FUTURE TRENDS IN MICROELECTRONICS
FUTURE TRENDS IN MICROELECTRONICS
FUTURE TRENDS
IN MICROELECTRONICS
Frontiers and Innovations

Edited by
SERGE LURYI
JIMMY XU
ALEX ZASLAVSKY

WILEY
## CONTENTS

### Preface
*S. Luryi, J. M. Xu, and A. Zaslavsky*  
ix

## I INNOVATIONS IN ELECTRONICS AND SYSTEMS

1. **Technology Innovation, Reshaping the Microelectronics Industry**  
   *K. Kim and U.-I. Chung*  
   4

2. **Challenges and Limits for Very Low Energy Computation**  
   *F. Balestra*  
   49

3. **Getting Rid of the DRAM Capacitor**  
   *N. Rodriguez, F. Gamiz, and S. Cristooveanu*  
   59

4. **Physics and Design of Nanoscale Field Effect Diodes for Memory and ESD Protection Applications**  
   *D. E. Ioannou, Z. Chbili, A. Z. Badwan, Q. Li, Y. Yang, and A. A. Salman*  
   73

5. **Sharp-Switching CMOS-Compatible Devices with High Current Drive**  
   *J. Wan, S. Cristooveanu, S. T. Le, A. Zaslavsky, C. Le Royer, S. A. Dayeh, D. E. Perea, and S. T. Picraux*  
   81

6. **Magnetic Tunnel Junctions with a Composite Free Layer: A New Concept for Future Universal Memory**  
   *A. Makarov, V. Sverdlov, and S. Selberherr*  
   93

7. **Silicon Carbide High Temperature Electronics – Is This Rocket Science?**  
   *C.-M. Zetterling*  
   102

8. **Microchip Post-Processing: There is Plenty of Room at the Top**  
   *J. Schmitz*  
   110

9. **EUV Lithography: Today and Tomorrow**  
   *V. Y. Banine*  
   120

10. **Manufacturability and Nanoelectronic Performance**  
    *M. J. Kelly*  
    133
Contents

II OPTOELECTRONICS IN THE NANO AGE 139

Ultrafast Nanophotonic Devices For Optical Interconnects 142
N. N. Ledentsov, V. A. Shchukin, and J. A. Lott

Will Optical Communications Meet the Challenges 160
of the Future?
D. K. Mynbaev

Optical Antennae for Optoelectronics: 173
Impacts, Promises, and Limitations
H. Mohseni

Spin Modulation: Teaching Lasers New Tricks 183
J. Lee, G. Boéris, R. Oszwaldowski, K. Výborný,
C. Göthgen, and I. Żutić

III HARVESTING ENERGY FROM THE SUN 191
AND THE ENVIRONMENT

Silicon Photovoltaics: Accelerating to Grid Parity 194
M. R. Pinto

Two- and Three-Dimensional Numerical Simulation of 210
Advanced Silicon Solar Cells
E. Sangiorgi, M. Zanuccoli, R. De Rose, P. Magnone, and C. Fiegna

Mechanical Energy Harvesting with Piezoelectric 230
Nanostructures: Great Expectations for Autonomous Systems
G. Ardila, R. Hinchet, L. Montès, and M. Mouis

Charged Quantum Dots for Photovoltaic 244
Conversion and IR Sensing
A. Sergeev, V. Mitin, N. Vagidov, and K. Sablon

Active Optomechanical Resonators 254
D. Princepe, L. Barea, G. O. Luiz, G. Wiederhecker, and N. C. Frateschi
IV PHYSICS FRONTIERS

State of the Art and Prospects for Quantum Computing
M. I. Dyakonov

Wireless, Implantable Neuroprosthesis: Applying Advanced Technology to Untether the Mind
D. A. Borton and A. V. Nurmikko

Correlated Electrons: A Platform for Solid State Devices
S. D. Ha, Y. Zhou, R. Jaramillo, and S. Ramanathan

Graphene-Based Integrated Electronic, Photonic and Spintronic Circuit
A. D. Güçlü, P. Potasz, and P. Hawrylak

Luttinger Liquid Behavior of Long GaAs Quantum Wires

Toward Spin Electronic Devices Based on Semiconductor Nanowires

An Alternative Path for the Fabrication of Self-Assembled III-Nitride Nanowires
A. Haab, M. Mikulics, T. Stoica, B. Kardynal, A. Winden, H. Hardtdegen, D. Grützmacb, and E. Sutter

InAs Nanowires with Surface States as Building Blocks for Tube-Like Electrical Sensing Transistors
N. V. Demarina, M. I. Lepsa, and D. Grützmacb

Lévy Flight of Photoexcited Minority Carriers in Moderately Doped Semiconductors: Theory and Observation
A. Subashiev and S. Luryi

Terahertz Plasma Oscillations in Field Effect Transistors: Main Ideas and Experimental Facts
W. Knap and M. I. Dyakonov

INDEX
IV PHYSICS FRONTIERS 263
State of the Art and Prospects for Quantum Computing 266
M. I. Dyakonov

Wireless, Implantable Neuroprosthesis: Applying Advanced Technology to Untether the Mind 286
D. A. Borton and A. V. Nurmikko

Correlated Electrons: A Platform for Solid State Devices 300
S. D. Ha, Y. Zhou, R. Jaramillo, and S. Ramanathan

Graphene-Based Integrated Electronic, Photonic and Spintronic Circuit 308
A. D. Güçlü, P. Potasz, and P. Hawrylak

Luttinger Liquid Behavior of Long GaAs Quantum Wires 319

Toward Spin Electronic Devices Based on Semiconductor Nanowires 328

An Alternative Path for the Fabrication of Self-Assembled III-Nitride Nanowires 340
A. Haab, M. Mikulics, T. Stoica, B. Kardynal, A. Winden, H. Hardtdegen, D. Griitzmacher, and E. Sutter

InAs Nanowires with Surface States as Building Blocks for Tube-Like Electrical Sensing Transistors 351
N. V. Demarina, M. I. Lepsa, and D. Griitzmacher

Levy Flight of Photoexcited Minority Carriers in Moderately Doped Semiconductors: Theory and Observation 359
A. Subashiev and S. Luryi

Terahertz Plasma Oscillations in Field Effect Transistors: Main Ideas and Experimental Facts 373
W. Knap and M. I. Dyakonov

INDEX 395
Preface

S. Luryi  
Dept. of Electrical and Computer Engineering  
State University of New York, Stony Brook, NY 11794-2350, U.S.A.

J. M. Xu and A. Zaslavsky  
School of Engineering, Brown University, Providence, RI 02912, U.S.A.

This book is a brainchild of the seventh workshop in the Future Trends in Microelectronics series (FTM-7). The first of the FTM conferences, "Reflections on the Road to Nanotechnology", had gathered in 1995 on Ile de Bendor, a beautiful little French Mediterranean island.\(^1\) The second FTM, "Off the Beaten Path", took place in 1998 on a larger island in the same area, Ile des Embiez.\(^2\) Instead of going to a still larger island, the third FTM, "The Nano Millennium", went back to its origins on Ile de Bendor in 2001.\(^3\) As if to compensate for small size of Bendor, the fourth FTM, "The Nano, the Giga, the Ultra, and the Bio", took place on the biggest French Mediterranean island of them all, Corsica.\(^4\) Normally, the FTM workshops gather every three years; however, the FTM-4 was held one year ahead of the usual schedule, in the summer of 2003, as a one-time exception. Continuing its inexorable motion eastward, the fifth FTM workshop, "Up the Nano Creek", had convened on Crete, Greece, in June of 2006.\(^5\) The inexorable motion was then interrupted to produce a semblance of a random walk in the Mediterranean and the FTM-6 "Unmapped Roads" went to the Italian island of Sardinia (June, 2009).\(^6\) The last FTM gathering, "Into the Cross Currents", returned to our earlier venue on Corsica (June 2012).

The FTM workshops are relatively small gatherings (less than 100 people) by invitation only. If you, the reader, wish to be invited, please consider following a few simple steps outlined on the conference website. The FTM website at www.ece.sunysb.edu/~serge/FTM.html contains links to all past and planned workshops, their programs, publications, sponsors, and participants. Our attendees have been an illustrious lot. Suffice it to say that among FTM participants we find five Nobel laureates (Zhores Alferov, Herbert Kroemer, Horst Stormer, Klaus von Klitzing, and Harold Kroto) and countless others poised for a similar distinction. To be sure, high distinction is not a prerequisite for being invited to FTM, but the ability and desire to bring fresh ideas is. All participants of FTM-7 can be considered authors of this book, which in this sense is a collective treatise.

The main purpose of FTM workshops is to provide a forum for a free-spirited exchange of views, projections, and critiques of current and future directions, among the leading professionals in industry, academia, and government.
For better or worse our civilization is destined to be based on electronics. Ever since the invention of the transistor and especially after the advent of integrated circuits, semiconductor devices have kept expanding their role in our lives. Electronic circuits entertain us and keep track of our money, they fight our wars and decipher the secret codes of life, and one day, perhaps, they will relieve us from the burden of thinking and making responsible decisions. Inasmuch as that day has not yet arrived, we have to fend for ourselves. The key to success is to have a clear vision of where we are heading. In the blinding light of a bright future, the FTM community has remained mindful of the fact that what controlled the past will still control the future – the basic principles of science. Thus, the trendy, red-hot projections of any given epoch deserve and require critical scrutiny.

Some degree of stability is of importance in these turbulent times and should be welcome. Thus, although the very term "microelectronics" has been generally re-christened "nanoelectronics", we have stuck to the original title of the FTM workshop series.

The present volume contains a number of original papers, some of which were presented at FTM-7 in oral sessions, other as posters. From the point of view of the program committee, there is no difference between these types of contributions in weight or importance. There was, however, a difference in style and focus – and that was intentionally imposed by the organizers. All speakers were asked to focus on the presenter's views and projections of future directions, assessments or critiques of important new ideas/approaches, and not on their own achievements. This latter point is perhaps the most distinctive feature of FTM workshops. Indeed, we are asking scientists not to speak of their own work! This has proven to be successful, however, in eliciting powerful and frank exchange. The presenters are asked to be provocative and/or inspiring. Latest advances made and results obtained by the participants are to be presented in the form of posters and group discussions.

Each day of the workshop was concluded by an evening panel or poster session that attempted to further the debates on selected controversial issues connected to the theme of the day. Each such session was chaired by a moderator who invited two or three attendees of his or her choice to lead with a position statement, with all other attendees serving as panelists. The debate was forcefully moderated and irrelevant digressions cut off without mercy. Moderators were also assigned the hopeless task of forging a consensus on critical issues.

To accommodate these principles, the FTM takes a format that is less rigid than usual workshops to allow and encourage uninhibited exchanges and sometimes confrontations of different views. A central theme is designed together with the speakers for each day. Another traditional feature of FTM workshops is a highly informal vote by the participants on the relative importance of various fashionable current topics in modern electronics research. This tradition owes its origin to Horst Stormer, who composed the original set of questions and maintained the results over four conferences. These votes are perhaps too bold and irreverent for general publication, but they are carefully maintained and made...
available to every new generation of FTM participants. Another traditional vote concerned the best poster. The 2012 winning poster was "Mechanical energy harvesting with piezoelectric nanostructures: Great expectations for autonomous systems" by Gustavo Ardila.

A joyful tradition of FTM meetings is the settling of scientific bets, a custom that dates back to the 1998 wager between Nikolai Ledentsov (pro) and Horst Stormer (con) about the putative future dominance of quantum dot-based lasers – a bet that Horst collected in 2004, at FTM-4. Another risky bet on the future dominance of SOI technology is to be adjudicated at a future (2015) workshop. The precise statement of this "good" bet (worth a six-magnum case of very good champagne) is: by 2015, SOI will cover more than 35% of the CMOS market including memories by value. This bet, proposed by Sorin Cristoloveanu, attracted three cons – Detlev Grützmacher, Dimitris Ioannou, and Enrico Sangiorgi – who will have to divide the spoils, should SOI fail to reach the mark. Several "bad" (penniless) bets that were supposed to be resolved at FTM-7 were put off as premature by the principals. Hopefully, by 2015 it will become clear whether or not "10% of all man-produced light will be white LED" (Michael Shur), "we shall have hi-fi real-time pocket translators" (Hiroshi Iwai), "most cars will have IP addresses" (Jimmy Xu), or "the Si industry will no longer be recognizable and there will be (essentially) no further improvement in devices" (Paul Solomon).

Not every contribution presented at FTM-7 has made it into this book (not for the lack of persistence by the editors). Perhaps most sorely we miss the exciting contribution by David Miller of Stanford University, entitled "The heat death of information processing and why interconnects matter more than logic", in which he illustrated how the energy required by data communication is imposing severe limits on overall information processing, with the demand for communications bandwidth increasing a hundredfold every decade. As a result, he posited very low energy optoelectronics densely integrated with logic as a way of allowing the continued scaling of information processing systems. Abstracts of his and all other presentations can be found on the program webpage, http://www.ee.sunysb.edu/~serge/ARW-7/program.html

The FTM meetings are known for the professional critiques – or even demolitions – of fashionable trends, that some may characterize as hype. The previous workshops had witnessed powerful assaults on quantum computing, molecular electronics, and spintronics. This time Michel Dyakonov offered another thorough and conclusive update on quantum computing. It seems that by now most of the hype associated with some of these trends has dissipated and perhaps we can take some credit for the more balanced outlook that has emerged since.

We have grouped all contributions into four chapters: one dealing with transistors and CMOS or CMOS-compatible technology, another with photonics and lasers, and still another with solar cell and energy harvesting. The last chapter collected the contributions focused on fundamental physics and other, less
technological subjects. The breakdown could not be uniquely defined, because some papers fit two or even three categories!

To produce a coherent collective treatise out of all of this, the interaction between FTM participants had begun well before their gathering at the workshop. All the proposed presentations were posted on the web in advance and could be subject to change up to the last minute to take into account peer criticism and suggestions. After the workshop is over, these materials (not all of which have made it into this book) remain on the web indefinitely, and the reader can peruse them starting at the www.ece.sunysb.edu/~serge/FTM.html home page.

Acknowledgments

The 2012 FTM workshop on Corsica and therefore this book were possible owing to support from:

- US National Science Foundation
- U.S. Department of Defense: ARO
- U.S. DoD European offices: EOARD, ONR–Global
- Industry: SAIT, Samsung Electronics Co.
- Academia: SUNY–Stony Brook, Brown University.

On behalf of all workshop attendees sincere gratitude is expressed to the above organizations for their generous support and especially to the following individuals whose initiative was indispensable: William Clark, John Gonglewski, Kinam Kim, Mark Pinto, Shawn Thorne, Robert True, and John Zavada.

Finally, the organizers wish to thank all of the contributors to this volume and all the attendees for making the workshop a rousing success.
References


Part I

Innovations in Electronics and Systems
I. Innovations in Electronics and Systems

The incredibly powerful silicon electronics is still full of steam, charging ahead despite the putatively insurmountable barriers and walls (of thermal, power, wiring, and scaling, for example). Equally evident is that the challenges ahead are real, getting greater, and often without a clear answer. One example is transition to EUV lithography, discussed in the chapter by Banine. What are the biggest challenges anticipated and the biggest changes assumed in the existing roadmap? If science is not (yet) the ultimate limit to nanoelectronics, is it manufacturability, discussed in the chapter by Kelly? What are the most desired innovations, and the most disruptive, interesting, and controversial directions? While the drive for ever greater device performance – discussed in a number of chapters devoted to novel memory and logic devices – is becoming prohibitively costly, could the advances in systems and integrated functionalities be sufficiently rewarding to justify and become the next driving force? These questions are addressed in Part I of this book, which opens with an insightful and comprehensive discussion of technology innovation by Kim and Chung, from the innovation powerhouse of Samsung Electronics.

Contributors

1.1 K. Kim and U-I. Chung
1.2 F. Balestra
1.3 N. Rodriguez, F. Gamiz, and S. Cristoloveanu
1.4 D. E. Ioannou, Z. Chbili, A. Z. Badwan, Q. Li, Y. Yang, and A. A. Salman
1.5 J. Wan, S. Cristoloveanu, S. T. Le, A. Zaslavsky, C. Le Royer, S. A. Dayeh, D. E. Perea, and S. T. Picraux
1.6 A. Makarov, V. Sverdlov, and S. Selberherr
1.7 C.-M. Zetterling
1.8 J. Schmitz
1.9 V. Y. Banine
1.10 M. J. Kelly
Technology Innovation, 
Reshaping the Microelectronics Industry

Kinam Kim and U-In Chung
Samsung Advanced Institute of Technology, Samsung Electronics Co. Ltd.
Giheung, Gyonggi-do, South Korea

1. Introduction

The remarkable evolution of electronics was enabled by the rapidly advancing silicon technology. The performance of such devices as CPUs and memories has improved tremendously over the past 20 years: CPUs improved by a factor of 2,400, DRAM by a factor of 1,000 and NAND Flash by a factor of 32,000. Mobile network speed has also increased by a factor of 840, stimulated by an increased usage of smartphones and tablets, which have grown by 58% and 260% in 2011, respectively. Furthermore, Internet communication traffic is anticipated to increase by 32% annually, with the traffic reaching an order of zeta ($10^{21}$) bytes by 2015.

These performance improvements have been made possible by silicon technology downscaling, which is nearing the 20 nm node and will soon reach 10 nm. Novel fabrication techniques, structures and materials will be required to continue scaling beyond 10 nm to satisfy the data handling, processing and storage requirements of the future.

Future device development will be constrained by power consumption, in addition to the traditional scaling issues, since faster and higher data handling and processing will unavoidably require a lot of power. For instance, 1.1 petaflop ($10^{15}$) computing in 2008 consumed 2.3 MW power for its operation, so an exascale ($10^{18}$) system might dissipate an estimated 2.3 GW, roughly equivalent to the electricity produced by two nuclear power plants. Most of this energy is lost in the inefficient metal interconnects, which could be reduced by combining photonics with Si technologies. This silicon-photonics convergence will create new applications and markets in the future.

Furthermore, silicon technology is bringing innovation to new areas, such as energy, health, and medical applications. Examples of this are ultrafast DNA sequencing, extremely compact and efficient medical imaging devices, and low-cost energy-efficient lighting, among others. These applications, fueled by Si technologies, could eventually take up a major portion of the semiconductor market in the next decade.
2. **Mainstream silicon technology: Memory**

- *Towards sub-20 nm DRAM*

  DRAM density has doubled every 18 months through scaling of the critical dimensions, approaching the 20 nm node today, as shown in Fig. 1. Major challenges to downscaling to the sub-10 nm regime will require device innovations in cell storage capacitors, cell array transistors (CATs) and patterning processes.\(^5\)

  The minimum cell capacitance must be maintained to provide an adequate signal for sensing and to meet retention time specifications.\(^6\) In order to maintain the cell capacitance, as DRAM downscaled from 90 nm to 20 nm, the capacitor structure has been changed from a simple 3D cylindrical shape to an extremely high aspect ratio (AR) supported cylinder or pillar structure, as shown in Fig. 2. For sub-10 nm DRAMs, structural innovation may not help any longer due to physical limitations in accommodating complex 3D structures. At sub-10 nm nodes, the distance between electrodes becomes \(\sim\)10 nm or less, which in turn requires the physical thickness of the storage electrodes and the dielectrics to be thinner than 5 nm or less. These requirements will be very difficult to meet. However, the limitation of cell capacitance can be compensated by the technological innovation in CAT with extremely low leakage currents and/or array architecture and its integration process to provide high sensing signals.

  From a device point of view, requirements related to the retention limit can be overcome by suppressing the storage node leakage current \(I_{\text{LEAK}}\). The data retention time \(t_{\text{RE}}\) is fundamentally determined by the amount of the stored charge and the time-dependent charge loss at the storage node.\(^7\) The sensing voltage \(\Delta V_{\text{BL}}\) is as follows:

  \[
  \Delta V_{\text{BL}} = \frac{C_S}{C_{\text{BL}} + C_S} \left( \frac{V_{\text{ARR}}}{2} - \Delta V_L \right),
  \]

  where \(C_S\) is the cell storage capacitance, \(V_{\text{ARR}}\) is the array voltage, \(C_{\text{BL}}\) is the bit

![Figure 1. DRAM technology roadmap of ITRS and top DRAM supplier.](image-url)
Future Trends in Microelectronics

line loading capacitance, and $\Delta V_L$ is the voltage loss due to charge loss during the retention. The charge loss can be expressed by $Q_{\text{LEAK}} = C_S(\Delta V_L) = I_{\text{LEAK}}t_{\text{RE}}$, from which the data retention time can be obtained as follows:

$$t_{\text{RE}} = \frac{C_S(V_{\text{ARR}}/2 - \Delta V_{\text{BL}}) - C_{\text{BL}}(\Delta V_{\text{BL}})}{I_{\text{LEAK}}}.$$  \hspace{1cm} (2)

As shown in Eq. (2), the leakage current $I_{\text{LEAK}}$ must be suppressed as much as possible (to less than sub-fA/cell) to ensure good data retention characteristics. The leakage $I_{\text{LEAK}}$ originates mainly from the GIDL, subthreshold current, and junction leakage of the CAT. In recent years, CAT structures have successfully evolved from a planar-channel to a recessed-channel (RCAT), and later to a metal buried-gate (BCAT) structure that eventually will evolve to a vertical-channel (VCAT) structure satisfying the $I_{\text{ON}}/I_{\text{LEAK}} \sim 10^{10}$ ratio requirement. Figure 2 shows the evolution of the cell array transistor through the years.

The sensing signal voltage $\Delta V_{\text{BL}}$ can also be improved by reducing $C_{\text{BL}}$. Vertical scaling of the bit line (BL) electrodes with lower resistivity metal and gap filling with lower-$\kappa$ dielectrics between BLs are the key points in reducing $C_{\text{BL}}$ and thereby enhancing $\Delta V_{\text{BL}}$.

Demands for higher speed devices with reduced power consumption require more stringent processes and technological breakthroughs in transistors. Aggressive scaling employing higher-$\kappa$ gate oxide with multi-thickness, ultra-shallow junctions with carefully controlled thermal activation, as well as novel mobility-boosting technologies based on strain effects are possible candidates for the enhancement of device performance and power reduction.

$$\sim 90 \text{ nm ArF} \ // \sim 40 \text{ nm ArF-i (DPT)} \ // < 20 \text{ nm EUV}$$

![Figure 2](image.png)

**Figure 2.** Key technology evolution of DRAM from 90 nm to sub-20 nm: lithography and critical size (top), cell capacitor (middle), and cell array transistor (bottom).
With the optimization of the 3D cell storage capacitor and array transistor structures, DRAM is anticipated to downscale to the sub-10nm regime. Nevertheless, in the future, patterning technology, e.g. lithography and etching, will become critical. Currently, the major concern is the productivity of such technologies: i.e. extreme ultraviolet (EUV) and/or double patterning lithography, and high aspect ratio capacitor hole etching. The slow progress of EUVL technology will result in the need for expensive double or quadruple patterning technologies (DPT or QPT) for several critical layers at the 22 nm node. The rapid increase in fabrication costs, as well as overlay and uniformity issues are becoming major challenges due to the drastic increase in the number of process steps.

- **Beyond sub-20 nm NAND flash**
  
  The conventional floating-gate (FG) NAND flash cell technology is currently at the 20 nm node. Moreover, sub-20nm cells using QPT have also been demonstrated using the word line (WL) air-gap technology in order to overcome coupling interference. However, as device technology enters the sub-20 nm region, cell-to-cell coupling interference, the numbers of stored electrons, WL-to-WL breakdown, endurance and data retention are becoming major challenges.

  Although circuit technologies including parallel programming, shadow programming and extended ECC (error correcting circuit) were successful in overcoming the cell-to-cell coupling down to the 20 nm node, they would not be effective in overcoming issues for beyond the 20 nm node planar NAND flash because coupling interference ratio is inversely proportional to the design rule for planar structures. Thus, as the design rule decreases, the coupling ratio will further increase, reaching the allowed design limit, which is about 5 at around 20 nm as shown in Fig. 3.

  However, a 3D NAND flash at the 10 nm node can be designed in such a way that the coupling interference ratio is at the same level as that of a 60 nm planar cell. The coupling interference ratio of the 3D vertical NAND (VNAND) remains far below the critical design limit even beyond the sub-10 nm region, as shown in Fig. 3. This is because the bit line coupling interference is almost eliminated by the gate-all-around (GAA) structure, and the WL-to-WL coupling interference can be reduced by the increased spacing between WLs.

  Among several proposed 3D vertical structures, the terabit CAT (TCAT) is thought to be the most promising. It is based on a damascened metal gate and a TANOS (TiN-AlO-nitride-oxide-Si) cell. The metal gate in this structure could have a wide programming window because TANOS has a better erase speed than SONOS (Si-oxide-nitride-oxide-Si). Furthermore, by shrinking the diameter of the channel hole, erase speed can be further improved as the electric field in the tunnel oxide is enhanced. The TCAT erase scheme is the same as that of planar cells, so that it can be compatible with the conventional NAND flash architecture. Furthermore, it has a better array distribution uniformity by a factor of two and ten times the endurance of the 20 nm node planar technology.
Future Trends in Microelectronics

Figure 3. Gate-all-around charge trap flash (CTF) structure (bottom left) and 3D VNAND with gate-all-around CTF structure (top left). The coupling interference ratio of 3D VNAND is much lower than in planar devices. The coupling interference ratio defines the charge build-up in cell while an adjacent cell is programming, and should not exceed the allowed design limit of ~ 5 (dashed horizontal line).

Further research will be necessary in order to achieve the required read margin and write speeds, which could be negatively affected by a low mobility and high leakage of channel poly-Si. To reduce the leakage current, the poly-Si has to be thin enough to be fully depleted, and its grain size has to be enlarged in order for its performance to be comparable to that of the single-crystal Si. Process technologies, such as high aspect ratio channel-hole formation, bowing and leaning free stacks, will also require further research.

- **STT-MRAM as a post-DRAM technology**

As the push for smaller DRAM cells faces formidable technological and economical challenges, a number of alternative post-DRAM devices are under consideration. Because of the severe charge leakage for downscaled storage elements, non-charge-based memory devices seem quite promising. One such device is the spin transfer torque magnetoresistive random access memory (STT-MRAM), which has a high operation speed, superior endurance and process compatibility with existing Si technology. The STT-MRAM technology utilizes magnetic tunnel junction (MTJ) tunneling for resistive storage as shown in Fig. 4, where MgO is widely used as the magnetic tunnel barrier. It can reach a similar area density as DRAM, since a cell consists of a single transistor as a selection device and a single MTJ as a storage element.

The key challenge in STT-MRAM is reducing the operating current to switch the spin states. The required currents to switch circular-shaped MTJ cells have been estimated to lie in the $J_c = 1-10 \times 10^6$ A/cm$^2$ range, as plotted in Fig. 5. The
Figure 4. STT-MRAM structure where the magnetic tunnel junction (MTJ) exhibits two different resistance states depending on the relative spin orientation of the MTJ layers.

various drive currents of underlying transistors at each technology node are also shown. In order to successfully write the STT-MRAM cell, smaller $J_C$ and/or larger transistor current drive will be required.\textsuperscript{15} If we assume transistor current drive of $\sim500$ $\mu$A/$\mu$m gate width (lower solid line in Fig. 6), $J_C$ will have to be reduced below $3\times10^6$ A/cm$^2$ for STT-MRAM, which is extremely difficult. Thus, we need materials and structures that would require lower currents to switch the spin state.

Figure 5. Current required to switch the spin state of the MTJ cell as a function of the technology node. Lines with symbols show the switching currents depending on the critical current density $J_C$ required to switch the state of the MTJ cell. Solid lines show the available transistor current drive $I_{ON}$ at each technology node.
The critical current density $J_c$ required to switch the magnetic moment of the in-plane free layer is given by the following equation:

\[ J_c = \left(\frac{2\alpha}{\hbar}\right) \left(\frac{2\alpha}{\eta}\right)(H_K V + \pi M_s^2 V), \tag{3} \]

where $\alpha$, $\eta$, $H_K$, $V$, and $M_s$ are the magnetic damping constant, spin torque efficiency, magnetic anisotropy constant, volume of the free layer, and the saturation magnetization, respectively. The $\pi M_s^2 V$ term in Eq. (3) arises from the surface demagnetization energy due to the out-of-plane precession of the in-plane magnetization during the switching. Since this term is absent in the perpendicular magnetic structure, switching current reduction will be possible by using materials with perpendicular magnetic anisotropy.

To further reduce the critical current density, a higher spin torque efficiency and a smaller damping constant for the perpendicular free layer are needed. However, reduction of $H_K V$ will cause data retention problems. As the small storage volume may be vulnerable to random thermal noise, the usage of perpendicular magnetic materials with high crystalline or interfacial anisotropy can improve the data retention.\(^{16,17,18}\) Otherwise, for in-plane magnetic anisotropy, thermal stability can be ensured through a 3D MTJ cell structure, which retains a large free-layer volume without an increased cell footprint by folding the free layer to a special geometry, as shown in Fig. 6.

The STT-MRAM will be the most promising candidate around the 12 nm technology node, or about 2018. In addition to reducing the switching current of the MTJs, etching damage and thermal stability of the MgO tunnel barrier also need to be improved in order to obtain adequate retention characteristics.

**Figure 6.** Two MTJ cell structures with equivalent thermal noise: (left) conventional planar MTJ cell with a long $3F$ footprint, where $F$ is the minimum feature size; (right) three-dimensional MTJ cell structure with $1F$ footprint that is robust against thermal noise because of its larger free-layer ratio.
• **ReRAM as a post-NAND technology**

Post-NAND devices are also being studied; one proposed technology is the resistive RAM (ReRAM) that makes use of resistance switching phenomena with structural simplicity. ReRAM is characterized by a resistive material acting as the memory element with a control element acting as the switch. Among the currently known material systems, tantalum oxide is the most promising candidate for the future nonvolatile ReRAM. In a Ta$_2$O$_5$/TaO$_x$ bi-layer structure, $10^{12}$ cycle endurance and ten-year retention time with fast program and erase pulses on the order of ~10 ns have been reported.$^{19}$

The resistive switching characteristics and structure of this device are shown in Fig. 7. The resistance switching in this type of ReRAM occurs in the Ta$_2$O$_5$ layer via a redox reaction. But, the superior cell performance comes from the

![Figure 7. Resistive switching characteristics and structure of the Ta$_2$O$_5$/TaO$_x$ device: (a) typical bipolar dc $I$–$V$ switching operation of a 30 nm cell (SEM in inset) from high to low resistance state (HRS to LRS); (b) schematic of the TaO$_x$ device consisting of a thin Ta$_2$O$_5$ insulating layer and a TaO$_x$ base layer. Resistive switching is due to the movement of internal oxygen ions or vacancies.](image-url)
presence of a TaOx layer that plays two key roles. First, it is able to hold a high density of oxygen ions as an oxygen reservoir that mitigates the oxygen ion depletion that causes reset-switching failures. Second, the TaOx layer limits the set switching with a finite current compliance, thereby preventing overshoot and hard breakdown during the set switching.

ReRAM devices can be operated in the metal-insulator-metal (MIM) configuration that can be integrated into a crossbar array. In this case, the device footprint can be reduced to $4F^2$; and if stacked into multi-layers, the feature size becomes $(4/n)F^2$, where $n$ is the number of the stacked layers, as shown in Fig. 8. However, the crossbar structure always has a signal disturbance problem due to unwanted stray currents from the neighboring cross-points. Thus, it needs a proper selection device, such as a diode or a threshold switch, serially connected with MIM storage element. As of this moment, it is critical to find an appropriate and reliable selection device and to integrate it with the ReRAM material in a crossbar structure.

Another feasible approach is a vertical ReRAM (VReRAM) structure (equivalent to 1T-nR structure), which has a similar layout to the 3D vertical NAND of Fig. 3. In the 1T-nR, the selection device must be integrated in the ReRAM cell because multiple cells are connected to one transistor. Because the role of the selection device is to filter out the stray current from the $n^2$ ReRAM cells, the specification of selection device for the 1T-nR structure is less tight than that of the crossbar structure. However, in the VReRAM structure, the hole sizes must be to be large enough to accommodate ReRAM cell materials, such as selection, resistive switching and base materials as well as the top electrode. This may keep VReRAM from further scaling.

When considering the technological issues of VReRAM and crossbar ReRAM, further research on self-rectifying ReRAM materials that would not require a selection device is urgently needed.

Figure 8. Cross-point ReRAM structure, where a resistive material acts as the memory element.
3. **Mainstream silicon technology: Logic**

In order to improve the performance of Si CMOS, logic devices have been continuously downscaled and integrated over the past 40 years, as shown in Fig. 9\textsuperscript{20,21}. However, as logic device size downscales into the nanometer regime, significant innovation in device structures and new materials have to be introduced in order to maintain the traditional rate of performance enhancement of past decades: e.g. Cu/low-\(\kappa\) dielectric at the 120 nm technology node; SiGe source/drain for the 65 nm node; and metal electrode/high-\(\kappa\) dielectric gate stacks replacing poly-Si/SiO\(_2\) gate stacks at the 45 nm node.

The conventional planar logic structure will continue to scale down to the 20 nm node through optimizing the HK/MG (high-\(\kappa\)/metal gate) technology, introduced at the 32 nm node. However, below 20 nm, equivalent oxide thickness (EOT) will decrease sharply because in order to suppress the short channel effects (SCEs), channel doping concentration has to be increased rapidly. The EOT constraints can be relaxed if a multi-gate structure or fin-FET is used to replace planar layout.

Recently, due to the rapid growth of mobile applications, low power consumption has joined speed and performance as a key metric. Power consumption \(P\) in logic devices is dominated by \(V_{DD}\) as follows:

\[
P \sim C_{OX}V_{DD}^2 f \text{ [active]} + I_{OFF}V_{DD} \text{ [standby]},
\]

**Figure 9.** Logic device downscaling deep into the nanometer regime, highlighting new device structures and materials, such as metal electrode/high-\(\kappa\) dielectric gate stacks replacing poly-Si/SiO\(_2\) gate stacks at 45 nm; high-\(\kappa\)/metal gate at 32 through 20 nm; and fin-FETs at the 14 nm node.
where $C_{OX}$ is the capacitance, $V_{DD}$ is the input voltage, $I_{OFF}$ is the off current and $f$ is the frequency. There are three approaches for reducing $V_{DD}$ and $I_{OFF}$ for low power mobile applications: (i) alternative transistor architectures, such as 3D or GAA layout for better electrostatic control; (ii) high mobility materials for alternative channels; and (iii) alternative devices with low subthreshold swing $SS < 60 \text{ mV/dec}$.

Of the three approaches, 3D devices with various performance boosting schemes are the furthest along, with fin-FETs being one of the best candidates for the next generation logic.

High mobility III-V or Ge channel materials are one of the promising directions for performance improvement as well as reducing power consumption. Even though there are many challenges to the 10 nm technology node or beyond, the improvement of gate stack/channel interface properties and introduction of heterogeneous process integration for new channel materials with higher electron and hole mobility should enable aggressive $V_{DD}$ scaling, thereby reducing power consumption.

One of the options to overcome thermal limit of sub-threshold slope below 60 mV/dec in a conventional CMOS is the tunneling FET (TFET). Tunneling FETs have been proposed for low $V_{DD}$ operation below 0.5 V, and would greatly benefit from low (effective) band gap energy, increasing the probability for tunneling. As a result, III-V or SiGe/Ge heterojunctions will possibly be the most effective materials for TFETs.

- **Fin-FET devices**

  In the 2011 ITRS roadmap, planar devices disappear below the $L_G = 20 \text{ nm}$ gate length due to short channel effects. As a solution for $L_G$ scaling beyond 20 nm, ITRS had considered the fully-depleted SOI (FD-SOI) or multi-gate FETs (MuGFETs), including fin-FET devices. Furthermore, FD-SOI will also likely disappear below $L_G = 15 \text{ nm}$, leaving MuGFETs as the only viable technology. These predictions arise mainly from the impact of SCEs, such as drain-induced barrier lowering (DIBL) illustrated in Fig. 10. Fully-depleted SOI can potentially improve DIBL, but cannot outperform fin-FETs because of the limited improvement in the junction depth.

  The influence of the drain electric field on the channel region, which is the physical mechanism underpinning SCEs, can be understood through the "natural length" of the device, expressed as $\lambda$ in the following equation:

  \[
  \lambda = \sqrt{\frac{\epsilon_S \epsilon_{si}}{\epsilon_{ox} t_{si} t_{ox}}},
  \]

  where $\epsilon_S$, $\epsilon_{OX}$, $t_{SI}$, $t_{OX}$ are the permittivities and thicknesses of silicon and oxide, respectively. By moving from bulk to the fin-FET structure, $t_{SI}$ can be markedly reduced. Fin-FETs will be the best choice for minimizing the natural length.