Yao[2], Yuejia Zhou[2], Weiqin Huang[2], Mingxiang Yu[2], Chuanlin Sun[2], Wenpu Luo[2], Siyuan Qi[2], Jinghao Yang[2], Zhiqi Liu[2], Junchen Dong[4], Yihan Zhang[3], Kechao Tang[2,5], Ru Huang[2,5]
[1]School of Integrated Circuits, Peking University, [2]Peking University, [3]Hong Kong University of Science and Technology, [4]Beijing Information Science and Technology University, [5]Beijing Advanced Innovation Center for Integrated Circuits

Analog Compute-in-Memory (ACiM) promises efficient multiplication acceleration for edge large models (LMs), but has long suffered from poor reliability. In this work, we demonstrate a high-throughput, high-accuracy multi-bit ACiM macro based on 3D stackable FeFETs via device-array-system co-optimization. At the device level, based on reliability enhanced 1FeFET-1T (1F-1T) cells, a novel drain-input multi-level (DIML) scheme is implemented for accurate weight readout with minimal latency. At the array level, design of 1F-1T array topology is combined with a Bidirectional Successive Programming scheme, effectively improving disturb immunity by 7 times. At the system level, we introduce deep variation-aware training (DVAT) to realize a fault-tolerant ViT model. Implemented on reliable 2-bit FeFETs with 10^{10} endurance and >10 years retention, the ACiM macro achieved a peak throughput of 3.824 TOPS while maintaining a high inference accuracy of 93.82% in image classification, showing immense potential for the acceleration of edge LMs.

9:55 AM

32-3 Nonlinear Ferroelectric Diode Offers High Security in Accurate and Energy-Efficient Compute-in-Memory

Wenjuan Zhou[1], Renhao Xue[1], Quanyi Tu[1], Zenghui Chang[1,2], Ruizhan Yan[1], Mansun Chan[2], Xiwen Liu[1]
[1]The Hong Kong University of Science and Technology (Guangzhou), [2]The Hong Kong University of Science and Technology

Neural network models implemented with compute-in-memory (CIM) architectures are vulnerable to readout attacks and reverse engineering. Existing encryption methods based on linear devices typically involve complex circuits, poor scalability, and high power consumption. In this work, we propose a secure CIM architecture leveraging the intrinsic nonlinearity of ferroelectric diodes (FeDs) to enable encryption directly at the device level. A transistor-free, multilayer encryption scheme is embedded within the FeD crossbar, incorporating XOR-based input–weight co-encryption, unclonable input obfuscation, and β-sensitive linearization. These mechanisms ensure that only correct keys yield accurate matrix-vector multiplication (MVM), while incorrect keys lead to masked and distorted outputs. Physical measurements on a 1.6 kb FeD crosspoint array show that inference accuracy drops from 86.33% to 18.04% after encryption, while decryption successfully restores it without any loss. This device-level approach is compatible with other nonlinear memories, offering a scalable and energy-efficient hardware security solution for CIM.

10:20 AM

32-4 Concealable and Bit Error-Free Breakdown-Based Physical Unclonable Functions Using Magnetic Tunnel Junctions

Kyung Min Lee[1,2], Ryun-Han Koo[3], Munhyeon Kim[3], Joon Hwang[3], Jae-Joon Kim[3], Jong-Ho Lee[3]
[1]Department of Electrical and Computer Engineering(ECE), Seoul National University, [2]R&D center, Samsung Electronics, [3]Seoul National University

This work is the first to demonstrate a concealable breakdown-based PUF using MTJs using a fabricated MRAM macro. PUF responses are formed by pulse-controlled breakdown, where breakdown and non-breakdown cells represent '1' and '0'. By leveraging the memory state of non-breakdown cells, the response can be concealed. The responses exhibit near-ideal randomness metrics and maintain bit-error-free readout under 10k repeated accesses at 125°C. Device-level characterization was performed to understand breakdown-related behaviors. A

strong PUF architecture using these responses is also proposed, with ML- and rank-based attack resistance. The design is compatible with commercial MRAM/OTP platforms for secure hardware integration.

11:10 AM

32-5 High-resolution, Data-driven 3D X-ray Imaging of Microchips using Ptychography (Invited)

Tomas Aidukas[1], Nicholas William Phillips[2,3], Ana Diaz[3], Emiliya Poghosyan[3], Elisabeth Mueller[3], Anthony Levi[4], Gabriel Aeppli[3,5,5,6], Manuel Guizar-Sicairos[3,6], Mirko Holler[3]
[1]PSI Center for Photon Science, Paul Scherrer Institute, [2]Commonwealth Scientific and Industrial Research Organisation, [3]Paul Scherrer Institute, [4]University of Southern California, [5]Eidgenössische Technische Hochschule Zürich (ETH Zürich), [6]École Polytechnique Fédérale de Lausanne (EPFL)

As transistor dimensions have shrunk over the years, X-ray microscopy resolution has also improved to meet the demands of the semiconductor industry. Although electron microscopy can achieve higher resolution, X-rays offer unique advantages, including non-destructive 3D imaging of fully intact integrated circuit dies and the future possibility of imaging under operando conditions. In this work**,** we demonstrate the capabilities of ptychographic X-ray computed tomography for 3D microchip imaging at 4 nm resolution. To achieve such performance, we introduce newly developed methods to overcome nm-scale experimental instabilities and depth-of-field limitations. Microchip-specific high-throughput data collection methods will also be introduced.

11:35 AM

32-6 A Monolithic Reconfigurable RRAM CIM Array Integrating PUF, TRNG, and a Lightweight Block Cipher for Secure Edge AI

Songqi Wang[1], Shuang Liang[2], Shaonan Wu[3], Zhiqi Yang[3], Jichang Yang[2], Xinyuan Zhang[2,4,5], Yi Li[6], Aoyang Zhang[3], Zhongrui Wang[4,7], Yuhao Zhang[2], Han Wang[2]
[1]Electrical and Electronic Engineering, The University of Hong Kong, [2]The University of Hong Kong, [3]Tsinghua University, [4]Southern University of Science and Technology, [5]Access, [6]the Univeristy Of Hong Kong, [7]Access

AI edge devices urgently need onchip security that is both area and energy efficient, yet current designs still separate security, compute, and storage blocks, inflating silicon area and power consumption. We designed SCIM, an RRAM-based secure compute-in-memory chip that fully integrates the cryptographic security algorithm suite (cipher module, PUF, and TRNG), leveraging the energy-efficient reconfigurable RRAM array. SCIM consists of five functions: (i) ISO/IEC-compliant PRESENT cipher, (ii) reconfigurable r-PUF, (iii) a TRNG with entropy throughput of 357 Mbps, (iv) OTP-style encryption/decryption, and (v) vector–matrix multiplication. Consolidating these primitives in array shrinks area by 69.1%. The r-PUF achieves a Hamming distance of 49.87% between chips with a concealment error of 20.12%. In-array PRESENT reduces latency by 25.9% and energy by 92.0% versus conventional designs. Moreover, XOR encryption of CIFAR-100 classification heads cuts accuracy by 78.23% and 82.04%. Collectively, SCIM delivers a low-area, multifunctional, and energy-efficient hardware root of trust for edge AI.

12:00 PM

32-7 First Demonstration of In-MVM XOR-Decryption based on 40nm RRAM Chip for Secure Point Cloud Processing

Yi Gao[1], Zongwei Wang[2,3], Ruiqing Xie[2], Yuhang Yang[2], Lin Bao[2], Ling Liang[2], Zheng Zhou[2], Yimao Cai[2]
[1]School of Integrated Circuits, Peking University, [2]Peking University, [3]YanXin MicroElectronics Co., Ltd. (YXME)

In this work, we propose the first 1T1R RRAM-based computing-in-memory (CIM) accelerator that integrates in-situ bit-wise XOR decryption into matrix-vector-multiplication to address the secure risks in edge-based point cloud processing (e.g., recognition). With the proposed hybrid encryption (ENC) scheme, our design effectively protects input point clouds, rotation keys, and PNN weights during on-chip inference. To enhance performance and functionality, we further introduce an MLC-extension bit-wise XOR mechanism for high-efficiency MVM and an in-situ transpose matrix decryption scheme for data recovery. With a fully pipelined design across point-wise encryption and recognition, the proposed secure point cloud system achieved significant improvement in speed (1.71×), energy efficiency (267×), and robust security, demonstrating great potential in edge-based secure point cloud applications.

12:25 PM

32-8 First Demonstration of ReRAM-based In-Memory Searchable Encryption System with Randomness-Controllable Programming

Yuyang Fu[1], Yingjie Yu[2], Jiancong Li[2,3], Jia Chen[4], Zhengzhe Li[2], Junjie Jin[2], Yibai Xue[2], Ao Xu[5], Yinan Wang[5], Zhongrui Wang[6], Yi Li[2], Xiangshui Miao[2]
[1]School of Integrated Circuits, Huazhong University of Science and Technology, [2]Huazhong University of Science and Technology, [3]the Hong Kong University of Science and Technology, [4]AI Chip Center for Emerging Smart Systems, [5]National University of Defense Technology, [6]Southern University of Science and Technology

We present the first ReRAM-based Searchable Encryption (SE) system with a randomness-controllable programming method to address the conflicting randomness requirements of encryption and search. Leveraging Computing-In-Memory (CIM) architecture, our design supports reconfigurable processing with high parallelism and energy efficiency. The one-step in-memory CVP encryption achieved 25× lower latency and 5,079× higher energy efficiency than GPU, while the Euclidean distance-based in-memory search achieved 5× speedup and 20,961× improvement. Evaluated on practical datasets, our system demonstrates a promising path toward secure and efficient post-quantum search architectures.

Session 33: ALT | BEOL-Compatible Devices and Interconnect Technology

9:00 AM, Continental 6
Co-Chairs: Martin O'Toole, ASML and Amy Novak, Advanced Micro Devices
Session of eight papers highlighting major advances in BEOL-compatible devices, interconnects, and power delivery systems. Showcasing BEOL compatible devices - switched capacitors, MIMs and transistors. Advancements in grain structure and integration to enable ruthenium interconnect to replace copper. Chip to chip interconnect technologies for 3D SoCs and CNT TSV usage for thermal management.

9:05 AM

33-1 High-Efficiency BEOL-Compatible On-Chip Switched Capacitor DC-DC Converter Using Planar and Vertical Oxide Transistors with Trench Capacitors

Hung-Chun Chou, Sunbin Deng, Chengyang Zhang, Jungyoun Kwak, Arun Dandapani Asokan, Emmanuel Quezada, Eknath Sarkar, Junmo Lee, Jaewon Shin, Dyutimoy Chakraborty, Visvesh S Sathe, Shimeng Yu, Suman Datta
Georgia Institute of Technology

This work presents a breakthrough in on-chip point-of-load (PoL) DC-DC voltage conversion for heterogeneous 3D integrated circuits by leveraging BEOL-compatible transistor and trench capacitor architectures. This work presents indium-tin-oxide (ITO) planar transistors and vertical indium-gallium-oxide (IGO) transistors, achieving a specific on-resistance of 87 and 0.04 $m\Omega \cdot mm^2$ for 12V and 3V voltage ratings, respectively. 3D trench $Hf_{0.25}Zr_{0.75}O_2$ flying capacitors with 97.9 $fF/\mu m^2$ capacitance density are demonstrated. By monolithically integrating enhancement- and depletion-mode ITO planar transistors, a 2:1 switched-capacitor DC-DC converters are experimentally demonstrated with ≥87.5% efficiency per stage. Circuit simulations using IGO vertical transistors and 3D trench $Hf_{0.25}Zr_{0.75}O_2$ capacitors exhibit 94.3% peak stage efficiency for 2:1 step-down conversion from 3 V, offering competitive BEOL-compatible active and passive device technologies for “active” backside power delivery.

9:30 AM

33-2 Monolithic BEOL-Compatible Multi-Layer High-Density MIM Capacitor for Power Integrity Optimization and Vertical Power Delivery in HPC 3D ICs

Peng An[1], Guoshan Yang[2], Guoxing Chai[2], Zhoutong Liu[2], Jiawei Cui[3], Sihang Liu[3], Jin Wei[3], Bo Shen[3], Yan Wang[2], Xiaosen Liu[2]
[1]School of Integrated Circuits, Tsinghua University, [2]Tsinghua University, [3]Peking University

This work presents a monolithic 3D-stacked high-density MIM (3DSMIM) capacitor for power integrity (PI) enhancement in high performance computing (HPC). TiN/HfO_2 stacking, combined with sidewall oxidation and reflow-based profile tuning, reduces ESR and improves reliability. A DTCO-guided co-design framework optimizes structure and process based on workload demands. The eight-layer 3DSMIM achieves 125.9 nF/mm^2 capacitance and supports scalable integration across on-die, RDL, and package substrates. The high-density 3DSMIM

capacitor with full BEOL and VPD compatibility is proposed to suppress the 1st droop in M3D computing workloads, enabling a fundamental reduction in VDD guardband and power consumption, ultimately improving IPC within the same thermal design power (TDP) budget for kW-class HPC processors.

9:55 AM

33-3 BEOL-Compatible ITO/ZnO Transistors with Record-Small Transition Region Width of 80 mV and Ultra-High Mobility of 150 $cm^2V^{-1}s^{-1}$

Chuanlin Sun[1,2], Tingchen Yi[3], Xinru Shi[4], Ruibin Duan[3], Junchen Dong[4], Zongwei Wang[3], Zheng Zhou[3], Yudi Zhao[4], Kai Zhao[4], Dedong Han[3], Yimao Cai[3], Xing Zhang[3]
[1]School of Information & Communication Engineering, Beijing Information Science and Technology University, [2]School of Integrated Circuits, Peking University, [3]Peking University, [4]Beijing Information Science and Technology University

In this work, we propose a novel post-treatment method named the compensation technique (CT) to reduce the transition region width (V_{TR}), which defines how much the operating voltage of oxide transistors can be scaled down. By applying CT, we achieve a record-small V_{TR} of 80 mV in the ITO/ZnO thin-film transistors, along with an ultra-high field-effect mobility (μ_{FE}) of 150 $cm^2V^{-1}s^{-1}$ and marked improvements in on-current and gate control. Through device modeling and material characterization, we found that CT treatment can effectively suppress oxygen vacancies in the ITO/ZnO active layer, thereby decreasing tail state traps that are the primary factor determining the V_{TR}. Meanwhile, CT also introduces additional hydrogen, which contributes to the enhanced μ_{FE}. Circuit-level simulations verify that reducing V_{TR} enables low-power operation. Our work paves the way for low-power, high-density 3D integration of oxide transistors in future electronics.

10:20 AM

33-4 High Performance 2.5V N&P 400°C SOI MOSFETs: A Breakthrough for Versatile 3D Sequential Integration

Daphnée Bosch, William Vandendaele, Elhadji Diallo, Anne-Sophie Royet, Antoine Albouy, Mathieu Opprecht, François-Xavier Darras, Véronique Benevent, Paul Maillard, Obaid-allah Adami, Mickael Ribotta, Sebastien Kerdiles, Remy Gassilloud, Xavier Garros, Perrine Batude
CEA-Leti

For the first time, n & p 2.5V SOI MOSFETs have been successfully fabricated for versatile sequential integration (thermal budget ≤400°C)**.** Devices exhibit excellent electrical performance**,** close to standard 2.5V SOI MOSFET. For L=0.2µm: D_{it} <5.10^{10} $cm^2.eV^{-1}$, μ_{0p} / μ_{0n} =190 /455 $cm^2.V^{-1}.s^{-1}$, I_{ON-p} /I_{ON-n}= 325 /490 $\mu A.\mu m^{-1}$ I_{OFF-P}/I_{OFF-n}= 60/20 $pA.\mu m^{-1}$ (R_{access} <420 Ω.µm). Thanks to NLA gate stack curing, low-frequency noise trap densities as low as N_{t-p} / N_{t-n} =2.10^{17} / 3.10^{16} $eV^{-1}.cm^{-3}$ are achieved along with a 2 years pMOS lifetime and >5 years nMOS lifetime, overcoming main low temperature Si CMOS devices showstopper.

11:10 AM

33-5 Towards SoCs Leveraging 200nm Wafer-to-Wafer Interconnection Pitch

Joeri De Vos, James Myers, Stefaan Van Huylenbroeck, Liesbeth Witters, Emmanuel Chery, Michele Stucchi, Peng Zhao, Koen D'Have, Maxim Vanrusselt, Serena Iacovo, Hung-Chieh Tsai, Sven Dewilde, Lieve Bogaerts, Cindy Heyvaert, Soon Aik Chew, Boyao Zhang, Shuo Kang, Moritz Brunion, Dwaipayan Biswas, Gerald Beyer, Geert Van der Plas, Zsolt Tokei, Eric Beyne
imec

This paper highlights the advantages of 3D System-on-Chip and shows that STCO is required for system-level benefits. A physical design study was performed evaluating the feasibility when multi-threading CPUs, showing that 250nm 3D pitches enable this performance benefit without area penalties. Two high density 3D technology enablers are described. For wafer-to-wafer Cu-SiCN hybrid bonding technology the challenges are outlined to enable scaling down the interconnect pitch from 300nm down to 200nm. Nano Through-Silicon Via technology enables fine-pitch 210nm and 120nm vertical connections between the frontside and backside of a wafer through extreme wafer thinning and overlay control in backside lithography.

11:35 AM

33-6 Grain-Orientation-Engineering of Atomic Layer Deposited Ruthenium Interconnect Technology

Youngchul Leem, Eun-Hyoung Cho, Young Min Lee, Taewon Jeong, Yujin Han, Jae Hong Lee, Dongmin Kim, Jeong Yub Lee, Keun Wook Shin, Kyung-Eun Byun, Gi-Young Cho
Samsung Advanced Institute of Technology

Ru is being actively investigated as a promising candidate to replace damascene Cu for interconnects at sub-10 nm technology nodes. We introduce ALD-based strategies for tailoring orientation of Ru grains in both metal lines and vias. Utilizing ALD, we successfully deposited a Ru film with over 99% texture quality on an amorphous underlayer, leading to metal lines with significantly improved electrical resistance. Furthermore, we demonstrate bottom-up via filling using supercycle-based area-selective deposition, enabling the growth of c-axis-oriented single-crystal Ru vias. Finally, our ALD process exhibits excellent conformality even in highly complex 3D structures, underscoring its suitability for next-generation interconnect integration.

12:00 PM

33-7 Demonstration of 16nm Pitch Two Metal Level Fully Self-Aligned Via Ru Interconnects

Gilles Delie, Stefan Decoster, Yannick Hermans, Giulio Marti, Vincent Renaud, Evi Vrancken, Bart Kenens, Bart De Wachter, Gayle Murdoch, Chen Wu, Seongho Park, Zsolt Tokei
Imec

We demonstrate for the first time a two metal level Ru interconnect using fully self-aligned vias (FSAV) landing on 16 nm metal pitch (MP). M_x lines are patterned using EUV-SADP and direct metal etch of Ru in a spacer is dielectric (SID) integration scheme. With a critical dimension (CD) of 8 nm M_x lines measure an average of 670 Ω/μm with a yield of >95% across a 300 mm wafer and average resistivity of 15.5 μΩ.cm extracted from TCR measurements. FSAV via performance is competitive with a measured resistance value of 14 Ω for >95% yield across the full wafer.

12:25 PM

33-8 An Electrical Thermal Co-transmission (ETC) TSV Interconnect with Annular Cu Conductor and CNT Core for Improving the Thermal Management of Heterogeneous Chiplet Integration

Ziyue Zhang[1], Han Wang[2], Yigang Hao[2], Yuwen Su[2], Xuyan Chen[2], Jiaxuan Zhang[2], Yingtao Ding[2], Zhiming Chen[2]
[1]School of Integrated Circuits and Electronics, Beijing Institute of Technology, [2]Beijing Institute of Technology

Thermal management presents a significant challenge for high-density integrated systems. Herein, we demonstrate a novel electrical thermal co-transmission (ETC) TSV structure featuring an annular Cu conductor and central CNT core, which not only serves as a vertical electrical interconnect but also provides enhanced heat dissipation. The ETC TSVs are fabricated through a simplified process flow based on double-sided processes, exhibiting excellent electrical performance characterized by low and stable parasitic capacitance, low leakage current, and low resistance. Furthermore, thermal measurements verify the ETC TSVs' capability to enhance heat dissipation, validating their effectiveness in improving the thermal management of high-density integrated systems.

Session 34: SMB | Integrated Photo Sensors and Applications

9:00 AM, Continental 7-9

Co-Chairs: Eng-Huat Toh, GlobalFoundries and Chiung Lo, xMEMS Lab

This session includes seven papers dealing with integrated photo sensors and applications. The first paper by S. Kim et al proposes a CMOS-compatible neuromorphic photoreceptor that monolithically integrated an amorphous silicon photodiode. The second paper by S. Zhang et al develops a gate-tunable, wide-spectrum organic Phototransistor array for multifunctional in-sensor vision computing. The third paper by P. Tan et al demonstrates a perovskite/CMOS integrated image sensor with multi spectral response from near-infrared to X-ray. The fourth paper by X. Guo et al proposes in-situ image encryption based on α-In_2Se_3/Bi_2O_2Se heterostructure ferroelectric phototransistors integrating photosensing, memory, and encryption functionalities at the edge-sensor level. The fifth paper by W. Xu et al demonstrates a single-electron-sensitive SiC detector featuring an iso-potential guard ring and low-leakage processing technique. The sixth papers by T. Fu et al present a three-tiers hybrid heterogeneous monolithic 3D integration platform combining sensing, memory and logic for spatial depth

perception. The last paper by T.-T. Lin proposes a non-contact vibration sensor with a time-of-flight CMOS single-photo avalanche diode array.

9:05 AM

34-1 Monolithically Integrated Photodiode–Spiking Circuit for Neuromorphic Vision with In-Sensor Feature Extraction

Seungyeob Kim[1], Seungmin Cheon[2], Seonjae Park[2], Taeseung Jung[2], Sanghun Jeon[2]
[1]EE, Korea Advanced Institute of Science and Technology, [2]Korea Advanced Institute of Science and Technology

We present a CMOS-compatible neuromorphic photoreceptor that monolithically integrates an amorphous silicon photodiode with a polysilicon spiking neuron circuit. The device converts continuous-wave illumination ranging from 0.1 to 12 mW/cm^2 into spike trains with frequencies from 200 Hz to 1.2 kHz, modulated by light intensity. Unlike systems requiring pulsed inputs or external spike encoders, this architecture enables one-to-one sensor-to-neuron mapping and reduces digital overhead. Spatial filtering was demonstrated using 3×3 convolutional kernels on 28×28 and 63×63 patterns. When applied to a spiking neural network trained to 95% accuracy, the spike-based inputs outperformed raw image inputs in classification performance.

9:30 AM

34-2 First Demonstration of a Gate-Tunable, Wide-Spectrum (405–850 nm), and 1.7% Average Linearity Error Organic Phototransistor Array for Multifunctional In-Sensor Vision Computing Under Complex Lighting Conditions

Shuaidi Zhang[1,2], Yue Tian[1,2], Haoxiong Ren[1], Xufan Li[1], Jinyao Wang[1], Miao Cheng[1], Yao Li[1], Chi Zhang[1], Yifan Xie[1], Ke Hu[1], Cailian Ma[1], Mengmeng Li[1], Lingfei Wang[1], Feng Zhang[1], Ling Li[1], Ming Liu[1]
[1]Institute of Microelectronics of the Chinese Academy of Sciences, [2]University of Chinese Academy of Sciences

In-sensor computing (ISC) presents an efficient and low-latency paradigm for vision processing, yet existing ISC implementations are sensor-functionality-constrained and task-specific, limiting their general applicability in complex visual environments. To tackle with such issues, gate-tunable linear organic wide-spectrum photo-transistors (GLOW-PTs) are fabricated with organic heterojunctions, featuring broad spectral sensitivity (400-850 nm), gate-controlled positive /negative photoresponse, and exceptional linearity (1.7% average linearity error). Multifunctional image processing tasks on such a unified hardware platform are firstly demonstrated under complex lighting conditions—low light, over/under-exposure, and detail-rich scenes that impose heavy computational loads. Specifically, we achieve near-software inference accuracy (84.8%) on the CIFAR-10 dataset via voltage-programmed on-chip convolution; boost nighttime object detection with an increase in mAP@0.5 by 4.4 percentage-point (pp) over the visible-only baseline on the LLVIP dataset by exploiting integrated infrared sensitivity; and perform dual-frame HDR image restoration using complementary positive/negative photoresponse on the SICE dataset. It significantly expands ISC platform capabilities, delivering versatile solutions for robust, efficient and intelligent vision computing under challenging lighting conditions.

9:55 AM

34-3 A Perovskite/CMOS Integrated Image Sensor with Multispectral Response from Near-Infrared (850 nm) to X-Ray (40 keV) and Record Spatial Resolution

Pengju Tan[1], Shuo Yang[1], Tianyu Liu[1], Ying Meng[1], Sihui Yang[1], Xunyong Yang[1], Mingzhu Hu[1], Hengwei Liu[1], Huaqing Yang[1], Lijuan Huang[1], Yuangan Chen[1], Yiran Hao[1], Yong Guan[1], Xuefei Feng[1], Yu Li[1], Xiaolong Zhao[1], Xiaohu Hou[1], Xin Tang[2], Xiaosong Liu[1], Shibing Long[1], Qin Hu[1]
[1]University of Science and Technology of China, [2]Beijing Institute of Technology

Here we demonstrate a novel perovskite/CMOS integrated image sensor with multispectral response from near-infrared (1.5 eV, 850 nm) to X-ray (40 keV, 0.031 nm) for the first time. This integrated CMOS chip consists of a pixel array of 640×512 with a pixel pitch of 15 μm. Perovskite photosensitive layer was heterogeneously integrated with CMOS through low-temperature solution process, leading to excellent uniformity with low photoresponse non-uniformity of 3.1% under visible light. Under soft X-ray, a high quantum efficiency of 1900% (@850 eV) is achieved and the imaging ability at temperature as high as 100 °C is firstly demonstrated. Hard X-ray imaging results demonstrate that the device achieves a record spatial resolution of 15.3 lp mm^{-1}. This image sensor holds great potential as intelligent electric eye to realize comprehensive medical diagnosis with multispectral imaging.

10:45 AM

34-4 Synergistic Ferroelectric-Photoelectric Multiple Polarization Encryptions with N^{128} Key Spaces in-situ α-In_2Se_3/Bi_2O_2Se Heterostructure Phototransistors

Xinrui Guo[1], Jiying Xiao[2], Kunpeng Si[3], Zuoyuan Dong[1], Yu Zhu[1], Shuo Liu[1], Meng Li[1], Yongji Gong[3], Pengying Chang[2], Lang Zeng[3], Ming He[1]

[1]Peking University, [2]Beijing University of Technology, [3]Beihang University

We demonstrate in-situ image encryption based on α-In_2Se_3/Bi_2O_2Se heterostructure ferroelectric phototransistors (FePTs) that integrate photosensing, memory, and encryption functionalities at the edge-sensor level.Our FePT heterojunction executes multiple encryptions via ferroelectric and photoelectric polarizations, benefiting from the unique ferroelectric-semiconducting features of both α-In_2Se_3 and Bi_2O_2Se. More particularly, the self-oxidation of Bi_2O_2Se to β-Bi_2SeO_5 forms the lattice-matched dielectric interlayer that facilitates the ferroelectric coupling of α-In_2Se_3/Bi_2O_2Se heterostructure, leading to the remarkable ferroelectric polarization window of 5.6 V, programmable conductance states of 128, and endurance cycles of 10^4 for ferroelectric semiconductors. And the anisotropic photoresponse yields an ultrahigh polarization ratio of 3400%. By exploiting these features, we manufacture an in-sensor encryption system with an exceptional key space of N^{128} (i.e., *N* is pixel), reaching an ultralow ciphertext correlation coefficient of 1% and decryption accuracy of 98% in a 100×100-pixel FePT array via single-round key exchange.

11:10 AM

34-5 Single-Electron-Sensitive 4H-SiC Detector for Energy Spectroscopy, Time-Resolved Characterization, and Scattering Electron Imaging in Electron Microscopy

Weizong Xu[1], Renjie Xu[2], Jiuzhou Zhao[2], Yue Guo[2], Tao Miao[2], Jie Jin[2], Ling Zhang[2], Jiawei Tan[2], Guangyang Gu[2], Hang Yu[2], Cheng Mao[2], Dong Zhou[2], Feng Zhou[2], Fangfang Ren[2], Chi Li[3,4], Jiandong Ye[2], Rong Zhang[2], Youdou Zheng[2], Hai Lu[2]

[1]School of Electronic Science and Engineering, Nanjing University, [2]Nanjing University, [3]National Center for Nanoscience and Technology, [4]University of Chinese Academy of Sciences

This work demonstrates a single-electron-sensitive SiC detector featuring an iso-potential guard ring and low-leakage processing techniques, achieving 0.05 pA dark current at bias of 200 V with 200°C stability, linear response covering a wide beam current range from 0.78 pA to 410 nA with a contrast ratio exceeding 10^4 over dark current, and a 95% electron detection efficiency. The detector demonstrates for the first time single-electron-detection-based temporal analysis on compositional and spatial information, where 92.5 ps time jitter in the time-to-digital converter and 6.62 ns time resolution in the detection system has been achieved, facilitating direct single electron energy spectroscopy. The detector superior ruggedness against 1.86 GeV heavy ions and 10^7 A/cm^2 30 keV electron beam irradiation further verifies that, this single electron detector could elevate the electron microscopy from statistical averaging not only to single-quantum-event studies, but also to operations in harsh environment.

11:35 AM

34-6 A Monolithic 3D-Integrated Photoelectric Sensor Featuring 8-bit Storage and Biomimetic Multi-modal Fusion for Spatial Depth Estimation

Tianyue Fu, Xiong Xiong, Xin Wang, Qianlan Hu, Jin Luo, Qianqian Huang, Yanqing Wu

Peking University

We present a three-tiers hybrid heterogeneous monolithic 3D integration platform that combines sensing, memory, and logic for spatial depth perception. The integrated hardware consists of: (1) a bottom-layer commercial 12-inch silicon-based logic platform, (2) a middle-layer oxide-semiconductor 2T0C DRAM, and (3) a top-layer two-dimensional (2D) floating-gate optoelectronic sensor. These functional layers are sequentially stacked at low thermal budget and vertically interconnected through interlayer vias, enabling high integration density on CMOS BEOL. The silicon CMOS logic retains its performance after the entire integration process, while the oxide-semiconductor device demonstrates long-term stability over 8 months, providing a robust foundation for high-performance in-memory multiply–accumulate (MAC) operations. The 2D sensor layer exhibits distinguishable 282-level optoelectronic storage states and a 43.8 times super-additive response due to temporally staggered

plasticity, illustrating the potential of vision–audio synergetic processing for efficient spatial depth estimation in a highly compact hardware architecture.

12:00 PM

34-7 SPAD ToF Array for High-Performance Non-contact Vibration Sensing

Tzu-Tao Lin[1], Tzu-Hsien Sang[2], Chia-Ming Tsai[2], Gray Lin[2], Sheng-Di Lin[2]
[1]Electronics, National Yang Ming Chiao Tung University, [2]National Yang Ming Chiao Tung University

Non-contact vibration sensing is demonstrated with the time-of-flight (ToF) ranging by using a low-jitter CMOS single-photon avalanche diode (SPAD) array and an ultra-short-pulsed laser. The scheme deploying multiple SPADs enhances the sampling rate without compromising the ranging precision and enables the vibration sensing at a distance. As a result, successful measurement of vibration up to 500 Hz with sub-millimeter precision is achieved. This work presents an unprecedented way for vibration sensing and shows the potential of SPAD-ToF ranger for cost-effective and accurate long-distance sensing of vibration.

Session 35: NC | Co-Design for Multimodal & Accurate In-Memory Computing

9:00 AM, Imperial A

Co-Chairs: Ravi Kumar, Infineon Technologies and Thomas Kämpfe, Fraunhofer IPMS / TU Braunschweig

This session has seven papers dealing with technology and system co-design for multimodal and accurate in-memory computing. The first paper by L. Yan, et al., from Peking University, discusses 9Mb 64-core phase change based CIM processor with improved write accuracy, used to implement a multimodal model. The second paper by D. Bondi, et al., from Politecnico di Milano, discusses compensation and mitigation of drift in PCM arrays using a weight mapping technique, for energy efficient in-memory compute. The third paper by Z. Pan, et al., from Peking University, discusses a 40nm PCM based CIM chip that fuses analog and digital CIM, supporting integer and floating-point workloads, with application in embodied AI agents. The fourth paper by Y. Zhou, et al., from UC San Diego, discusses a filament-free bulk RRAM for 3D multilevel compute-in-memory, implementing edge AI tasks. The fifth paper by L. Bao, et al., from Peking University, demonstrates hybrid integration of 2T0C DRAM with RRAM on a CMOS platform, enabling bit-reconfigurable masked matrix vector multiplication acceleration. The sixth paper by S. Kim, et al., from POSTECH, demonstrates a monolithic 3D integrated Electro-Chemical RAM array with excellent yield and device characteristics, showing promise as a candidate for AI accelerators. The seventh paper by E. C. Park, et al., from Hanyang University, demonstrates a fully integrated Hafnia-based neuromorphic computing platform which is multi-modal while offering a scalable and energy-efficient computing solution.

9:05 AM

35-1 Efficient Multimodal Fusion with Intelligent Predictive Programming Based on 64-Core PCM Compute-in-Memory Processor

Longhao Yan[1], Zitong Zheng[2], Xi Li[3], Xile Wang[2], Zhe Zhan[2], Zelun Pan[2], Jiukai Fang[2], Minghui Yin[2], Song Jia[2], Yaoyu Tao[2,2], Zhitang Song[3], Yuchao Yang[2,2,2,4]
[1]School of Integrated Circuits, Peking University, [2]Peking University, [3]Chinese Academy of Sciences, [4]Chinese Institute for Brain Research

Multimodal models that fuse diverse data sources, such as images and text, are pivotal for advancing toward artificial general intelligence. In this work, we fabricate a 9Mb 64-core PCM-based Compute-in-Memory (CIM) processor using a 40nm process. We propose an intelligent predictive programming method that not only actualizes diverse target conductance distributions but also accurately predicts the pre-drift conductance write range in reverse via a machine learning model, improving write accuracy by 26.3% compared to conventional drift models. Furthermore, based on this programming method, we present a software-hardware co-design scheme that leverages the analog and stochastic conductance characteristics of PCM. This scheme efficiently implements the Positive-Incentive distribution, which enhances semantic fusion, and the multimodal model with the CIM architecture. This chip is utilized to realize a medical instrument detection task. Finally, compared to a conventional GPU, our work achieves a 587-fold improvement in energy efficiency and an 85-fold reduction in latency.

9:30 AM

35-2 Linear compensation of drift and mitigation of drift variations in PCM arrays for highly accurate, energy efficient in-memory computing

Danilo Bondì[1], Stefano Giangrasso[1], Alberto Salvadore[1], Giandomenico Panettieri[1], Andrea Al Muktash[1], Letizia Pistolesi[1], Artem Glukhov[1], Lorenzo Cattaneo[1], Nicola Lepri[1], Piergiulio Mannocci[1], Matteo Baldo[2], Andrea Giovanni Bonfanti[1], Daniele Ielmini[1]
[1]Politecnico di Milano, [2]STMicroelectronics

Phase change memory (PCM) offers high density and low current operation for energy efficient in-memory computing (IMC). However, conductance drift is still a major challenge affecting the accuracy of IMC-based edge artificial intelligence (AI) accelerators. This work presents a new weight mapping technique to linearly compensate drift, thus ensuring the linearity of matrix-vector multiplication (MVM). We show a new integrated circuit for MVM capable of current subtraction in the analog domain. High inference accuracy is demonstrated for mapped deep neural networks (DNNs) even after extensive annealing, thanks to differential mapping and increase of read voltage V_{read} to mitigate drift-induced variations.

9:55 AM

35-3 A 40nm Mixed-Precision PCM Chip with Fused Analog/Digital Compute-in-Memory and Adaptive Drift Compensation for Embodied AI Applications

Zelun Pan[1], Longhao Yan[2], Xile Wang[2], Bowen Wang[2], Yueying Luo[2], Zhe Zhan[2], Xi Li[3], Yuqi Li[2], Lei Cai[2], Teng Zhang[2], Minghui Yin[2], Yaoyu Tao[2], Zhitang Song[3], Yuchao Yang[2]
[1]School of Integrated Circuits, Peking University, [2]Peking University, [3]Chinese Academy of Sciences

Embodied AI agents necessitate multiple mixed-precision models running on one chip, presenting challenges when adopting compute-in-memory technologies in terms of memory density, data types, energy efficiency and robustness. We present a 40nm PCM-based computing-in-memory (CIM) chip that fuses analog CIM (ACIM) and digital CIM (DCIM) in one chip, supporting mixed integer and floating-point precision for diverse workloads. To improve memory density and robustness, analog write-verify circuits that enable fast yet accurate conductance programming and hardware-software co-designed chip architecture with adaptive drift compensation are implemented. The chip delivers 13.1× computational loss reduction against conductance drift, and high energy efficiency of 0.821 TFLOPS/W with DCIM and 67.97 TOPS/W with ACIM, providing a promising solution for embodied AI.

10:45 AM

35-4 8-Layer Vertical Filament-Free Bulk RRAM with High Dynamic Range and Energy Efficiency for 3D Multilevel Compute-in-Memory

Yucheng Zhou[1], Yue Zhou[2], Ashwani Kumar[2], Weihong Xu[2], Chang Eun Song[2], Victor Palin[2], Andrea Padovani[3], Gaurav Thareja[4], Ivan K. Schuller[2], Gert Cauwenbergs[2], Tajana Rosing[2], Duygu Kuzum[2]
[1]Department of Electrical and Computer Engineering, University of California, San Diego, [2]University of California, San Diego, [3]DISMI, University of Modena and Reggio Emilia, [4]Applied Materials Inc.

Bulk switching RRAM technologies have been developed to address nonidealities of filamentary RRAM for embedded compute-in-memory applications. However, high density 3D integration and scalability to nano regime have yet to be experimentally demonstrated. Here, we present a scalable, filament-free 3D 8-layer vertical bulk RRAM (b-RRAM) technology optimized for embedded compute-in-memory (CIM) applications. This forming-free device features reliable cycling, multi-level switching, and enhanced speed via hydrogen doping. Guided by multiscale device simulations to optimize the switching stack, we demonstrate 40×40 nm^2 b-RRAM cells with MΩ-level resistance and current nonlinearity, enabling accurate, energy-efficient matrix-vector multiplications (MVM) in selector-less crossbars. A hyperdimensional computing-based continual learning algorithm is implemented on 3D b-RRAM for edge AI tasks, achieving ~90% accuracy—comparable to high-precision floating-point (FP) baselines—while delivering substantial energy savings.

11:10 AM

35-5 Hybrid Integration of IAZO/Si-based 2T0C DRAM with RRAM on CMOS Platform Enabling Bit-reconfigurable Mask for Multitask In-memory Computing

Lin Bao[1,2], Jiaqi Zhao[3], Zongwei Wang[3,4], Haisu Zhang[3], Bowen Jiang[5], Yingtong Ji[3], Yingchen Ji[3], Jingwei Sun[3], Yuhang Yang[3], Ling Liang[3], Zheng Zhou[3], Cuimei Wang[3], Yandong He[3], Shanguo Huang[5], Yimao Cai[3]
[1]School of Integrated Circuits, Peking University, [2]State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, [3]Peking University, [4]YanXin MicroElectronics Co., Ltd. (YXME), [5]Beijing University of Posts and Telecommunications

Embodied intelligence demands integrated multimodal sensory processing for real-time decision-making, necessitating computationally efficient and architecturally adaptable solutions. Multitask neural networks (MNNs) address this through weight-shared feature extraction layers, enabling dynamic activation of task-specific subnetworks via dynamic masked matrix-vector multiplication (DM-MVM). This work introduces a hybrid in-memory computing (H-IMC) architecture tailored for bit-reconfigurable DM-MVM acceleration. Leveraging monolithic 3D integration, each H-IMC unit combines oxide-semiconductor transistor, Si transistor, and RRAM in a 2T0C-1R configuration, co-locating volatile (2T0C) and non-volatile (RRAM) memory. Experimental validation demonstrates > 2700 s dynamic mask retention and 5-bit static weight storage, fulfilling DM-MVM's dual requirements. A dedicated H-IMC array is designed for MNN, achieving experimental verification of the H-IMC paradigm. With device-technology co-optimization (DTCO), the H-IMC-based MNN yielded 93.6% average accuracy in multi-task processing, with 30× latency reduction and 32× energy efficiency gains over conventional 1T1R architectures, paving the way for multi-task IMC technology.

11:35 AM

35-6 First M3D-Integrated 4K ECRAM Chip with Sub-100 nm Channels Achieving 100% Yield and Highly Uniform Performance for Analog AI Computing

Seungkun Kim[1], Woochan Song[2], Jeonghoon Son[2], Minil Kang[3], Minseong Um[3], Jinho Byun[2], Junyoung Choi[2], Seongchan Bae[2], Bohyeon Kang[2], Rock-Hyun Baek[2], Jae-Yoon Sim[2], Hyung-Min Lee[3], Seyoung Kim[2]
[1]Department of Materials Science and Engineering, Pohang University of Science and Technology (POSTECH), [2]Pohang University of Science and Technology (POSTECH), [3]Korea University

For the first time, we demonstrate a monolithic three-dimensional (M3D)-integrated 4K electrochemical random-access memory (ECRAM) array with sub-100 nm channels, fabricated atop of the Si-CMOS logic circuits using a back-end-of-line (BEOL)-compatible process (≤ 200 ℃). These devices exhibit significantly low device-to-device variation ($\sigma/\mu \sim 0.026$), along with outstanding endurance (5×10^6 pulses). By scaling channel dimensions down to 100 nm through electron-beam lithography, we attain a remarkable 100-fold improvement in programming speed. Full array-level analysis reveals that our proposed M3D-integrated ECRAM-based processing-in-memory (M3D-ECPIM) chip achieves 100% yield and exhibits highly uniform analog switching behavior across the entire 4K array. Additionally, the first chip-level evaluation of symmetry point and half-bias selectivity confirms that both meet the stringent specifications required for neural network training. Finally, the M3D-ECPIM chip demonstrates excellent pattern programming accuracy with only 0.05% error.

12:00 PM

35-7 First Hardware Demonstration of Fully Integrated Hafnia-based Multimodal Neuromorphic Computing Platform

Eun Chan Park[1], Minsuk Song[2], Jangsaeng Kim[3], Jongwoo Lee[2], Ryun-Han Koo[4], Jiseong Im[4], Wonjun Shin[5], Daewoong Kwon[2]
[1]Department of Electronic Engineering, Hanyang University, [2]Hanyang University, [3]Sogang University, [4]Seoul National University, [5]Sungkyunkwan University

We demonstrate for the first time a hafnia-based multimodal neuromorphic computing (HMNC) platform via co-integration of volatile and nonvolatile FeTFTs. Material-device co-optimization of the FeTFTs enables diverse functionalities for multimodal data processing. Full hardware demonstration achieves robust performance for multimodal MNIST data (88%) and emotion/sentiment recognition (66%), offering a scalable and energy-efficient neuromorphic solution.

Session 36: PMA | GaN and III-V Integration for Next-Generation RF Devices

9:00 AM, Imperial B

Co-Chairs: Jose Jimenez, Qorvo and Yuhao Zhang, The University of Hong Kong

This session includes 7 papers highlighting advances in GaN and III-V device technologies spanning mobile applications, integration platforms, high-power density devices, and emerging ultra-wide bandgap RF transistors. The first paper by Haochen Zhang benchmarks GaN RF technology in a representative mobile handset application, demonstrating superior efficiency above 50%. The next three papers highlight new integration concepts and platform technologies: the second paper by Han Wui Then introduces a fully-developed GaN chiplet technology based on 300 mm GaN-on-Si, the third paper by Bin Hou reports an Au-free monolithic E/D-mode GaN platform on HR-Si enabling both RF and digital logic applications with record power density, and the fourth paper by Jaeyong Jeong presents the monolithic 3D integration of III-V HEMTs on glass using an ultra-thin dielectric bonding layer for sub-THz applications. The following two papers focus on high power density and thermal management: the fifth paper by Hong Zhou demonstrates low thermal resistance AlGaN/GaN/AlN-on-SiC RF HEMTs with record Pout = 41 W/mm and fT·BV = 31 THz·V, while the sixth paper by Jeong-Kyu Kim highlights the dual role of topside diamond as both passivation and heat spreader in AlGaN/GaN HEMTs. Finally, the seventh paper by Seiya Kawasaki reports the first RF operation of ultra-wide bandgap AlGaN-channel FETs with average Al content exceeding 0.75.

9:05 AM

36-1 First Integration of GaN Low-Voltage PA MMIC into Mobile Handsets with Superior Efficiency Over 50%

Haochen Zhang[1], Yue Sun[2], Hongtu Qian[1], Jianan Liu[2], Shuiling Fan[1], Xiao Han[1], Yongsheng Zhang[1], Hui Zhang[1], Xinchuan Zhang[1], Junzhuo Qiu[1], Yi Pei[1], Shui Liu[2], Haiding Sun[1], Kevin Jing Chen[3], Nanqian Zhang[1]
[1]Dynax Semiconductor Inc., [2]Xiaomi Communications Company Ltd., [3]The Hong Kong University of Science and Technology

The demand for compact, efficient and agile power-amplifiers PAs continues to rise. GaN-on-Si technologies are anticipated to deliver PAs with superior power performance, smaller form factors, and lower costs. However, their typically high operating-voltage requirements are often incompatible with the power-control units used in current handsets. In this work, we demonstrate a GaN-on-Si low-voltage HEMT technology achieving PAE of over 80% and P_{out} of 2.84W/mm at drain voltages below 10V at 2.2GHz. In a three-stage MMIC PA, we verify system-level power performance in real handset applications for the first time. At 2.0GHz, the MMIC achieves a gain of 39.3dB, P_{out} of 39.3dBm, and PAE of 60.1%, while maintaining the required linearity for efficient handset operation. The outstanding device- and circuit-level power performance of our GaN-on-Si HEMTs surpasses that of other technologies for low-voltage RF applications, strongly indicating the potential of GaN-on-Si RF technologies in next-generation mobile-communication systems.

9:30 AM

36-2 GaN Chiplet Technology Based on 300mm GaN-on-Silicon

Han Wui Then, Samuel Bader, Ahmad Zubair, Heli Vora, Prafful Golani, Pratik Koirala, Michael Beumer, Trianggono Widodo, Aris Orbase, Patrick McIllece, Andrey Vyatskikh, Thomas Hoff, Rambert Nahm, Kirby Maxey, Dimitri Frolov, Isha Khandelwal, Shreyas Jayaprakash, Kaushik Narayanan, Jag Rangaswamy, Marko Radosavljevic
Intel Corporation

We demonstrate for the first time a GaN chiplet technology based on 300mm GaN-on-silicon process for high-performance, high-density, efficient power electronics and high-speed/RF electronics. This GaN chiplet technology features: (a)industry's thinnest GaN chiplet with an underlying silicon substrate that is only 19µm thick, harvested from a fully-processed, thinned and singulated 300mm GaN-on-silicon wafer; (b)industry's first library of fully-functional integrated on-die CMOS digital circuits, from logic-gates, multiplexers, and flip-flops to ring-oscillators, using monolithically integrated GaN N-MOSHEMT and Si PMOS process; (c)promising results in TDDB, pBTI, HTRB and HCI studies indicating that the 300mm GaN MOSHEMT technology can meet required reliability metrics.

9:55 AM

36-3 Demonstration of 5.74/11.0 W/mm Au-free GaN Monolithic-integrated E/D-mode Platform on HR-Si for RF and Digital Logic Applications

Bin Hou[1], Jiale Du[2], Hao Lu[2], Ling Yang[2], Yifan Chen[2], Tianqi Pan[2], Kaiwen Chen[2], Hailin Liu[2], Qingyuan Chang[2], Qing Zhu[2], Meng Zhang[2], Mei Wu[2], Mao Jia[2], Minhan Mi[2], Jiejie Zhu[2], Xiaohua Ma[2], Yue Hao[2]
[1]School of Microelectronics, Xidian University, [2]Xidian University

In this work, we demonstrate a GaN based monolithic integrated platform on high-resistivity (HR) Si substrate for RF and logic applications using Au-free metal technology. By employing an optimized Au-free metal system with ohmic regrowth technology, an ultra-low ohmic contact resistance (R_C) of 0.058 Ω·mm was achieved. An ALE technology was applied in the E-mode operation with controlled etching depth and low etching damage. Leveraging high-quality epitaxial growth and atomic layer etching (ALE) technology, an output power density (P_{out}) of 5.74 W/mm (E-mode) and 11.0 W/mm (D-mode) Au-free GaN RF devices were attained at 3.6 GHz under operating voltage of 45 V. Finally, the Au-free p-GaN E/D-mode logic circuits were successfully realized, exhibiting excellent logic and noise performance. These results highlight the potential of the Au-free monolithic integration platform on HR-Si for mixed-signal applications, spanning RF analog and digital logic circuits.

10:20 AM

36-4 Monolithic 3D Integration of III-V HEMTs on Glass Using Ultra-Thin Dielectric Bonding Layer: A High-Frequency and Low-Loss Active Glass Platform for Sub-THz Applications

Jaeyong Jeong[1], Yurim Choi[2], Seongju Kim[2], Hyeongjun Kim[2], Jimin Kwon[2], Sanghyeon Kim[1]
[1]KAIST, [2]UNIST

The growing demand for sub-THz wireless systems requires not only high-frequency RF devices but also compact, low-loss integration platforms. In this work, we demonstrate a monolithic III-V/glass RF platform using a low-temperature wafer bonding technique, where III-V enables high-frequency transistors and glass offers a low-loss, low-cost, and dimensionally stable platform. The InGaAs HEMTs fabricated on glass achieve fT of 343 GHz and fMAX of 659 GHz. The glass platform features on-glass CPW lines, achieving 0.1 dB/mm at 40 GHz. Thermal simulations, validated by thermoreflectance microscopy, show that the introduction of through-glass vias can reduce thermal resistance of 3D integrated InGaAs HEMTs by 65.5%, bringing thermal performance close to that of conventional InP substrates. This is the first report of sub-THz III-V device operation on a glass platform, highlighting its potential for future sub-THz wireless systems.

11:10 AM

36-5 Low Thermal Resistance X-band AlGaN/GaN/AlN-on-SiC RF Power HEMTs with Record Pout = 41 W/mm and fT×BV = 31 THz·V

Hong Zhou[1], Kun Zhang[1], Chaoqun Zhang[1], Qifeng Lyu[2], Hui Zhang[2], Naiqian Zhang[2], Peng Dong[3], Jia Cao[3], Yali Mao[4], Hehe Gong[5], Zhihong Liu[1], Chao Yuan[4], Yachao Zhang[1], Yi Pei[2], Yuhao Zhang[5], Yue Hao[1], Jincheng Zhang[1]
[1]Xidian University, [2]Dynax Semiconductor Inc., [3]Chengdu Aerospace Bomu Electronic Science and Technology Co., Ltd, [4]Wuhan University, [5]University of Hong Kong

This work demonstrates a new record P_{out} of 41 W/mm at 10 GHz in an AlGaN/GaN/ AlN-on-SiC RF power HEMT. This breakthrough is enabled by deploying an ultra-wide-bandgap (UWBG) AlN buffer and reducing the combined channel and buffer thickness to just 400 nm. By optimizing the V:III ratio and growth rate during AlN epitaxy, we suppress conventional island-like nucleation, achieving rapid film coalescence, low dislocation density, and reduced thickness. Thermoreflectance (TR) measurements reveal a low device R_T of 2.1 K·mm/W, which is less than 1/3 of conventional device with μm-thick GaN buffer and island-nucleated AlN layer. Meanwhile, the use of UWBG AlN buffer and thin GaN channel boost device breakdown voltage (BV) and carrier confinement. As a result, the fabricated RF HEMTs achieve a P_{out} of 41 W/mm at 10 GHz with an associated power-added efficiency (PAE) of 51% and a record high f_T×BV figure-of-merit (FOM) of 31 THz·V.

11:35 AM

36-6 Demonstration of the Dual Role of Topside Diamond as a Passivation Layer and Heat Spreader in AlGaN/GaN HEMTs

Jeong-Kyu Kim[1], Kelly Woo[2], Mohamadali Malakoutian[2], Rohith Soman[2], Wiley Yu[2], Devansh Saraswat[2], Junrui Lyu[2], Chuanzhe Meng[2], Rafael Martinez[2,3], Srabanti Chowdhury[2]
[1]Electrical Engineering, Stanford University, [2]Stanford University, [3]Broadcom

Self-heating remains a critical barrier to unlocking the full performance and long-term reliability of AlGaN/GaN HEMTs. In our prior work, we explored topside diamond heat spreaders; however, their placement above conventional thick passivation layers limited thermal coupling to the active region. Here, we demonstrate—for the first time—a topside polycrystalline diamond (PCD) layer that serves simultaneously as an effective surface passivation layer and a heat spreader, enabling closer thermal integration. The PCD was grown on fully-fabricated GaN HEMTs at temperatures below 450 °C. Electrical characterization (DC, pulsed-IV, small-signal, and load-pull) confirms full compatibility with device operation. Gate resistance thermometry shows a substantial ~55 °C reduction in average channel temperature at ~18 W/mm. Notably, DC-RF dispersion is dramatically suppressed, with knee current mismatch between DC and pulsed-IV reduced from ~30% to ~5%. These results position topside PCD as a transformative dual-function solution for next-generation thermal and electrical stability in RF GaN HEMTs.

12:00 PM

36-7 First RF Operation of AlGaN-channel Polarization-Doped FETs with Average Al-content Over 0.75

Seiya Kawasaki, Masanobu Hiroki, Kazuyuki Hirama, Yoshitaka Taniyasu
Basic Research Laboratories, NTT, Inc.

We present the first RF operation of polarization-doped FETs (pol-FETs) with average AlN mole fractions (x_{ave}) of over 0.75, enabled by a downward-graded AlGaN underlayer and a selectively-regrown n^+-graded-AlGaN contact. The maximum drain current is 0.78 A/mm, which is among the highest reported for AlGaN-channel pol-FETs ($x_{ave} \geq 0.5$). The fabricated devices with x_{ave} up to 0.89 showed an f_{max} exceeding 60 GHz and saturation output power density (P_{sat}) of over 0.3 W/mm at 2.4 GHz. The highest f_{max} of 78.7 GHz and the highest P_{sat} of 1.2 W/mm were obtained at x_{ave} of 0.85 and 0.78, respectively.

Session 37: EDT | Advances in Memory and Computing Using Hafnia based Ferroelectrics

1:30 PM, Grand Ballroom A

Co-Chairs: Suhwan Lim, Samsung and Nazila Haratipour, Kepler Computing

This session includes 8 papers that describe recent progress in ferroelectric based memory & computing. The first paper in this session is an invited contribution, by Jesus del Alamo of Microsystems Technology Laboratories, MIT, describes polarization switching dynamics in scaled oxide channel FEFETs. The second paper, by Dong Zhang, National University of Singapore, presents a scaled split-gate FEFET to optimize gate-area-ratio for achieving a record-low operating voltage of 0.5 V and record-high memory window-to-operation voltage. The third paper, by Yang Feng, National University of Singapore, proposes a novel BEOL-compatible 3-layer stacked nanosheet FEFET with oxide channel. In this novel geometry, they optimize the ferro-to-channel area-ratio to improve gate controllability and erase efficiency. The fourth paper, by Chen Sun, National University of Singapore, demonstrates 3D ferroelectric vertical NAND with dual-layer oxide channel and laminated FE stack to improve the memory window and on/off current ratio. The fifth paper, by Hwiho Hwang, Department of Semiconductor Engineering, Hanyang University, demonstrates a ferroelectric memcapacitor crossbar array based on NAND flash structure for in-memory computing. The sixth paper, by Jin Luo, School of Integrated Circuits, Peking University, proposes a ferroelectric charge-steering memcapacitor device-based delta-computing-in-memory. The seventh paper, by Jiahui Duan, University of Notre Dame, proposes a NAND FEFET-based XOR-cipher encryption scheme leveraging in-situ memory operations for minimal overhead. The last paper in this session, by Jeong-Han Kim, Department of Electronic Engineering, Hanyang University, proposes a vertical ferroelectric tunnel junction (VFTJ) array enabling hardware-level implementation of homomorphic encryption for secure federated learning.

1:35 PM

37-1 Single-domain Switching Dynamics in BEOL Nanoscale Ferroelectric Field-effect Transistors (Invited)

Yanjie Shao[1], Minsoo Kim[2], John C-C Huang[1], Dimitri A Antoniadis[1], Jesus A del Alamo[1]
[1]MIT, [2]Samsung Electronics

We study polarization switching dynamics in back-end-of-the-line (BEOL) oxide-channel ferroelectric field-effect transistors (FE-FETs) based on 10-nm-thick FE-$HfZrO_2$ (FE-HZO). We demonstrate single-domain switching in nanoscale devices at a voltage amplitude < 1.8 V and pulse width < 10 ns. Leveraging this device platform, we have studied FE polarization accumulation and relaxation in response to a pulse train in wide as well as very narrow devices where we can observe single-domain switching. Remarkably, polarization accumulation and relaxation dynamics is virtually identical in both types of devices, suggesting a large stochasticity in single domain switching which dominates that of a large domain ensemble. Our results suggest that a broad distribution of coercive field exists even within a single domain which would imply that the stochasticity of switching in large-area devices arises entirely from that of the single domain.

2:00 PM

37-2 First Demonstration of Oxide Semiconductor BEOL-Compatible Dual-Gate FeFET Achieving a Record-Low Operation Voltage of 0.5 V and a Record-High Memory Window-to-Operation Voltage Ratio of 1.62

Zhang Dong[1], Xinming Li[1], Xiaopeng Li[2], Jiezhi Chen[2], Xiao Gong[1]
[1]National University of Singapore, [2]Shandong University

In this work, we present and experimentally demonstrate the first oxide-semiconductor (OS) BEOL-compatible dual-gate FeFET (DG-FeFET) with channel length (LCH) down to 40 nm, achieving record-breaking performance metrics. The key highlights of this work include: (1) A novel gate-area-ratio (GAR) engineering strategy is introduced to modulate the voltage distribution across the gate stack, effectively overcoming the intrinsic operation voltage limitations of conventional FeFETs. (2) The DG-FeFET with a 5 nm HZO layer achieves a record-low operation voltage 0.5 V among all reported FeFETs, while maintaining a decent memory window of 0.5 V, a high current window of 104, excellent endurance exceeding 109 cycles, and fast programming speed down to 100 ns. (3) The DG-FeFET with an 8 nm HZO layer exhibits a substantial memory window of 2.03 V at an ultra-low operation voltage of 1.25 V, leading to a record-high memory window-to-operation voltage ratio of 1.62 among reported FeFETs.

2:25 PM

37-3 First Demonstration of BEOL-compatible Stacked Nanosheet FeFETs Enabling Enhanced Area Efficiency and Record Memory Window

Yang Feng[1,2], Xiaolin Wang[1], Jiezhi Chen[3], Xiao Gong[1,4]
[1]National University of Singapore, [2]Shandong university, [3]Shandong University, [4]A*STAR

For the first time, we propose and demonstrate the BEOL-compatible stacked nanosheet FeFETs. By integrating a three-layer oxide semiconductor thin-film transistor onto a fully ALD-deposited MFM capacitor, the fabricated FeFETs with a metal–ferroelectric–metal–insulator–semiconductor stack achieve several key advancements: (1) The enlarged channel area redistributes more applied voltage across the ferroelectric layer, yielding a MW greater than 4 V with an ~10 nm-thick HZO; (2) over 80% channel footprint area is saved compared to planar devices while achieving the same MW; (3) improved electrostatic control mitigates negative shift of V_{th} and MW degradation as L_{ch} scaling down to 50 nm; and (4) enhanced reliability of the FE layer enables excellent endurance, exceeding 10^{10} cycles, while maintaining a stable MW without dielectric breakdown. Our work highlights the impact of transistor structure engineering on the performance of BEOL-compatible FeFETs, demonstrating their promise for scalable, high-performance integration in future memory and computing architectures.

2:50 PM

37-4 First Demonstration of 3D Fe-VNAND with Function-Separated Dual-Layer Oxide Semiconductor Channel and Laminated Ferroelectric Stack

Chen Sun[1], Xinming Li[1], Dong Zhang[1], Wei Shi[1], Yue Chen[1], Gan Liu[1], Rui Shao[1], Qiwen Kong[1], Yuxuan Wang[1], Leming Jiao[1], Zuopu Zhou[1], Yang Feng[1], Xiaolin Wang[1], Kai Ni[2], Xiao Gong[1]
[1]National University of Singapore, [2]University of Notre Dame

In this work, we demonstrate the first 3D Fe-VNAND architecture featuring a fully ALD-based dual-layer OS channel and a laminated Fe stack. The OS channel integrates oxygen-deficient ZnO to enhance polarization switching and high-mobility In_2O_3 to ensure robust vertical string conduction. By fine-tuning the oxygen deficiency in planar ZnO FeFETs, a remarkable 84% increase in MW is achieved. This MW can be further boosted by optimizing the Fe stack lamination, realizing an ultra-wide MW of 16.1 V. In 3D Fe-VNAND, the addition of an In_2O_3 layer to form an In_2O_3-ZnO dual-layer OS channel effectively overcomes conduction limitations in string current caused by high-resistance space regions between cells, enabling a substantial improvement in I_{ON}/I_{OFF} ratio and achieving a large MW of 9.6 V under a 10 V DC sweep. Furthermore, the non-overlapping WL-to-S/D structure in Fe-VNAND, facilitated by the dual-layer OS design, results in outstanding retention and endurance exceeding 10^9 cycles.

3:40 PM

37-5 Ferroelectric Memcapacitor Crossbar Array with NAND Flash Structure for In-Memory Computing

Hwiho Hwang[1], Junsu Yu[2], Sangwook Youn[1], Hyungjin Kim[1]
[1]Hanyang University, [2]Samsung Electronics

In this paper, 48×24 ferroelectric memcapacitor crossbar is experimentally demonstrated based on NAND flash structure, with TiN/HZO/SiO_2/poly-Si gate stack. With the optimal read voltage, a wide charge memory window of ~12.27 fC/μm^2 is achieved. Read, write, and inhibit operations are successfully verified across the fabricated array, enabling reliable 4-bit multi-level programming with retention over 10 years. Vector-matrix multiplication operations are confirmed with an R^2 value of 0.984, with the verification of output sensing circuit. Also, CIFAR-10 classification is implemented on the fabricated 48×24 array by pre-trained weight transfer with 4-bit device states, achieving accuracy of 87.75%. The feasibility of vertically stacked structure based on 3D-NAND flash is also verified by TCAD and SPICE simulations.

4:05 PM

37-6 A Novel Ferroelectric Charge-Steering Memcapacitor Array based Δ-CIM with Enhanced Parallelism and Self-Adaptive Sparsity for Edge LLMs

Jin Luo, Shaodi Xu, Peiyu Chen, Yuxin Lin, Haoyu Wang, Bingrui Song, Weikai Xu, Boyi Fu, Yuan Wang, Le Ye, Qianqian Huang, Ru Huang
Peking University

In this work, a novel **f**erroelectric (FE) charge-**s**teering mem**cap**acitor (**FS-Cap**) device-based delta-computing-in-memory (Δ-CIM) is proposed, experimentally demonstrating co-designed spatial and temporal delta-encoding schemes. It combines differential capacitance (ΔC) storage with delta-input (ΔIN) encoding to achieve high parallelism and energy efficiency for edge large language models (LLMs). At device level, the three-terminal FS-Cap leverages asymmetric doping for differential dual-ended extraction of MOS capacitor, while optimized FE-assisted gate-injection storage can steer the induced-charge between dual-ends. It enables multi-level ΔC and allows single-device unification of signed multiply-accumulate (MAC) and content-addressable memory (CAM) operations with enhancing noise suppression for improved parallelism. Moreover, the system implementation of FS-Cap array based Δ-CIM incorporates bit-transition-aware ΔIN encoding, exploiting adaptive sparsity to achieve 84.5% charging energy reduction in experiments. Based on the FS-Cap Δ-CIM architecture, inference and fine-tuning achieve 33.9× enhancement in SNR and parallelism, along with superior energy efficiency, showing great potential for advanced AI systems.

4:30 PM

37-7 Securing 3D NAND Without Density Loss via In-Situ Encryption Using a Single Transistor XOR Cell

Jiahui Duan[1], Sanwar Ahmed Ovy[2], Po-Kai Hsu[3], Zijian Zhao[1], Thomas Kampfe[4], Xiao Gong[5], Shimeng Yu[3], Vijaykrishnan Narayanan[6], Sumitha George[2], Kai Ni[1]
[1]University of Notre Dame, [2]North Dakota State University, [3]Georgia Institute of Technology, [4]Fraunhofer IPMS, [5]National University of Singapore, [6]Pennsylvania State University

In this work, we push lightweight XOR-based in-situ encryption to extreme density by proposing a single transistor XOR memory cell in 3D NAND, enabling secure data storage without density loss. Using a FeFET as an example technology, we demonstrate that: i) a 1FeFET memory can realize the XOR by exploiting the ability to charge the source/drain separately and control current flow direction, eliminating the need for conventional encrypted cells

that rely on complementary devices; ii) with a XOR-based cipher, encryption and decryption can be mapped to in-situ array operations, where ciphertext is stored as the threshold voltage of FeFETs, and decryption is achieved through read operations using key-dependent complementary source/drain bias; iii) the proposed technique is scalable to MLC storage by encrypting and decrypting data bit by bit; iv) using an integrated NAND FeFET array, we experimentally demonstrate encryption and decryption operations for both single-level cell and MLC storage.

4:55 PM

37-8 First Demonstration of Federated Learning with Homomorphic Encryption Using Vertically-stacked Ferroelectric Tunnel Junction

Jeong-Han Kim[1], Jiseong Im[2], Minsuk Song[3], Yelin Yoo[3], Jongwoo Lee[3], Ryun-Han Koo[2], Jangsaeng Kim[4], Daewoong Kwon[3]
[1]Department of Electronic Engineering, Hanyang University, [2]Seoul National University, [3]Hanyang University, [4]Sogang University

A novel vertical ferroelectric tunnel junction (VFTJ) array enabling hardware-level implementation of homomorphic encryption (HE) for secure federated learning (FL) is proposed. A highly integrated 4-stack TiN/SiO_2 structure with vertically etched hole-type architecture was fabricated. Distinct ferroelectric and ionic switching induced by interfacial TiO_x formation significantly enhanced tunneling electroresistance (TER) and switching reliability. The intrinsic randomness of devices enabled the realization of polynomial samplers (ternary, Gaussian, uniform) essential for HE schemes. The proposed VFTJ-based FL achieves accuracy comparable to the software baseline, exhibiting less than 4% accuracy reduction under highly non-IID conditions.

Session 38: ALT | System Optimization and Process Innovation

1:30 PM, Grand Ballroom B

Co-Chairs: Veeraraghavan Basker, Applied Materials and Vincent Chang, TSMC

With the advent of AI applications, system-level optimization is key to achieve peak performance and minimize energy consumption. This session starts with an invited paper reviewing a comprehensive understanding on interactions of device level up to packaging optimization to maximize the performance of Qualcomm Snapdragon® chips. In addition, new approaches including Dual-sided signal and Carbon nanotube backside clock distribution will be covered. Further, innovative approaches of process simplification with known materials used to address key challenges like multi-Vt modulation and EOT scaling will be presented. Finally, novel materials to address the critical issue of thermal dissipation for AI applications will be covered.

1:35 PM

38-1 Comprehensive Enhancements for AI Mobile Edge Computing: Innovations from Device-Level Advancements to Advanced Packaging and System Optimization (Invited)

Youseok Suh, Jie Deng, Jun Yuan, Frank Yang, Jihong Choi, Giri Nallapati, Ronald Kuse, PR. Chidi Chidambaram
Qualcomm

The increasing demand for real-time, intelligent applications has made AI mobile edge computing a critical enabler of next-generation user experiences. By performing inference directly on-device, it minimizes latency, enhances data privacy, and improves responsiveness. This paper presents a systematic approach to boosting product performance for AI mobile edge computing, covering innovations from device-level architecture to system integration and advanced packaging. Through processor design advancements, System-Technology Co-Optimization (STCO), Design-Technology Co-Optimization (DTCO), Design-Process Co-Optimization (DPCO), we demonstrate significant improvements in performance, power, area, and cost (PPAC). We conclude by showcasing Qualcomm's latest achievements with the leading AI mobile Edge computing chip, Snapdragon® X Elite.

2:00 PM

38-2 Extending Vertical FET for Advanced Logic Scaling with Architecture Innovations: Dual-sided Global Signal, Flip VFET and Omega Nanosheet

Ziqiao Xu[1], Yanbang Chu[2], Yimeng Wang[2], Siyuan Liu[2], Zhihao Wang[2], Xinyue He[2], Yu Liu[2], Haoran Lu[2], Yandong Ge[2], Xun Jiang[2], Yongqin Wu[2], Lijie Zhang[2], Weihai Bu[2], Yibo Lin[2], Runsheng Wang[2], Ming Li[2], Heng Wu[2], Ru Huang[2]
[1]School of Integrated Circuits, Peking University, [2]Peking University

For the first time, VFET is comprehensively studied in terms of the interconnect routability, device scalability and current drivability. Based on dual-sided VFET (DSVFET) with better device symmetry and area efficiency, dual-sided global signal (DGS) routing by flipping certain nets to the wafer backside was successfully developed. With DGS, the studied RISC-V core gains ~15% less power at iso-frequency. The scalability of DSVFET was also explored. By scaling down the vertical Nanosheet pitch, 14% less energy-delay-product and 24% smaller area were confirmed at block level. For ultimate scaling, a novel Flip VFET with self-aligned back-to-back VFET stacking was proposed, delivering ~33% even smaller area. Furthermore, to overcome the drivability concerns, the Omega NS channel for DSVFET was proposed and experimentally demonstrated for the first time, providing 37.5% more effective gate width and 6% frequency gain without wasting area. This work provides new solutions to the key issues of the VFET.

2:25 PM

38-3 Active BSCDN Benchmark Framework with Backside-Compatible CNFET Logic Technology

Minho Park[1], Yehyun Shin[2], Ikkyum Kim[2], Junghyun Yoon[2], Seunghun Baek[2], Seongmin Eum[2], Heesoo Yang[2], Yurim Choi[2], Jaeyong Jeong[3], Sanghyeon Kim[3], Haksoon Jung[2], Seongju Kim[2], Heechun Park[2], Jimin Kwon[2]
[1]Department of Electrical Engineering, UNIST, [2]UNIST, [3]KAIST

This study presents a benchmark framework for digital systems integrating an active backside clock distribution network (BSCDN) using carbon nanotube field-effect transistor (CNFET) logic. Unlike conventional CDN schemes, the proposed architecture relocates clock buffers and flip-flops entirely to the backside, enabling complete clock routing in BS metal layers and freeing frontside resources for signal interconnect. Complementary CNFETs were fabricated with backside-compatible low-temperature processes and modeled using TCAD and neural network-based compact models. Standard cells were characterized and placed between BS metal layers, with signal routing connected through nano TSVs. Benchmarking with OpenCores designs shows that active BSCDN reduces clock skew and latency while lowering chip area, wirelength, and power consumption compared to frontside and passive BSCDN counterparts. Thermal analysis highlights the importance of heat-aware design, as backside integration increases power density. This work demonstrates the feasibility and benefits of active BS logic for efficient and scalable digital system design.

2:50 PM

38-4 First Demonstration of LaO_x Atomic Flip Dipole Technique with Single Dipole Shifter in CMOS GAA NSFETs to Achieve ~66 mV/level Multi-V_Ts within 330 mV Range

Yu Wang, Jiaxin Yao, Yanzhao Wei, Qingzhu Zhang, Feixiong Wang, Shuang liu, Lei Cao, Yuzhe Yu, Shuai Yang, Wenjuan Xiong, Junfeng Li, Huaxiang Yin, Xiaolei Wang, Jun Luo
Institute of Microelectronics of the Chinese Academy of Sciences

In this paper, for the first time, we proposed one novel atomic LaO_x flip technique to achieve WFM-less multi-V_Ts modulation by single dipole shifter in CMOS GAA NSFETs. With mitigating atomic LaO_x dipole layer from down-side HfO_2 surface to up-side, named in Atomic Flip Dipole Technique (AFDT), the dipole force direction is successfully inverted and causing a band-edge impact on p-type WFM modulation, resulting in a linear 4-level V_Ts modulation within a ~250 mV range of NSFETs. Furthermore, a flip/interfacial hybrid structure for CMOS integration by single LaO_x dipole shifter is demonstrated, enabling a 6(N) + 6(P) multi-V_Ts modulation result with an average of ~66mV/level within a ~330 mV(N)/400 mV(P) range with less carrier mobility penalty, which provides a promising method to lessen limitations in p-type dipole modulation technology and one lower cost dipole V_T modulation solution for CMOS NSFET integration with higher transistor performance.

3:40 PM

38-5 First Demonstration of Quasi-Single-Crystal (QSC) HfZrON Gate Dielectric (κ=67) with sub-5 Å EOT and 4×10^{-3}A/cm^2 J_g for Si NSFET Scaling

Kun Zhong[1,2], Zhaohao Zhang[2,3], Tianjiao Xin[4], Junshai Chai[2,3], Hong Yang[2,3], Lei Cao[2,3], Siyuan Liu[2,3], Yadong Zhang[2,3], Jiacheng Shi[5], Rui Chen[5], Qiang Zhang[5], Chunsong Zhao[5], Qingzhu Zhang[2,3], Jianfeng Gao[2,3], Junfeng Li[2,3], Yongliang Li[2,3], Yan Cheng[4], Huaxiang Yin[2,3], Xiaolei Wang[2,3], Jun Luo[2,3]
[1]State Key Laboratory of Fabrication Technologies for Integrated Circuits, Institute of Microelectronics of the Chinese Academy of Sciences, [2]University of Chinese Academy of Sciences, [3]Institute of Microelectronics of the Chinese Academy of Sciences, [4]East China Normal University, [5]Huawei Technologies Co Ltd

For the first time, we demonstrate and integrate an ultra-thin dielectric with sub-5 Å EOT enabled by Quasi-Single-Crystal (QSC) HZON into GAA-NSFETs. Through interfacial nitrogen engineering, a distinct QSC structure is achieved, featuring exceptionally large grain size (up to ~180 nm) and high *t&o*-phase ratio (96%), which yields record-high κ value of 67. Sub-5 Å EOT, with extremely low leakage (9×10^{-2}A/cm^2 with ~4.6 Å EOT; 4×10^{-3}A/cm^2 with ~4.9 Å EOT) and high reliability is realized on QSC-HZON MOSCAPs. The QSC-HZON is further integrated into *n*-NSFETs, achieving ~35% I_{on} enhancement, providing a straightforward way for NSFET's performance improvement and further scaling.

4:05 PM

38-6 Efficient Power Delivery Network of Enhancement and Depletion Mode GaN FeMHEMT by Ferroelectric Gate Stack for Voltage Regulators

Che-Chi Cheng[1], Heng Liu[2], Zhao-Feng Lou[2], Cheng-Hong Liu[2], Min-Hung Lee[2,2,2]
[1]Program for Semiconductor Devices, Materials, and Hetero-integration, National Taiwan University, [2]National Taiwan University

Dynamical Modulating E-mode and D-mode of power management devices of FeMHEMT are proposed by ferroelectric gate stack for the voltage regulators of power delivery network (PDN) applications. The Vth of E-mode and D-mode are 3.7V and -1.1V, respectively, with A_{MFM}:A_{GaN} ~ 1:940 to overcome polarization lock effect. The non-degradation of V_{BD} with ferroelectric gate stack has been validated to qualify for 12V operation of PDN applications. The robust stability has been validated for ΔVth =3V after 10^9 disturb cycles. The co-integration of E-mode and D-mode GaN FeMHEMTs to realize a DC-DC converter for a voltage regulator for PDN in a heterogenous 3D integrated system can provide high efficiency and minimize the heat issue.

4:30 PM

38-7 Sub-micron AlN Films with High Thermal Conductivity and Electrical Resistivity for Efficient Heat Dissipation in 3D ICs

Sam Vaziri[1], Xiaheng Huang[2], Songrui Hou[2], Isha Datye[2], Chengxing He[2], Ji-Feng Ying[2], Xiangjin Wu[3], Chen-Feng Hsu[2], Luke Gyubin Min[3], Heungdong Kwon[3], Sarah Chen[2], Chao-Ching Chiang[2], Wei-Yen Woon[2], Szuya Liao[2], Mehdi Asheghi[3], Kenneth Goodson[3], Eric Pop[3], H.-S. Philip Wong[3], Xinyu Bao[2]
[1]Corporate Research, Taiwan Semiconductor Manufacturing Company, [2]Taiwan Semiconductor Manufacturing Company, [3]Stanford University

Poly-AlN films are explored as a promising scalable thermal material for 3D ICs. Sub-500 nm thick poly-AlN films, PVD sputter-deposited below 200°C on various substrates, achieve exceptional cross-plane and in-plane thermal conductivities up to 100 $Wm^{-1}K^{-1}$ and 30 $Wm^{-1}K^{-1}$, respectively. Systematic optimization of crystallinity and mitigation of oxygen-related defects are key to achieving high-performance AlN. The improved AlN films exhibit high electrical resistivities >1014 Ω·cm, comparable to those of PECVD-deposited SiO_2. Additionally, ultra-smooth surface roughness of sub-5 Å is achieved, ensuring integration compatibility. Thermal simulations demonstrate that poly-AlN reduces hotspot temperatures by ~15% compared to SiO_2 in a two-tier structure, establishing it as a promising material for thermal management in advanced 3D ICs.

4:55 PM

38-8 Thermal Mitigation Strategy for 3D Stacked Devices: Leveraging Dense BN Low-k Dielectrics with Improved Thermal Conductivity

Taehoon Kim[1], Jaewon Kim[2], Jae Hong Lee[2], Yoonhoo Ha[2], Woochang Lee[2], Sangjun Lee[2], Heegoo Kim[2], Jaeyoung Hwang[3], Wonsik Lee[3], Hyejin Jang[3], Chang-Seok Lee[2]
[1]Device Research Center, Samsung Advanced Institute of Technology, Samsung Electronics, [2]Samsung Electronics, [3]Seoul National University

Rapid advancements in semiconductors—such as line miniaturization, 3D architectures, and chip stacking—are pushing thermal management challenges to new limits, necessitating innovative materials, designs, and techniques to mitigate heat-buildup in devices. A dense boron nitride (BN) dielectric with a 55-times higher thermal conductivity and a 1.63-times higher density than conventional p-SiCOH materials is presented. Using multi-scale simulations based on the complementary FET (CFET) device structure, potential device temperature

reductions up to 17.3°C is demonstrated. Combined with advanced packaging cooling solutions, this innovation marks a step forward toward effective heat regulation for more efficient, reliable, and high-performing electronics.

Session 39: MT | MRAM and RRAM

1:30 PM, Continental 1-3
Co-Chairs: Johannes Mueller, GlobalFoundries and Kai-Shin Li, Taiwan Semiconductor Research Institute
This session brings together eight cutting-edge papers showcasing advancements in MRAM and RRAM technologies. The first four papers focus on MRAM, followed by four contributions on RRAM. Samsung opens the session with two significant contributions: Automotive-grade eMRAM at the 8nm node, demonstrating enhanced reliability for harsh environments and magnetic immunity improvements through bit-level magnetic shielding, addressing a key reliability concern. The third paper, from the IBM–Samsung MRAM Alliance, introduces efficiency gains in double spin-torque MTJs, pushing the limits of performance. The MRAM section concludes with TSMC's breakthrough in field-free switching for high-endurance, high-retention c-type SOT devices. The RRAM segment begins with a collaboration between NTHU, DTP, and TSMC, presenting low-mask-adder RRAM integration at the 5nm node. Next, Peking University, YanXin MicroElectronics, and ICEM tackle the access device bottleneck by embedding BJT functionality into RRAM arrays. National Yang Ming Chiao Tung University follows with a compact 2T16R bit-cell, achieving high compute-in-memory (CiM) performance. The session concludes with POSTECH's innovative hybrid memory, combining ferroelectricity and conductive-bridging in an FTJ/CBRAM device.

1:35 PM
39-1 World First 8nm 128Mb Embedded STT-MRAM for Automotive Application
Hyunseok Whang[1], Seungpil Ko[1], Jaechul Shim[1], Woochang Lim[1], Junho Park[1], Tae Young Lee[2], Geonhee Bae[1], Chanhyeong Lee[1], Kyounghun Ryu[1], Suho Choi[1], Baeseong Kwon[2], Jae Wook Lee[1], Yoonsung Han[1], Jong Hyuk Lee[1], Hyun-chang Lee[2], Sohee Hwang[2], Kiseok Suh[2], Shinhee Han[2], Gyuseong Kang[2], Hyun Jin Shin[2], Jonghoon Jung[2], Cheol Kim[1], Joon Myoung Lee[1], Young Hyun Kim[1], Jeong-Heon Park[1], Kangho Lee[2], Sangjin Hyun[1], Yoonjong Song[1], Su Jin Ahn[1], Jaihyuk Song[1]
[1]R&D center, Samsung Electronics Co, [2]Foundry Business, Samsung Electronics Co.
Here we demonstrate an 8nm logic compatible 128Mb embedded magnetic random access memory (eMRAM) for automotive applications. We successfully scaled down the eMRAM to 0.017μm^2 unit cell. With advanced integration process, we achieved sub-ppm level of short failure and ppm-level of read failure. With optimized pinned layer and advanced magnetic tunnel junction (MTJ) stack engineering, we achieved sub-ppm level of write error rate (WER) while maintaining the retention failure rate substantially below the criteria required for reliable automotive applications. We demonstrate the full functionality of write and read operation over the temperature range from -40℃ to 150℃.

2:00 PM
39-2 Wafer-level Magnetic Shielding for Highly Reliable STT-MRAM with Enhanced Magnetic Immunity
Jeong-Heon Park, Jae Hoon Kim, Dmytro Apalkov, Myoungkwan Chae, Whan Kyun Kim, Jeong Mok Kim, Taewoo Kim, Cheol Kim, Younghyun Kim, Hyunsung Jung, Seungpil Ko, Kiseok Suh, Shin Hee Han, Kangho Lee, Yoon Jong Song, Su Jin Ahn, Seung Hun Lee, Sangjin Hyun, Jai Hyuk Song
Samsung Electronics
Magnetic immunity remains a concerning barrier to the adoption of STT-MRAM in IoT and edge applications, where devices often operate in unshielded and unpredictable environments. We propose a wafer-level magnetic shielding approach that overcomes the inherent limitations of conventional package-level shields. By integrating a soft magnetic material directly into the backend-of-line (BEOL) process, our method offers superior proximity-based protection against external magnetic fields.

2:25 PM
39-3 Progress and Gaps in Double Spin-torque MTJs for Last Level Cache Applications
Guohan Hu, Philip Trouilloud, Christopher Safranski, Stephen Brown, Matthias Gottwald, Cesar Heinsohn, Gukcheon Kim, Junhoe Kim, Laura Lehmann, Jim Liang, Michael Robbins, Jonathan Sun, Pouya Hashemi, Daniel

Worledge
IBM T J Watson Research Center
We demonstrate 25% efficiency improvement in double spin-torque magnetic tunnel junctions (DS-MTJs) compared to standard single-MTJs, without degrading other properties. Optimized stacks enabled reliable setting of the two reference layers into the desired anti-parallel alignment for a 4-kbit DS-MTJ array, where switching of all 4k devices with 2 ns write pulses was demonstrated for the first time. Activation energy ~ 60 kT at CD ~ 40 nm was achieved with new free layer designs. Devices with optimized DS-MTJ free layers also exhibit steep write-error-rate (WER) slopes, comparable to those of the best single-MTJs reported to date.

2:50 PM

39-4 Novel Field Free Type-C SOT-MRAM Devices with Circular MTJ and 8kb Array Demonstration

MingYuan Song[1], Chen-Yu Hu[2], Chiao-Yun Lo[3], Ho-Lin Tsai[3], I-Jung Wang[3], Guan-Long Chen[3], Chen-Feng Hsu[2], Hsin-Yu Chen[2], Yao-Jen Chang[3], Shan-Yi Yang[3], Yu-Chen Hsin[3], Yuan-Jen Lee[2], Jeng-Hua Wei[3], Shih-Chieh Chang[3], Xinyu Bao[2]
[1]CR, TSMC, [2]TSMC, [3]ITRI

For the first time, a novel type-C SOT-MRAM device is demonstrated, featuring a circular-shaped MTJ with built-in magnetic anisotropy. The type-C SOT-MRAM exhibits field-free operation while maintaining strong data retention and excellent performance. The high-speed write capabilities with pulse durations as short as 1 ns are demonstrated, along with excellent thermal stability (>70k_BT), validated across multiple devices and an 8kb memory array. Furthermore, the Type-C device offers a 48% reduction in cell size and a 25% decrease in switching current, presenting a highly compelling solution for emerging applications that require cost-effective, high-performance non-volatile memories.

3:15 PM

39-5 Ultra High Density 1D1R FinFET Dielectric RRAM in 5nm FinFET CMOS Logic Process

Wei-Hwa Lin[1], Pin-Hsuan Shih[2], Tzu-Ying Sung[2], Yue-Der Chih[3], Yih Wang[3], Jonathan Chang[3], Ya-Chin King[2,2], Chrong Jung Lin[2,2]
[1]Institute of Electronics Engineering, National Tsing Hua University (NTHU), [2]National Tsing Hua University (NTHU), [3]Taiwan Semiconductor Manufacturing Company (TSMC)

An ultra-high-density 5nm FinFET Dielectric (FinD RRAM) RRAM with 1D1R cross-point array is firstly developed with fully compatible 5nm FinFET CMOS Logic process . Also, FinFET Schottky diode is innovated and built-in FinD RRAM to achieve cross-point array by 1D1R-structure with small cell size of 1/2 FinFET or 0.02μm^2. The FinD RRAM features ultra-small cell size in high-density memory, 5nm FinFET process compatibility for manufacturing, fast resistive switching for HPC applications, low-operation voltage for SoC , and superior reliability. The FinD RRAM can be adapted to 3nm FinFET process and beyond, making it a promising RRAM solution in advanced embedded memory applications.

4:05 PM

39-6 High-Density RRAM for Advanced Logic Process: A Hybrid-Driven Cell with Self-Aligned Isolation Scalable to FinFET Technology

Jingwei Sun[1], Zongwei Wang[2,3], Zhixing Cai[2], Shengyu Bao[2], Zimeng Wu[2], Yanbang Chu[2], Heng Wu[2], Yingchen Ji[2], Ruiqing Xie[2], Lin Bao[2], Zheng Zhou[2], Ling Liang[2], Zhizhen Yu[4], Gaoming Feng[4], Ao Guo[4], Yifei Lu[4], Chen Li[4], Shoumian Chen[4], Yuhang Zhao[4], Yimao Cai[2], Ru Huang[2]
[1]School of Integrated Circuits, Peking University, [2]Peking University, [3]YanXin MicroElectronics Co., Ltd., [4]ICEM

RRAM embedded in logic process faces a trade-off between density and reliability due to the mismatch between the switching voltage/current requirements of the RRAM and the driving capability of the access transistor. In this study, we demonstrate a hybrid-driven (HYDRV) RRAM cell that leverages in-cell parasitic BJTs to boost RESET current, enabling aggressive transistor down-scaling. Further, a self-aligned deep-trench isolation (DTI) process is introduced to minimize inter-cell pitch and enhance electrical isolation. The proposed HYDRV cell is silicon-verified at 28nm, and FinFET node, achieving record-density of 36.76Mb/mm^2 and 47.62Mb/mm^2, respectively. Despite aggressive scaling, the HYDRV cell exhibits excellent reliability, including endurance >10^5 cycles and

retention >10years@125℃. This work presents a scalable and robust RRAM design suitable for high-density, high-reliability embedded memory in advanced CMOS technologies.

4:30 PM

39-7 A Novel High-Density and Stackable 3D Vertical RRAM-Based CIM Macro with 72.9-TOPS/W Energy-Efficiency for AI Inference Accelerator

P. H. Shih, Y. T. Yang, M. J. Huang, Y. R. Liu, Y. K. Wang, J. C. Guo, Steve Chung

National Yang Ming Chiao Tung University

This work demonstrates a 1kb 3D vertical RRAM-based Compute-In-Memory (CIM) macro utilizing a novel 2TnR (n=16) unit cell. Vertically-stacked MIM is forming free with low SET/RESET operating voltages and excellent HRS/LRS ratio (>10^6). The 2T16R structure is non-destructive in read with 1T for programming and 1T for read, reducing the operation voltage, while retaining a large memory window. The CIM macro achieves high-speed operation (140MHz) with stable 3-bit-per-cell operations. This CIM macro supports AI inference with 97.45% and 84.28% accuracy on MNIST and CIFAR-10 datasets, respectively. It also achieved 72.9TOPS/W energy efficiency, providing a promising solution for neural-network applications.

4:55 PM

39-8 Atomic-Gap Polarization Switching in 3D-Stackable FTJ/CBRAM Hybrid Memory for High-Performance, Reliable Multi-bit Storage

Jiho Kim[1], Laeyong Jung[2], Jongseon Seo[2], Hyunsang Hwang[2]

[1]Material Science and Engineering, Pohang University of Science and Technology (POSTECH), [2]Pohang University of Science and Technology (POSTECH)

We propose a hybrid memory device that combines a ferroelectric tunnel junction (FTJ) with a sub-G_0-operating conductive-bridge RAM (CBRAM), to address the critical limitations of each device individually. By precisely controlling filament growth, we realized polarization switching within an atomic-scale (~1 nm) tunnel gap, dramatically enhancing device performance. The hybrid device demonstrated a significantly improved on/off ratio (>10^5) and J_{on} (>10^4 A/cm^2) compared to conventional FTJ. Moreover, unlike conventional CBRAM, hybrid device achieved highly uniform switching states and superior retention (>10 years at RT), enabled by the internal polarization field from ferroelectric HZO. In addition, 3-bit operation and robust endurance (>10^9 cycles) were achieved via uniform, high-contrast switching involving only a few Cu atoms. Reliable switching in a 3-layer vertically stacked structure validated the potential for high-density 3D integration.

Session 40: RSD | Advanced Transistor Reliability

1:30 PM, Continental 4

Co-Chairs: Runsheng Wang, Peking University and Stephen Ramey, Intel

This session includes 8 papers that describe the advances of reliability in silicon and oxide semiconductor transistors. The first 4 papers cover silicon-based devices. The first, from IMEC, demonstrates the significant NBTI improvement potential of novel anneal conditions especially relevant for low thermal budget process flows. The second paper is from the National Tsing Hua University and TSMC detailing the use of a new recording device for evaluating in-process plasma effects in a 5nm CMOS technology. The third paper, from IMEC, details a new experimental method for extraction of the leakage currents from floating body effects in GAA transistors which is important for DRAM. The final paper in the silicon group, from Applied Materials and the Indian Institute of Technology, describes GAA process integration modulation for channel strain with implications for NBTI. The next 4 papers cover oxide-based semiconductors (OS). The first paper in the OS group is from IMEC and proposes a novel annealing scheme for IGZO transistors, which enables the hydrogen resilience needed to meet the reliability specs for embedded DRAM applications. The second paper, from Georgia Tech and Samsung, presents an analysis framework for Vth degradation due to DC/AC BTI in IGO transistors with on-the-fly measurement and new modeling. The third paper, from National University of Singapore and collaborators, studies the degradation due to DC/AC BTI in ITO transistors, moving beyond conventional Vth analysis to include Gm and Rc dynamics. The last paper, from Peking University, proposes a hydrogen evacuation layer to enable hydrogenation immunity in IGO transistors with improved BTI reliability.

1:35 PM

40-1 Novel Low Temperature SiO_2 Formation Process by Oxygen and Hydrogen Radicals for Core and I/O RMG Stacks: Achieving the Ultimate NBTI Reliability with a Charge-Free IL

Jacopo Franco[1], Hiroaki Arimura[1], Joao Pedro Bastos[1], Valeri Afanasiev[1,2], Jean-Francois de Marneffe[1], Min-Soo Kim[1], Ben Kaczer[1], Naoto Horiguchi[1]
[1]imec, [2]KU Leuven

Low thermal budget RMG integration is required for future CMOS architectures. We demonstrate a low-temperature (450°C) plasma-based SiO_2 formation process by oxygen and hydrogen radicals yielding a low-defectivity, fixed charge-free interfacial layer, and outmatching conventional SiO_2 formation processes, including 900°C dry oxidation. IL thickness control is demonstrated for Core and I/O gate stacks, which would not be possible with self-limited chemical oxidation. The novel IL enables ultimate NBTI reliability due to suppressed hole trapping and minimized interface state generation, without requiring any post-growth treatment. The study sheds light on the impact of SiO_2 positive fixed charge (trapped protons) on NBTI.

2:00 PM

40-2 5nm FinFET Differential Plasma Recorder Device

Wei Chang[1], Chieh-Hsin Lin[2], Ching-Hsing Fang[2], Tzu-Hao Yu[2], Ching-Yun Chang[3], Wun-Jie Lin[3], Jam-Wem Lee[3], Ko-Ji Chen[3], Ming-Hsiang Song[3], Yue-Der Chih[3], Yih Wang[3], Jonathan Chang[3], Ya-Chin King[2], Chrong Jung Lin[2]
[1]Institute of Electronics Engineering, National Tsing Hua University, [2]National Tsing Hua University, [3]Taiwan Semiconductor Manufacturing Company

A newly developed FinFET-based wafer-level differential plasma charge recorder fabricated using a 5nm FinFET CMOS logic process is first presented in this paper. The device enables precise monitoring of charge polarization and quantification of plasma-induced charges generated in the 5nm FinFET BEOL dual damascene metallization processes. A high-resolution differential recorder device with differential pair structure has been developed and utilized to accurately record the transient charging behavior, distribution of plasma levels, polarization of charges, and interference in adjacent wells, and charging effect from BEOL Cu corrosion during dynamic BEOL process steps across the entire 12-inch wafers. The promising recording device offers a reliable and precise method for wafer-level plasma-induced charge monitoring in 5nm FinFET technology and beyond.

2:25 PM

40-3 First Systematic Characterization of Floating Body Effects in GAA Nanosheet 3D DRAM Access Transistors

Daniele Garbin[1], Geert Eneman[1], Romain Ritzenthaler[1], Nouredine Rassoul[1], Pierre Eyben[1], Eren Canga[1], Daisuke Matsubayashi[1], Andrea Fantini[1], Barry O'Sullivan[1], Loris Angelo Labbate[1], Jana Loyo Prado[1], Roger Loo[1,2], Wouter Devulder[1], Emmanuel Dupuy[1], Tobias Peissker[1], Antoine Pacco[1], Hemant Kumar Raut[1], Alexey Milenin[1], Lennaert Wouters[1], Evi Vrancken[1], Matteo Beggiato[1], Erik Rosseel[1], Gouri Sankar Kar[1], Attilio Belmonte[1]
[1]imec, [2]Ghent University

As DRAM scaling moves toward dense 3D architectures with Si-based select transistors, floating body effects (FBE) raise reliability concerns due to transient leakage and retention risks. We present the first experimental extraction of floating-body-induced bipolar current (FBC) in nanosheet GAA 3D DRAM access transistors, enabled by large arrays. Two device populations, differing in quasi-static leakage, are analyzed using a protocol isolating transient BJT currents. FBC amplitude scales with pre-charge time and leakage, consistent with TCAD. Extrapolation toward target I_{OFF} suggests FBE suppression in optimized devices, providing reliability guidance for future 3D DRAM integration

2:50 PM

40-4 Reliability-Aware Front-End-of-Line & Circuit Design Considerations for Next Generation Gate-All-Around Technology

Pratik B Vyas[1], Megha S[2], Lars P. Tatum[1], Maasil Ahmad Wani[3], Arnav Shaurya Bisht[3], Raj Sambhav[3], Mainak Saha[3], Benjamin Colombeau[1], Ashish Pal[1], Souvik Mahapatra[3], El Mehdi Bazizi[1]
[1]Applied Materials Inc., [2]Indian Institute of Technology Bombay, [3]Indian Institute of Technology, Bombay

As logic technology transitions from a FinFET to a Gate All-Around (GAA) transistor architecture, device design considerations become critical both from a nominal as well as end-of-life performance perspective. In this paper, we compare and quantify the nominal and NBTI (Negative Bias Temperature Instability) reliability performance evolution going from FinFET to GAA architecture.

3:40 PM

40-5 Improving IGZO transistor PBTI reliability with hydrogen anneal achieving overdrive voltage of 1.2V for a lifetime of 5 years at 95°C

Adrian Chasin[1], Jacopo Franco[1], Daisuke Matsubayashi[1], Stanislav Tyaginov[1], Michiel Van Setten[1], Bart De Wachter[1], Harold Dekkers[1], Anastasiia Kruv[1], Pietro Rinaudo[1,2], Ying Zhao[1], Luca Panarella[1,2], Valeri Afanasiev[1,2], Alexandru Pavel[1], Yiqun Wan[1], Subhali Subhechha[1], Attilio Belmonte[1], Ben Kaczer[1], Gouri Sankar Kar[1]
[1]imec, [2]KULeuven

We propose a novel annealing scheme for IGZObased transistors, which enables the hydrogen resilience needed to meet the reliability specs for embedded DRAM applications: an overdrive voltage of 1.2V for a lifetime of 5 years at 95°C. First, we show that standard IGZO compositions (1:1:1 at%) and In-poor (In~10%) suffer from opposite reliability limitations: whereas standard IGZO only meets the reliability target for the electron trapping component, In-poor only meets the reliability target for the H-doping process. Second, we describe how a combination of hydrogen (FGA) and oxygen anneals reduces the H-doping process during PBTI stress by ~30x, without performance penalty.

4:05 PM

40-6 DC and AC PBTI / NBTI Analysis for BEOL Oxide FETs via On-the-Fly Measurement and Modeling

Hyun Jae Lee[1], Hyeonwoo Park[2], Chengyang Zhang[2], Jaewon Shin[2], Eknath Sarkar[2], Hwan Kim[3], Changik Im[3], Min Ji Hong[3], Daewon Ha[3], Shimeng Yu[2], Asif Islam Khan[2], Suman Datta[2]
[1]Electrical and Computer Engineering, Georgia Institute of Technology, [2]Georgia Institute of Technology, [3]Samsung Electronics

We present the first comprehensive framework for AC BTI lifetime projection in amorphous oxide semiconductor (AOS)-channel devices. Our approach combines: (i) On-the-fly kHz-regime measurement of stress/relax transients; (ii) An analytical AC bias-temperature-instability (BTI) model based on a stretched-exponential form originally developed for DC BTI; and (iii) closed-form extrapolation to year-scale predictions under realistic access scheme. Projection with DRAM-relevant conditions (400 MHz, Duty factor = 0.3), a ΔV_{th} of +6.3 mV at 1 kilo-second PBTI results show that BTI can be successfully cared. Moreover, the model reveals which components dominate degradation across time scales, offering actionable insights for AOS device design.

4:30 PM

40-7 Beyond ΔV_{th}: Unlock New Insights into Hydrogen-Induced BTI in ITO FETs through G_m and R_C Dynamics

Gan Liu[1], Qiwen Kong[2], Chen Sun[2], Dong Zhang[2], Leming Jiao[2], Yuxuan Wang[2], Bich-Yen Nguyen[3], Xiao Gong[2,4]
[1]ECE, National University of Singapore, [2]National University of Singapore, [3]Soitec, [4]Agency for Science, Technology and Research (A*STAR)

This study delves into the hydrogen (H)-related instabilities in ITO FETs under DC and AC BTI conditions, spanning a wide range of L_{CH}. Moving beyond conventional V_{th} analysis, we unveil, for the first time, the dynamic changes in transconductance (G_m) and contact resistance (R_C) during stress, highlighting their dependence on downscaling L_{CH}. Through *in situ* R_C extraction, we uncover that R_C changes exhibit distinct behaviors under PBTI and NBTI conditions. Based on the above observations, we propose two dominant mechanisms driving these H-related phenomena under different BTI scenarios: H drift governing DC stress and H release governing AC stress.

4:55 PM

40-8 Hydrogenation-Immune Oxide Semiconductor FETs via Hetero-Integration-Compatible Hydrogen Evacuation Layer

Jingyu Fan[1], Zhendong Jiang[2], Jinwen Liu[2], Mengran Liu[2], Yuhan Zhang[2], Shengjie Yang[2], Zhihao Lv[2], Xin Zheng[2], Jinxiong Li[2], Qian Li[2], Kuanchang Chang[2], Xinwei Wang[2], Zeyu Cai[2], Shengdong Zhang[2], Meng Zhang[3], Yan Yan[3], Guijun Li[3], Lei Lu[2]

[1]School of Electronics and Computer Engineering(SECE), Peking University, [2]Peking University, [3]Shenzhen University
The hetero-integration of oxide semiconductors (OSs) and silicon (e.g., poly-Si and crystalline-Si) exhibits promising potential on advancing displays and 3DICs, while the inevitable hydrogen (H) doping results in a critical process incompatibility. In contrast to incumbent mitigation solutions to the H-induced micro-scale short-channel effect (SCE) and severe bias-temperature instabilities in OS field-effect transistors (FETs), a novel hetero-integration-compatible Hydrogen Evacuation Layer (HEL) is proposed to enable the thorough hydrogenation-immunity in OS FETs. The H-induced SCE is suppressed with OS channel length downscaling to 60 nm, and positive/negative bias temperature instabilities (PBTI/NBTI) and self-heating instability (SHI) are reduced to 0.03 V and 0.3 V, respectively. This paves the way for Miro- and Nano-scale OS-Si hetero-integration.

Session 41: ALT/ODI | Focus Session | Silicon Photonics for Energy Efficient AI Computing

1:30 PM, Continental 5
Co-Chairs: Matteo Buffolo, University of Padova and Byounghak Hong, Samsung Electronics
This session brings together papers highlighting the progress of silicon photonics and heterogeneous integration for next-generation optical interconnects, covering polymer and Si waveguide-based CPO modules, modulators, photodetectors and integrated membrane lasers/EAMs, while also exploring optical switching, advanced optical routing approaches, integrated GeSnOI lasers and large-scale scalable photonic circuits for sustainable AI systems and enhanced connectivity.

1:35 PM
41-1 Photonic Interconnect for Next-Generation AI Systems (Invited)
Benjamin G Lee
Research, NVIDIA
AI systems employ tightly integrated GPU clusters using low-cost high-reliability intra-cluster electrical signaling. Reach ultimately limits cluster size and aggravates thermal density. Efficient, low-cost, distance-insensitive off-chip communication can alleviate thermal challenges, increase cluster size, and simplify packaging. Photonics can help, but it must meet stringent efficiency, cost, and density targets.

2:00 PM
41-2 Co-packaged Optics Module with Single-Mode Polymer Waveguide (Invited)
Akihiro Horibe, Yoichi Taira, Sayuri Kohara, Marushima Chinami, Hiroyuki Mori, Hidetoshi Numata, Jean Benoit Héroux, Adrian Paz Ramos, Hsianghan Hsu, Neng Liu, Mark Schultz, John Knickerbocker
IBM Research
We present co-packaged optics (CPO) modules featuring single-mode polymer waveguide (PWG) fiber interfaces to address the growing demand for high-bandwidth, energy-efficient interconnects in AI systems. Compared to conventional fiber connections, PWGs offer higher optical lane density and easier assembly. We evaluated two substrate configurations and conducted mechanical and thermomechanical reliability testing. Results showed stable adhesion, low insertion loss (<1.2–2.0 dB), and resilience under thermal stress. A 3-port CPO module with a cut-out substrate demonstrated the feasibility of optics-last assembly. These findings support PWG-integrated CPO modules as a promising solution for future heterogeneous chiplet-based AI hardware platforms.

2:25 PM
41-3 Silicon Photonics Modulators and Photodetectors for Next-Generation CPO and Optical I/O (Invited)
Joris Van Campenhout[1], Minkyu Kim[1], Conor Coughlan[1], Amir Shahin[1], Zohauddin Ahmad[1], Cedric Bruynsteen[2], Leandro Da Silva[2], Nishant Singh[2], Dharmander Malik[1], Grigorij Muliuk[1], Didit Yudistira[1], Javad Rahimi Vaskasi[1], Shalini Jakanadan[1], Guy Lepage[1], Chiara Marchese[1,1], Shuchi Kaushik[1], Roger Loo[1], Yosuke Shimura[1], Hakim Kobbi[1], Rafal Magdziak[1], Huseyin Sar[1], Dieter Bode[1], Swetanshu Bipul[1], Patrick Carolan[1], Marta Agati[1], Dries Van Thourhout[3], Peter Verheyen[1], Maumita Chakrabarti[1], Peter De Heyn[1], Peter Ossieur[1], Filippo Ferraro[1], Yoojin Ban[1]
[1]IMEC, [2]imec-Ghent University, [3]Ghent University - imec
We review the recent performance of high-speed, high-efficiency electro-optical modulators and photo-detectors, targeting next-generation co-packaged optics (CPO) operating at 400G/lane and future Optical I/O (OIO),

employing a “wide-and-slow” parallel interface. We focus on compact devices, including Si ring and disk modulators, as well as electro-absorption modulators implemented in Ge(Si) or through heterogeneous InP-on-Si integration. Finally, we provide a performance overview of recently developed Si-integrated Ge photodetectors, capable of supporting both next-generation CPO and Optical I/O.

2:50 PM

41-4 Heterogeneously Integrated Membrane Lasers and EAMs on Si Platform for Energy-Efficient Optical Link (Invited)

Shinji Matsuo, Tatsurou Hiraki, Koji Takeda, Takuro Fujii, Takuma Aihara, Yoshiho Maeda, Tomonari Sato
NTT, Inc.

We demonstrate a photonic integration platform that uses two complementary techniques: micro-transfer printing (MTP) and wafer-scale fabrication involving the epitaxial growth of compound semiconductors. The platform utilizes InP-based membrane distributed feedback (DFB) lasers, and electro-absorption modulators (EAMs) integrated on silicon wafers. In the first approach, MTP enables selective placement of independently optimized III-V coupons onto a Si photonic circuit, demonstrating 128-Gbit/s NRZ data transmission. In the second approach, a wafer-scale process is employed to realize an optical link composed of membrane laser, EAM, and photodetector (PD) on a Si wafer interconnected via low-loss SiOx waveguides. This configuration achieves an energy cost as low as 0.14 pJ/bit at 50-Gbit/s NRZ signal. These results demonstrate that membrane-based photonic integration offers a scalable and energy-efficient solution for both inter-package and intra-package optical interconnects.

3:40 PM

41-5 What's Next for Optical Switching? (Invited)

George Papen, Cedric Fung Lam, Hong Liu, William (Max) Mellette, Ryohei Urata, Lieven Verslegers, Kevin Yasumura
Google

We discuss device technologies for future optical circuit switches focusing on large scale system applications such as data center networks and machine learning supercomputers. Important device parameters including insertion loss, crosstalk, port count, reconfiguration time, polarization sensitivity etc. can affect the final system performance and reliability.

4:05 PM

41-6 Large-scale Silicon Photonic Integrated Circuits for Sustainable AI Infrastructure (Invited)

Guangwei Cong[1], Shu Namiki[2]
[1]Photonics-Electronics Integration Research Center, The National Institute of Advanced Industrial Science and Technology (AIST), [2]The National Institute of Advanced Industrial Science and Technology (AIST)

This talk will discuss how silicon photonics can inherently resolve the current bottlenecks not just interconnects but switches and even compute and argue that the integration of photonics and electronics under the holistic optimization of architecture will be the key.

4:30 PM

41-7 Scalable Integrated Photonics for Enhanced Connectivity (Invited)

Anna Tauke-Pedretti[1], Radoslav Bogoslovov[2], Amil Patel[3], Chelsea Haughn[3], Tyler Michalak[4]
[1]DARPA, [2]Bogoslovov Consulting, [3]Booz Allen Hamilton, [4]Booz Allen Hamiliton

Advanced integrated photonics solutions significantly improve connectivity density for energy-efficient computing, sensing and data transmission architectures. By enhancing three-dimensional optical routing and robust chip-to-chip connections within the microsystem, information transmission can be optimized, enabling scalable architectures that reduce energy consumption while supporting complex AI (artificial intelligence) applications.

4:55 PM

41-8 GeSnOI Lasers for Laser-Integrated Photonic Chips (Invited)

Donguk Nam[1], Gwangmin Yu[1], Hyo-Jun Joo[2], Jiawen Liu[3], Melvina Chen[2], Youngmin Kim[2], Simone Assali[4], Carlo Sirtori[3], Yanko Todorov[3], Oussama Moutanabbir[4]
[1]KAIST, [2]Nanyang Technological University, [3]École Normale Supérieure, [4]École Polytechnique de Montréal

The explosive growth of artificial intelligence and machine learning applications has significantly increased the demand for high-speed and energy-efficient data communication within data centers. Optical interconnects, particularly co-packaged optics, have emerged as promising solutions due to their capability to overcome limitations of traditional electrical interconnects. However, the absence of efficient CMOS-compatible on-chip lasers remains a major bottleneck. In this paper, we discuss recent advancements in GeSn-based lasers as a viable candidate to address this challenge.

Session 42: ODI | Advanced Image Sensors

1:30 PM, Continental 6

Co-Chairs: Geunsook Park, Omnivision and Daniel Van Blerkom, SWIRLabs

This session includes 8 papers on the latest in image sensor technology. The first is an invited paper on progress in flash LiDAR using heterodyne detection. The next two papers present HDR imagers using LOFIC pixels that reach or exceed 120dB. This is followed by papers on specialty imagers, the first one describing an all-organic flexible imager, followed by an extremely high frame-rate burst CIS sensor. The final three papers of the session cover the latest technologies for shrinking pixels to sub-micron dimensions. Of special note is the last paper, which shrinks the dual-photodiode pixel to 0.7um.

1:35 PM

42-1 3D FMCW Wide-angle Flash Lidar: Towards System Integration and In-pixel Frequency Measurement (Invited)

Laurent Frey, Anis Daami, Josep Segura Puchades, Florence Rigal, Vincent Moulin
CEA-Leti

The Frequency Modulated Continuous Wave Light detection and ranging (FMCW Lidar) usually scans the scene point by point, and measures the distance by a Fourier Transform (FT) of the heterodyne signal. A promising solution for higher frame rate with larger image resolution is the flash version of the FMCW Lidar, using floodlight illumination and an image sensor. We first review our recent achievements of FMCW flash Lidar system developments with commercially available components and post-processing. Because FT is difficult to implement in small pixels, we then introduce the principle of a heterodyne image sensor with in-pixel frequency measurement combined with a multi-chirp laser modulation strategy, targeting video rate real-time measurements.

2:00 PM

42-2 A 120 dB Dynamic Range 3D Stacked 2-Stage LOFIC CMOS Image Sensor with Illuminance-Adaptive Signal Selection Function

Kohei Takizawa[1], Yoshihito Hirai[2], Masaya Yoshida[2], Takezo Mawaki[2,2], Ken Miyauchi[2], Rihito Kuroda[2,2]
[1]Graduate School of Engineering, Tohoku University, [2]Tohoku University

This work presents a 3D stacked 2-stage lateral overflow integration capacitor (LOFIC) CMOS image sensor with illumination adaptive signal selection function. To reduce the high data rates of conventional wide dynamic range sensors, this work proposes an illuminance-adaptive signal selection circuit that non-destructively determines the light intensity level using electrons accumulated in the first stage of LOFIC. This allows the developed sensor to selectively output one or two most appropriate signals out of three, reducing the data rate while maintaining a wide dynamic range. Furthermore, a 3D-stacked Si trench capacitor is employed to achieve over 8.6Me^- FWC with 5.6 μm pixel pitch. The fabricated chip demonstrates a dynamic range of 120 dB with selected signal readout and a maximum SNR of 67.5 dB.

2:25 PM

42-3 A 129 dB Dynamic Range Triple Readout CMOS Image Sensor with FWC Enhancement Technology

Takaya Yamanaka[1], Ryosuke Nakamura[1], Shunsuke Kasashima[1], Chie Tokumitsu[1], Shuto Kuwahara[1], Masatoshi Kitajima[1], Rikuya Matano[1], Takeshi Ishizaki[1], Makoto Aoki[1], Naoshi Yagi[1], Hiroaki Iga[1], Shohei Nabeyoshi[2], Masashi Takami[2], Yorito Sakano[1], Yusuke Oike[1]

[1]Sony Semiconductor Solutions Corporation, [2]Sony Semiconductor Manufacturing Corporation

We present a 2.1 µm pixel CMOS image sensor for automotive applications achieving 129 dB single exposure dynamic range with triple readout. The advanced sub-pixel architecture incorporates FDTI, conformal doping and 3D-MIM technologies, significantly enhancing full-well capacity. The sensor enables a seamless triple-image composition with 29 dB SNR at connection points, suitable for high-temperature automotive environments.

3:15 PM

42-4 Flexible 256×256 All-Organic-Transistor Active-Matrix Optical Imager with Integrated Gate Driver

Weihong Yang[1], Jun Li[2], Peijin Huang[2], Simon Ogier[3], Xiaojun Guo[2]

[1]Department of Electronic Engineering, Shanghai Jiao Tong University, [2]Shanghai Jiao Tong University, [3]SmartKem Ltd.

Solution-processed organic thin film transistors (OTFTs) provide a promising platform for truly flexible, large-area integrated sensor systems. Here, an all-organic-transistor active-matrix imager using OTFTs for both the backplane and optical sensing layer is developed. Through reducing the density of states at the channel interface for a steep subthreshold swing and low dark current, the resulting organic phototransistor (OPT) presents a high detectivity of 2.2×10^{16} Jones. The OPT is stacked on top of an OTFT switch with a high ON/OFF ratio of 4.7×10^{10} to form the active matrix and the gate driver is also integrated. Finally, a 256 × 256 (213 PPI) flexible active-matrix imager is demonstrated for fingerprint and low-distortion imaging with the constructed real-time imaging system.

3:40 PM

42-5 A Global Shutter Burst CMOS Image Sensor with 6-Tpixel/s Readout Speed, 256-recording Frames and -170dB Parasitic Light Sensitivity

Masuto Kitamura[1], Yasunori Kawaguchi[1], Naoya Kuriyama[2], Tomoaki Maeda[3], Toshifumi Imamura[3], Takezo Mawaki[4], Shigetoshi Sugawa[4], Rihito Kuroda[4]

[1]Shimadzu Corp., [2]Lapis Semiconductor Co., [3]A-R-Tec Corp., [4]Tohoku University

This paper presents an ultra-high-speed (UHS) global shutter burst CMOS image sensor (CIS) featuring pixelwis analog memory arrays. The developed CIS with 628H x 480V pixels achieves a maximum frame rate of 20 Mfps and a readout speed of 6.03 Tpixel/s. A recording length of 256 frames and a parasitic light sensitivity (PLS) of -170 dB were also achieved simultaneously in a UHS camera. This low PLS is achieved through comprehensive metal shielding of the pixel circuit and memory regions, and by the spatial separation between the photodiode and memory regions, implemented using Si trench capacitors. The introduced bias adjustment circuit compensates for voltage variations among pixel positions due to the ground resistance and pixel circuit current during the pixel driving period, enabling high-resolution video recording with an effective 628H x 480V pixels and a 48 µm pitch.

4:05 PM

42-6 Silicon-on-Insulator Pixel FinFET Technology for a High Conversion Gain and Low Dark Noise 2-Layer Transistor Pixel Stacked CIS

Ryohei Takayanagi[1,2], Takashi Kamo[1], Ryosuke Yamachi[1], Mayu Sakurai[1], Hidetoshi Oishi[1], Takuya Iriguchi[1], Hiroshi Takahashi[1], Taikei Enomoto[1], Yuki Kageyama[1], Yusuke Tanaka[1], Yoshiaki Kikuchi[1], Junpei Yamamoto[2], Hideomi Kumano[2], Shinichi Yoshida[1], Toshiki Yamamoto[2], Shota Kawamoto[2], Yusuke Murakawa[2], Hiroshi Mizuno[2], Toru Nagaki[2], Hiroshi Tomita[2], Noriteru Yamada[2], Yoshiaki Kitano[1], Kazunobu Ota[1], Tomoyuki Hirano[1]

[1]Sony Semiconductor Solutions Corporation, [2]Sony Semiconductor Manufacturing Corporation

This study presents a 2-Layer transistor pixel stacked 0.8-µm dual-pixel (DP) CIS with silicon-on-insulator (SOI) fin field-effect transistor (FinFET) technology. The application of SOI FinFETs as pixel transistors, featuring a body-less configuration on buried oxide, reduces parasitic capacitance at the floating diffusion node, thereby enhancing conversion gain and noise characteristics. The SOI FinFET achieves improved transconductance and source

follower gain compared to a previous pixel FinFET. The resolution of challenges associated with the SOI structure is demonstrated through a 0.8μm DP CIS with SOI FinFETs.
4:30 PM

42-7 A 0.43μm Quad-Photodiode CMOS Image Sensor by 3-Wafer-Stacking and Dual-Backside Deep Trench Isolation Technologies
Tzu-Jui Wang, Chia-Chi Hsiao, Pi-Ju Cheng, Kuan-Chieh Huang, Wei-Cheng Hsu, Yi-Han Liao, Chi-Hsien Chung, Ming-Chieh Hsu, Keng-Yu Chou, Wen-Hau Wu, Wei-Chieh Chiang, Chi-En Chen, Shang-Fu Yeh, Meng-Hsiu Wu, Calvin Yi-Ping Chao, Chen-Jong Wang, Dun-Nian Yaung
Taiwan Semiconductor Manufacturing Company
Scaling pixel pitch below 0.5um has become highly challenging in conventional 2-wafer stacked CMOS image sensors due to the limited silicon area shared among photodiodes, photodiode-photodiode isolation, and the associated functional transistors, while maintaining the excellent pixel performance. In this work, several advanced pixel technologies, including 3-wafer stacking, dual-backside deep trench isolation, and enhanced composite metal grid, were proposed and employed, to realize the world's smallest 0.43 μm pitch quad-photodiode pixel, achieving exceptional performance metrics of full well capacity of 6000 e-, dark current of 1.3 e-/s, and read noise of 1.5 e-rms, without degradation in conversion gain.

4:55 PM
42-8 A 2-layer, 0.7μm-pitch Dual Photodiode Pixel CMOS Image Sensors with Metaphotonic Color Router
Jonghyun Go[1], KeunYeong Cho[1], ChangKyu Lee[1], Sookyoung Roh[2], Jiyoun Song[1], Jongeun Park[1], Taehoon Kim[1], Jinyoung Kim[1], Yongsang Park[1], Minkwan Kim[1], Minho Jang[1], Myunghae Seo[1], Yongkun Jo[1], Sejin Park[1], YoonSeok Kim[1], Hoocheol Lee[1], Jihyun Kwak[1], Dami Park[1], Gyunha Park[1], Sooyeon Kim[1], Jihun Lim[1], Hyuk Hur[1], Haneul Kim[1], Sungmin Son[1], HaeJung Lee[1], Jae Ho Kim[1], Hajin Lim[1], Hojin Lee[1], Sungmo Ahn[2], Seokho Yun[2], Jaihyuk Song[1]
[1]Semiconductor R&D Center, Samsung Electronics, [2]Samsung Advanced Institute of Technology
In this article, a world's smallest 0.7μm-pitch dual photodiode pixel is presented. We integrated 2-layer pixel with hybrid Cu-Cu bonding process only, without introducing pixel-level deep contacts. By optimizing layout of Cu pad layer, we suppressed capacitive coupling between neighboring floating diffusion nodes, still achieved similar conversion gain
compared to that of 0.7μm-pitch, 1-layer single photodiode pixel. We overcome the degradation of the auto-focus (AF) separation ratio by incorporating multi-focal, metaphotonic color routers (MPCR).

Session 43: SMB | Bioelectronic Devices and Systems

1:30 PM, Continental 7-9
Co-Chairs: Humberto Campanella, Silicon Austria Labs GmbH and Roozbeh Tabrizian, University of Michigan
This session includes 7 papers that describe the advances in the area of bioelectronic devices and systems. The first paper by Pan et al. demonstrates a novel NbSeVOx-based extended-gate FET for high-resolution sensing of cTnI, BNP, and NT-proBNP antibodies. This biosensor demonstrates a dynamic range exceeding 12 orders of magnitude. The second paper by Mei et al. presents a monolithically integrated array of highly-sensitive dual-gate IGZO thin-film transistors for parallel detection of multiple biomarkers. The third paper by Wang et al. demonstrates an ultra-large-scale low-temperature poly-silicon based active-matrix digital microfluidic chip. This system enables droplet actuation with a minimum volume of 1.2 nL. The fourth paper by Chaudhari et al. presents a wafer-scale fabrication of solid-state nanopores using extreme ultraviolet lithography, down to 10nm resolution. Their approach is validated through successful translocation of DNA molecules. The fifth paper by Shi et al. demonstrates a neuromorphic bio-signal processing system. The system utilizes dual-gate thin-film transistors and achieves 96% classification accuracy across four heartbeat signal classes. The sixth paper by Liu et al. presents biomolecule detection with aligned carbon nanotube FETs using a wafer-scale and high-throughput approach. The sensor shows a steep subthreshold swing less than 80 mV/dec. The seventh paper by Yao et al. demonstrates a high-performance dual-gate ion-sensitive ferroelectric FET using hafnia-zirconia on indium-oxide platform. The sensors achieves super-Nernstian sensitivity exceeding 1000 mV/PH.

1:35 PM

43-1 Ultra-Wide Range and High-Sensitivity Protein Detection for Heart Disease Using Ultrathin $NbSeVO_x$ EGFET Biosensor Arrays

Tung-Ming Pan[1,2], Wen-Hsin Chien[3], Wei-Che Weng[3], Sheng-Che Huang[3], Jim-Long Her[3]
[1]Electronics Engineering, Chang Gung University, [2]Division of Urology, Chang Gung Memorial Hospital, [3]Chang Gung University

We report the first RF co-sputtered $NbSeVO_x$ film on n^+-Si. After annealing at 600 °C, the film exhibited high pH sensitivity (63.6 mV/pH) with low hysteresis (1.1 mV) and drift (0.13 mV/h). The EGFET sensor demonstrated strong responses to cTnI (5.76 mV/pC_{cTnI}), BNP (5.01 mV/pC_{BNP}), and NT-proBNP (6.88 mV/$pC_{NT\text{-}proBNP}$) across a wide range of 10^{-6} to 10^5 pg/mL. Serum test results closely matched ELISA measurements, confirming the sensor's accuracy. This platform offers stable, noninvasive, and ultrasensitive protein detection down to ag/mL levels over 12 orders of magnitude.

2:00 PM

43-2 800 ppi 1-T Dual-Gate IGZO Ion-Sensitive Thin-Film Transistor Array with Super-Nernstian Sensitivity of 110 mV/pH for Rapid Parallel DNA Detection

Zikang Mei[1,2], Wei Tang[3,3], Jun Li[3,3], Kaiyutai Zhou[4], Ziwen Chen[3,4], Wei Jiang[4], Xiaojun Guo[3,3]
[1]Department of Electronic Engineering, Shanghai Jiao Tong University, [2]State Key Laboratory of Radio Frequency Heterogeneous Integration, Shanghai Jiao Tong University, [3]Shanghai Jiao Tong University, [4]Chinese Academy of Agricultural Sciences

This work presents a monolithically integrated highly-sensitive dual-gate (DG) indium gallium zinc oxide (IGZO) ion-sensitive thin-film transistor (ISTFT) array for rapid parallel label-free DNA detections. The DG design facilitates asymmetric capacitance coupling effect for super-Nernstian sensitivity (110 ± 20 mV/pH) while maintaining a small footprint. A large-scale (256 × 256) active-matrix ISTFT array with an ultra-high resolution of 800 pixels-per-inch (ppi) is implemented by using one ISTFT (1-T) pixel circuit, which is also integrated with gate driver on array (GOA) circuit on panel. Furthermore, a DNA biochip is demonstrated to enable on-chip parallel electrochemical detections of multiple DNA biomarkers with faster response compared to conventional LAMP and PCR methods.

2:25 PM

43-3 Ultra-large-scale Active-matrix Digital Microfluidic System Integrated with Droplet Position Sensing and AI Automatic Path Planning

Dongping Wang[1], Shengzhe Jiang[1], Hongbo Ma[1], Arokia Nathan[1,2], Jun Yu[1]
[1]Shandong University, [2]Cambridge University

This paper presents an ultra-large-scale low temperature poly-silicon (LTPS)-based active-matrix digital microfluidic (AM-DMF) chip, featuring 900 × 1200 pixels, marking the first-ever reported million-scale AM-DMF chip. Novel channel thin-film transistors (TFT) are used in pixel circuitsto suppress leakage current, thereby improving droplet actuationefficiency and enabling the minimum controllable droplet volume of 1.2 nL. A capacitance-based droplet position sensing module is co-integrated within each pixel and a priority-based A-star (PBA*) path-planning algorithm is employed to support real-time routing of large-scale droplets, establishing a robust closed-loop control framework. This work not only demonstrates the feasibility of extending TFT to non-display applications but also opens new avenues for high-throughput on-chip liquid biological samples processing.

3:15 PM

43-4 Fabrication of Solid-State Nanopores with EUV Lithography

Ashesh Ray Chaudhuri, Eric Beamish, Bert Du Bois, Pau Guell, Ayesha Walikar, Emma Vecchio, Thomas Vossen, Alexey Podkovskiy, Wouter Renckens, Wouter Botermans, Sanjin Marion, Haris Osman, Paru Deshpande, Simone Severi, Pol Van Dorpe
imec

We present the first successful demonstration of wafer-scale fabrication of solid-state nanopores using Extreme Ultraviolet (EUV) lithography. Specifically, we fabricated pores with a diameter down to approximately 10 nm within silicon nitride (Si3N4) membrane. Through a combination of direct patterning via EUV lithography and subsequent fabrication steps employing a spacer approach, we achieved highly uniform pores of precisely controlled

dimensions across the entire wafer. The fabricated nanopores were thoroughly characterized in an aqueous environment, assessing their current-voltage (I-V) characteristics and noise performance. Device functionality was subsequently validated through the successful translocation of DNA molecules. Based on these findings, we project that this fabrication process demonstrates significant potential for realizing both single nanopores and multi-pore arrays with various tuneable dimensions within a single wafer, thereby targeting a wide range of applications.

3:40 PM

43-5 Monolithic 3D-Integrated Neuromorphic Biosignal Processor Using Hydrogen-Controlled Dual-Gate TFTs on a Flexible Substrate

Runxiao Shi[1], Shiran Guo[2], Linghan Wang[3], Siyang Liu[3], Wangran Wu[3], Yushen Hu[2], Weifeng Sun[3], Man Wong[2]
[1]School of Integrated Circuits, Southeast University, [2]The Hong Kong University of Science and Technology, [3]Southeast University

This work presents a neuromorphic bio-signal processing system that monolithically integrates a sensing electrode array (SEA), analog front-end (AFE), and artificial neural network (ANN) on a flexible substrate. The system utilizes dual-gate (DG) thin-film transistors (TFTs) with indium gallium zinc oxide as the semiconductor, fabricated via a temperature gradient process to suppress hydrogen release during multi-layer silane-based deposition. The AFE and ANN are stacked beneath the SEA, forming 3D architecture. A DG-TFT dynamic compensation method ensures the AFE maintains high performance under process variations, achieving a gain of 62.3 dB, a 1.1 kHz bandwidth, and a 74.5 dB common-mode rejection ratio within a compact 1.5 mm^2 footprint. The ANN uses DG-TFTs for in-memory computing, with one gate for data storage and the other for weight storage, achieving 96% classification accuracy across four heartbeat signal classes.

4:05 PM

43-6 Aligned Carbon Nanotube BioFETs: Toward High-Throughput Single-Molecule Sensing

Lijun Liu[1,2], Sybren Santermans[1], Dennis Lin[1], Luca Mana[1,2], Marina Y. Timmermans[1], Himanshu Sharma[1], Erika Giorgione[1,2], Jean-Francois de Marneffe[1], Aryan Afzalian[1], Tim Stakenborg[1], Stef Van Loon[2], Lei Zhang[1], Pol Van Dorpe[1,2], Cesar Javier Lockhart de la Rosa[1], Gouri Sankar Kar[1], Koen Martens[1,2]
[1]imec, [2]KU Leuven

We present the first demonstration of biomolecule detection with Aligned Carbon Nanotube FETs using wafer-scale, solution-based assembled ACNTs. The devices feature a 70 nm gate length and 250 nm width. Unlike previously reported single-molecule CNT bioFETs, based on single-tube fabrication approaches that lack scalability, our approach addresses manufacturability. A key challenge, removal of the ACNT wrapping polymer, is tackled using a Transient Assisted Plasma (TAP) clean. We show improved CNT cleanliness, improved device performance, and DNA sensing yield. The electrolytic gate reveals the effect of TAP clean. Devices showed a subthreshold swing<80 mV/dec, peak transconductance>1.3 mS/μm. Binding DNA oligos to CNTs induces a positive shift of the FET threshold voltage, which serves as the DNA signal. Through both real-time and end-point measurements, we obtain a clear signal of >20 mV, in par with Si fin-bioFET. Moreover, we show the potential of single molecule CNT bioFET sensing with quantum transport simulations.

4:30 PM

43-7 Dual-gate Ion-Sensitive Ferroelectric Field-Effect Transistor with Ultra-High Sensitivity >1000 mV/pH and Low Operation Voltage <3 V

Zikang Yao[1,2], Danyang Chen[3,3], Mengwei Si[3], Longchun Wang[3], Xiaojun Guo[3], Jingquan Liu[3], Wei Tang[3], Xiuyan Li[3]
[1]National Key Laboratory of Advanced Micro and Nano Manufacture Technology, Shanghai Jiao Tong University, [2]School of Integrated Circuits (School of Information Science and Electronic Engineering), Shanghai Jiao Tong University, [3]Shanghai Jiao Tong University

This work, for the first time, demonstrates high performance of dual-gate ion-sensitive ferroelectric field-effect transistor (DG-ISFeFET) with antiferroelectric (AFE) $Hf_xZr_{1-x}O_2$ (HZO) film and In_2O_3 (InO) semiconductor. Super-Nernstian sensitivity of >1000 mV/pH is achieved for pH sensing with a low operation voltage of ~3 V, which benefits from current amplification induced by polarization switching in AFE HZO in addition to dual-gate structure optimization. This work paves a novel way to ultra-high ISFET by leveraging CMOS compatible ferroelectric technology.

IEDM Committee List

Executive Committee

General Chair
Jan Hoentschel
GlobalFoundries

Technical Program Chair
Meng-Fan (Marvin) Chang
National Tsing Hua University & TSMC

Technical Program Vice Chair
Jungwoo Joh
Texas Instruments

Publications Chair
Geert Eneman
imec

Publications Co-Chair
Srabanti Chowdhury
Stanford University

Publicity Chair
Gaudenzio Meneghesso
University of Padova

Publicity Co-Chair
John Paul Strachan
Forschungszentrum Jülich

Courses Chair
Sandy Liao
TSMC

Courses Co-Chair (Short Courses)
Kang-ill Seo
Samsung

Courses Co-Chair (Tutorials)
Sangbum Kim
Seoul National University

Focus Session and Special Events Chair
Uygar Avci
Intel

Focus Session & Special Events Co-Chair
Olga Blum
Sandia National Laboratories

Virtual Arrangements Chair
Dechao Guo
IBM

Asian Arrangements Chair
Rihito Kuroda
Tohoku University

Asian Arrangements Co-Chair
Vita Pi-Ho Hu
National Taiwan University

European Arrangements Chair
Elena Gnani
University of Bologna

European Arrangements Co-Chair
Michael Waltl
TU Wien

Advanced Logic Technology

Byounghak Hong, *Chair*
Samsung

Veeraraghavan Basker
AMAT

Daphnée Bosch
Cea-Leti

Vincent Chang
TSMC

Gabriela Dilliway
SemiTech Insights

Kwan-Yong Lim
Qualcomm

Chung-Hsun Lin
Intel

Amy Novac
Advanced Micro Devices

Martin O'Toole
ASML

Shay Rebh
IBM

Changhwan Shin
Korea University

Koji Watanabe
Tokyo Electron

Liesbeth Witters
IMEC

Emerging Device and Compute Technology

Tania Roy, *Chair*
Duke University

Ching-Tzu Chen
IBM

Albert Cheng
TSMC

Saptarshi Das
Penn State

Nazila Haratipour
Kepler Computing

Adrian Ionescu
EPFL

Asif Khan
Georgia Tech

Soo Gil Kim
SK Hynix

Suhwan Lim
Samsung

David Michalak
TNO

Farnaz Niroui
MIT

Maud Vinet
Quobly

Memory Technology

Nanbo Gong, *Chair*
IBM

Attilio Belmonte
imec

Mattia Boniardi
STMicroelectronics

Katherine Chiang
TSMC

Wei-Chih Chien
Natcast

Ali Khakifirooz
Intel

Seiyon Kim
SK Hynix

Kyunghwan Lee
Samsung

Kai-Shin Li
TSRI

Johannes Mueller
GlobalFoundries

Hongmei Wang
Micron

Shimeng Yu
Georgia Tech

Yanli Zhang
Infineon

Modeling and Simulation

Jing Wang, *Chair*
Nvidia

Aryan Afzalian
imec

Dipanjan Basu
Synopsys

Lado Filipovic
TU Wien

Mohammad Hasanuzzaman
IBM Research

Tzu-Hsuan (Bruce) Hsu
Macronix

Changwook Jeong
UNIST

Raseong Kim
Intel

Dongyean Oh
SK Hynix

Divya Prasad
AMD

Benoît Sklénard
CEA-Leti

Pin Su
National Yang Ming Chiao Tung University

Yan Zhang
MIT Lincoln Lab

Neuromorphic and Novel Computing

Daniele Ielmini, *Chair*
Politecnico di Milano

Christopher Bennett
Sandia National Laboratories

Valeria Bragaglia
IBM Research Europe-Zurich

Erika Covi
University of Groningen

Giuseppe Desoli
STMicroelectronics

Tuo-Hung (Alex) Hou
National Yang Ming Chiao Tung University / TSRI

Thomas Kämpfe
Fraunhofer IPMS

Seyoung Kim
POSTECH

Ravi Kumar
Infineon

Ming-Hsiu (Eric) Lee
Macronix

Kaushik Roy
Purdue University

Xin Si
Southeast University of China

Huaqiang Wu
Tsinghua University

Optoelectronics, Displays, and Imaging Systems

Jun Ogi, *Chair*
Sony

Dan Buca
Forschungszentrum Jülich

Matteo Buffolo
University of Padova

KeunYeong Cho
Samsung

Michael Cich
Google

Gauri V. Karve
Imec

Frederic Lalanne
STMicroelectronics

Stefan Lischke
IHP Microelectronics

Geunsook Park
Omnivision

Haiding Sun
University of Science and Technology of China

Susanna Thon
Johns Hopkins University

Daniel Van Blerkom
SWIRLabs

Keita Yasutomi
Shizuoka University

Power, Microwave/Mm-Wave and Analog Devices/Systems

Hongping Zhao, *Chair*
Ohio State University

Marina Antoniou
University of Warwick

Colombo Bolognesi
ETH Zurich

Julien Buckley
CEA-Leti

Ho-Young Cha
Hongik University

Kevin Chen
Hong Kong University of Science and Technology

Mengyuan Hua
Southern University Science & Technology

Vibhor Jain
GlobalFoundries

Jose Jimenez
Qorvo

Christian Koller
Infineon

Munetaka Noguchi
Mitsubishi Electric

Zhikai Tang
Texas Instruments

Hann-Huei Tsai
National Institutes of Applied Research

Yuhao Zhang
University of Hong Kong

Reliability of Systems and Devices

Bonnie Weir, *Chair*
Broadcom

David Chu
TSMC

Grace Chung
Samsung

Sumy Jose
NXP

Mario Lanza
University of Singapore

Maria Toledano Luque
GlobalFoundries

Kai Ni
University of Notre Dame

Andrea Padovani
University of Modena

Stephen Ramey
Intel

Sachin Saptnekar
University of Minnesota

Motoyuki Sato
Tokyo Electron

Runsheng Wang
Peking University

Huimei Zhou
IBM

Sensors, MEMS, and Bioelectronics

Sheng-Shian Li, *Chair*
National Tsing Hua University

Taeko Ando
Ritsumeikan University

Humberto Campanella
Silicon Austria Labs GmbH

Sophie Giroud
CEA-Leti

Wan-Thai Hsu
Soundskrit

Dion Khodagholy
University of California, Irvine

Mario Kiuchi
Sumitomo Precision Products

Chiung Lo
xMEMS

Pierpaolo Palestri
University of Modena and Reggio Emilia

Roozbeh Tabrizian
University of Florida

Eng-Huat Toh
GlobalFoundries

Xinran Wang
Nanjing University

Cunjiang Yu
University of Illinois, Urbana-Champaign