Advanced Relay Design and Technology for Energy-Efficient Electronics

Jaeseok Jeon

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2011-81
http://www.eecs.berkeley.edu/Pubs/TechRpts/2011/EECS-2011-81.html

July 7, 2011

Copyright © 2011, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

Advanced Relay Design and Technology for Energy-Efficient Electronics

By

Jaeseok Jeon

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Tsu-Jae King Liu, Chair

Professor Elad Alon

Professor Liwei Lin

Fall 2011

Advanced Relay Design and Technology for Energy-Efficient Electronics

Copyright © 2011

by

Jaeseok Jeon

Abstract

Advanced Relay Design and Technology for Energy-Efficient Electronics

by

Jaeseok Jeon

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Tsu-Jae King Liu, Chair

As the era of traditional Complementary-Metal-Oxide-Semiconductor (CMOS) technology scaling is coming to an end, continual improvements in integrated-circuit (IC) performance and cost per function are becoming difficult to achieve without increasing power density. This necessitates the investigation of alternate device technologies that surmount the fundamental CMOS energy-efficiency limit and hence enable ultra-low-power ICs. To that end, a nano-electro-mechanical (NEM) relay technology is promising, because of its immeasurably low *off*-state leakage current and abrupt turn-*on* behavior, which provide for zero static power consumption and potentially very low dynamic power consumption.

In this dissertation, relay design and process technology improvements, which led to the successful demonstration of relay-based digital IC building blocks, are discussed from both device- and circuit-level perspectives. A non-volatile (NV) memory relay design that can enable embedding of NV memory with relay-based logic circuits is also discussed. In addition, multi-electrode relays that can lead to smarter design and compact implementation of zero-leakage digital integrated circuits are discussed.

Dedicated to my family

Table of Contents

1. Introduction ---------- 1
References 6

2. Four-Terminal Relay Design Improvements ---------- 7
2.1. Introduction 7
2.2. 4-T Relay Design 7
2.2.1. Relay Fabrication Process 9
2.2.2. Dependence of Gate Switching Voltage on Body Bias 12
2.2.3. Dependence of Gate Switching Voltage on Drain Bias 14
2.2.4. Parasitic Actuation by Channel-to-Body Potential 15
2.3. Revised 4-T Relay Design 15
2.4. Summary 19
2.5. References 20

3. Four-Terminal Relay Design and Operation for Embedded Memory Applications ----- 23
3.1. Introduction 23
3.2. Memory Relay Design 23
3.3. Testing Results 25
3.4. Summary 27
3.5. References 28

4. See-Saw Relay Design for Digital Integrated Circuits ---------- 30
4.1. Introduction 30
4.2. Prototype See-Saw Relay Design 31
4.2.1. Device Design and Fabrication 31
4.2.2. Testing Results from Device Perspective 32
4.2.3. Testing Results from Circuit Perspective 35
4.3. See-Saw Relay Design Optimization 39
4.3.1. Monotonicity Analysis 39
4.3.1.1. Assumptions 40
4.3.1.2. Problem Statement 41
4.3.1.3. Case Study I 46
4.3.1.4. Case Study II 48

4.3.1.5. Case Study III 50
4.3.1.6. Fundamental Insights from the Case Studies 52
4.3.2. Validation of Monotonicity Analysis 52
4.3.2.1. Monotonicity Analysis *vs.* Nonlinear Constrained Optimization 52
4.3.2.2. Monotonicity Analysis *vs.* Finite-Element-Method Analysis 53
4.3.3. Summary of Design Optimization Study 53
4.4. Corrugated See-Saw Relay Design 55
4.4.1. Device Design and Fabrication 58
4.4.2. Testing Results 60
4.5. Summary 62
4.6. References 63

5. Multi-Electrode Relays for Compact Implementation of Digital Integrated Circuits --- 68
5.1. Introduction 68
5.2. 2-Input, Dual-Source/Drain Relay 69
5.3. 3-Input Relay 73
5.4. 4-Input Relay 77
5.5. Multi-Electrode Relay Scaling 78
5.6. Summary 79
5.7. References 80

6. Conclusions --- 81
6.1. Contributions of This Work 81
6.2. Future Directions 82
6.3. References 84

Acknowledgements

First and foremost, I would like to express my sincere, cordial, and affectionate thanks to my advisor, Professor Tsu-Jae King Liu, for her unbounded support, energy, enthusiasm, patience, encouragement, and confidence that have made my Ph.D. studies at Berkeley more fruitful and pleasant. My special thanks to Professor Elad Alon, Professor Ming Wu, and Professor Liwei Lin for graciously serving on my dissertation committee and for being my reference. I am also grateful to Professor Clark T.-C. Nguyen for his supportive supervision during my first semester here at Berkeley.

It has been my great pleasure to work with the NEM relay team. Special thanks to Fred Chen, Hossein Fariborzi, and Professor Vladimir Stojanović at MIT, Chengcheng Wang, Kevin Dwan, and Professor Dejan Marković at UCLA, and Matthew Spencer and Professor Elad Alon at Berkeley, for useful insights and discussions on relay-based integrated circuit design.

I am also indebted to all staff members of the Berkeley Microfabrication Laboratory for their constant guidance and valuable help in completing my research projects, especially Dr. William Flounders, Evan Stateler, Joseph Donnelly, Siavash Parsa, Jimmy G. M. Chang, Kim Chan, Susan Kellogg-Smith, Robert M. Hamilton, Marilyn Kushner, Jay Morford, Madeleine Leullier, Laszlo Petho, David Lo, Rosemary Spivey, Brian McNeil, and Adrienne Ruff.

I am thankful to Dr. Patrick M. Smith and Dr. Gloria Wong of Kovio, Inc. and Dr. Sang-Gil Ryu for pulsed laser annealing of the relays.

I have been fortunate to interact with excellent post-doctoral researchers, especially Dr. David Garmire, Dr. Hyuck Choo, Dr. Vincent Pott, Dr. Bongsang Kim, Dr. Wooyoung Choi, Dr. Louis Hutin, and Dr. Ruzica Jevtic. It was an always enjoyable experience collaborating with them.

I would like to thank the device group alumni for their thoughtful guidance, especially Dr. Joanna T. Lai, Dr. Donovan Lee, Dr. Anupama Bowonder, Dr. Alvaro Padilla, Dr. Reinaldo Vega, Dr. Hei Kam, Dr. Xin Sun, Dr. Cheuk Chi Lo, and Dr. Yu-Chih Tseng.

I am grateful to my colleagues in the device group: Nattapol Damrongplasit, Nuo Xu, Sunghwan Kim, Rhesa Nathanael, Jack Yaung, Zachery Jacobson, Byron Ho, I-Ru Tim Chen, Minhee Cho, Wookhyun Kwon, Eungseok Park, Patrick T. Kwong, Jaewon Jang, Changhwan Shin, Peter Matheu, Jemin Park, Yenhao Phillip Chen, Kanghoon Jeon, and Hongki Kang.

This work was supported by the Defense Advanced Research Projects Agency (DARPA) / Microsystems Technology Office (MTO) Nano-Electro-Mechanical Systems (NEMS) program.

Chapter 1

Introduction

Over the last 50 years, the evolution of information-processing technology has transformed every aspect of our lives -- the way we think, work, communicate, commute, and entertain ourselves. We use smart phones to communicate, computers at work, and digital cameras to keep alive our special memories. However, as electronic products have become a pervasive part of our daily lives, available any place at any time, reducing their energy consumption has become a major challenge. Take the data center, for example: in the year of 2000, the total electricity consumed for all data centers around the world (including that used for the cooling and auxiliary equipment) was 0.5 % of total electricity consumption in the world. The percentage was doubled to 1 % in the year of 2005, as shown in Figure 1.1(a), [1] and roughly octupled to 3.7 % (~195 TWh) in the year of 2008 [2]. This represents a significant increase over the past decade, considering the fact that it is now comparable to the total electricity consumption of Australia (~220 TWh), as shown in Figure 1.1(b) [2].

Since the integrated circuit (IC) was invented, continual improvements in IC performance and cost per function have enabled more capable and affordable electronic products. However, they are no longer achievable today without increasing power density (Figure 1.2), as traditional Complementary-Metal-Oxide-Semiconductor (CMOS) technology scaling has essentially ended.

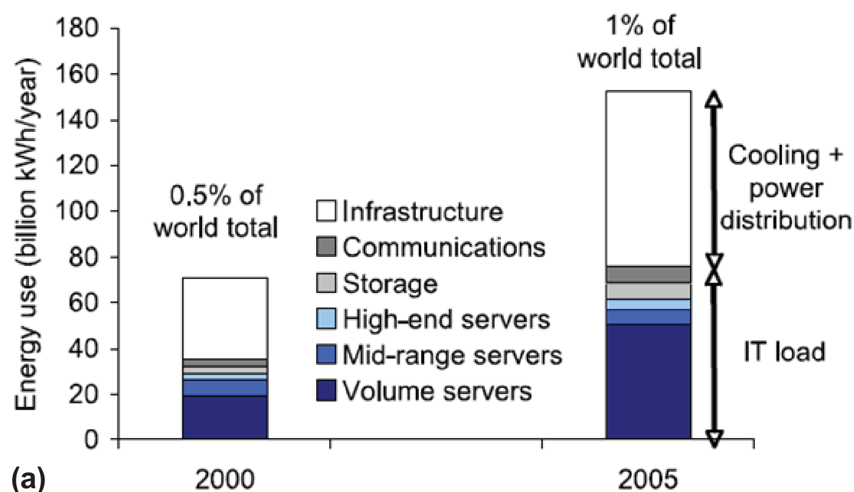


Figure 1.1: (a) Adapted from [1]: Total electricity usage of all data centers in the world in 2000 *vs.* 2005 (the electricity used for cooling and auxiliary equipment is included). Note that total electricity consumption in the world was increased from 13,238 TWh in 2000 to 15,747 TWh in 2005.

Figure 1.1: (b) Adapted from [2]: In 2008, the total power consumed for information communication technology (822 TWh) represents 3.7 % of the world total power consumption. The electricity usage of the data centers around the world (195 TWh) is comparable to that of Australia (220 TWh).

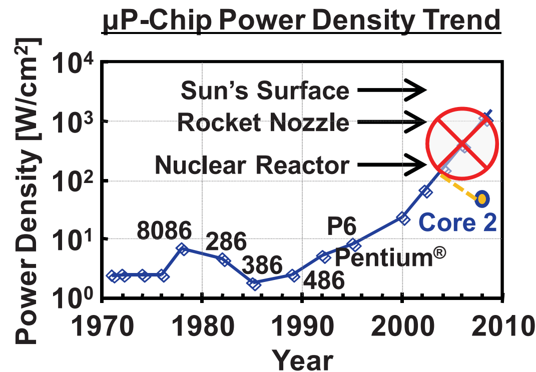


Figure 1.2: Courtesy of Shekhar Y. Borkar at Intel, Corp. Today, typical CMOS microprocessors operate at around the power density of a nuclear reactor, *e.g.* 65 Watts and 130 Watts for Intel® Core™2 Duo E6320 (die size of 143 mm^2) and Intel® Core™ i7-990X (die size of 239 mm^2), respectively [3].

At the 130 nm CMOS technology node and beyond, the power-supply voltage (V_{DD}) at which the CMOS chip operates has not been reduced as aggressively as the device pitch (Figure 1.3(a)), *e.g.* V_{DD} = 1.3 V [5] *vs.* 1 V [6] for the 130 and 32 nm nodes, respectively. This is because the threshold (V_{TH}) of a CMOS transistor (*i.e.* the voltage at which the transistor turns *on* and *off*) cannot be scaled down any lower, without incurring significant increase in transistor *off*-state leakage and hence increase in static power dissipation (Figure 1.3(b)).

As power consumption (and resulting heat generation) becomes more exacerbated with newer generations of the CMOS chip, parallel-computing has become widely-employed in microprocessors [8]. For example, a dual-core microprocessor has two parallel processing units, each of which is designed to run at a slower clock speed, in order to reduce its energy per operation; the overall system performance can be recovered by operating both of the processors in parallel (Figure 1.4).

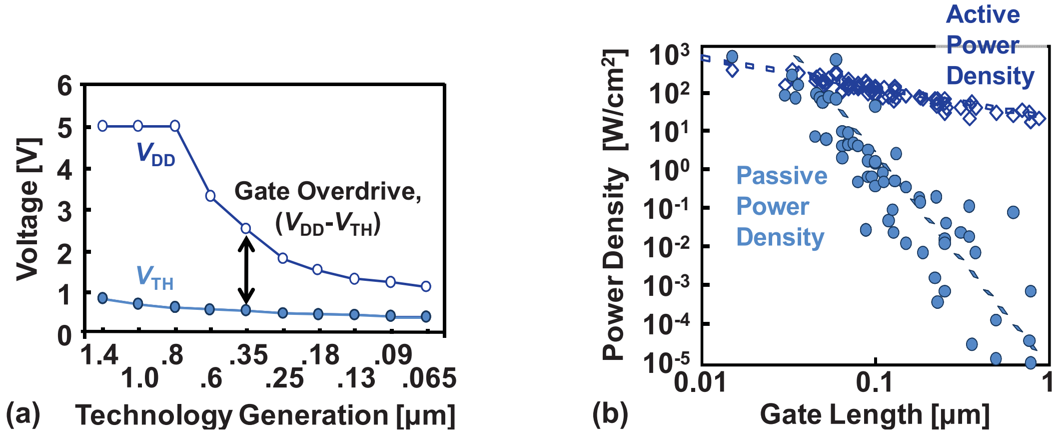


Figure 1.3: (a) Adapted from [4]. Scaling of V_{DD} and V_{TH} with advances in CMOS technology. (b) Adapted from [7]. As the gate length (and hence density) of transistors has decreased, the static power density has increased even faster (due to an exponential increase in transistor *off*-state leakage) than the dynamic power density has increased.

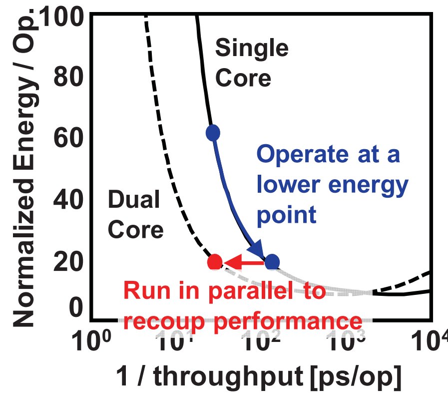


Figure 1.4: A single-core microprocessor that runs at a lower clock speed can be made to operate at lower energy per operation. Two single-core processors can be run in parallel to recover the overall system performance.

The paradigm shift to multi-core computing, however, inherently has a fundamental problem at the device-level. Take the CMOS transistor, for instance. As V_{DD} of a transistor is reduced, the active energy consumed to switch capacitors (E_{ACTIVE}) decreases at the expense of a longer time to perform logical operations (t_{DELAY}):

$$E_{\mathrm{ACTIVE}} = \alpha \cdot f \cdot L_{\mathrm{D}} \cdot C \cdot V_{\mathrm{DD}}^2,$$

$$t_{\mathrm{DELAY}} = \frac{f \cdot L_{\mathrm{D}} \cdot C \cdot V_{\mathrm{DD}}}{2I_{\mathrm{ON}}},$$

where α is the activity factor, f is the fan-out, L_{D} is the logic depth, C is the capacitance per stage, and I_{ON} is the on-state current. The V_{DD} reduction, however, forces the transistor to dissipate even higher passive energy (E_{PASSIVE}) due to increasing *off*-state leakage (I_{OFF}):

$$E_{\mathrm{PASSIVE}} = f \cdot L_{\mathrm{D}} \cdot I_{\mathrm{OFF}} \cdot V_{\mathrm{DD}} \cdot t_{\mathrm{DELAY}}.$$

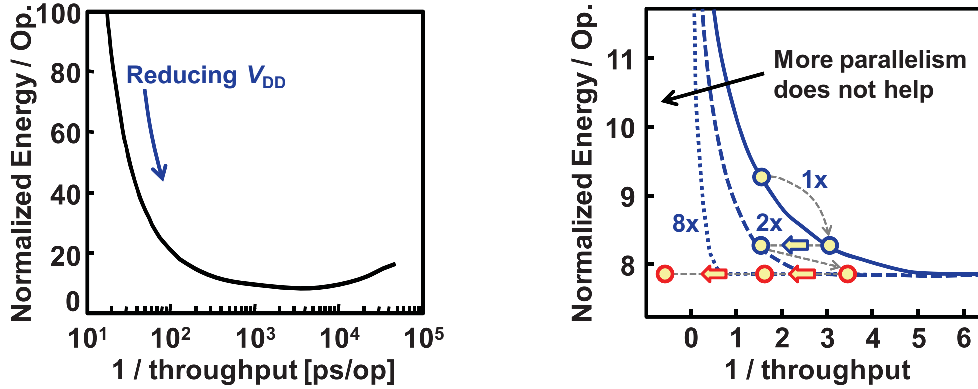


Figure 1.5: (a) As V_{DD} is decreased, E_{ACTIVE} decreases, t_{DELAY} increases, and $E_{PASSIVE}$ increases. Therefore, any CMOS technology has a fundamental energy-efficiency limit due to transistor *off*-state leakage. (b) Once the energy-efficiency limit is reached, the parallel computing methodology would be no longer beneficial for improving power density.

As a result, the minimum point where E_{ACTIVE} and $E_{PASSIVE}$ balance each other out yields the fundamental energy-efficiency limit, as shown in Figure 1.5(a).

The parallel computing methodology would eventually become ineffective, because addition of a processor core (that operates albeit at lower throughput) would not lead to any lower energy consumption per logic operation, as Figure 1.5(b) shows. This will necessitate alternate device technologies that surmount the fundamental CMOS energy-efficiency limit and hence enable emerging, ultra-low-power IC applications. To that end, a nano-electro-mechanical (NEM) relay technology is the closest to ideal, because of its immeasurably low *off*-state leakage current and abrupt turn-*on* behavior, which provide for zero static power consumption and potentially very low dynamic power consumption. A scaled relay, in principle, can be implemented to switch *on* or *off* at very low voltages (< 100 mV) and hence to have a higher energy-efficiency (~10× or more) than a similarly-sized CMOS transistor [9].

Figure 1.6(a) shows a simplified diagram of a mechanical relay: In the *off*-state, no current can flow because of the air gap that separates the source electrode from the drain electrode, as

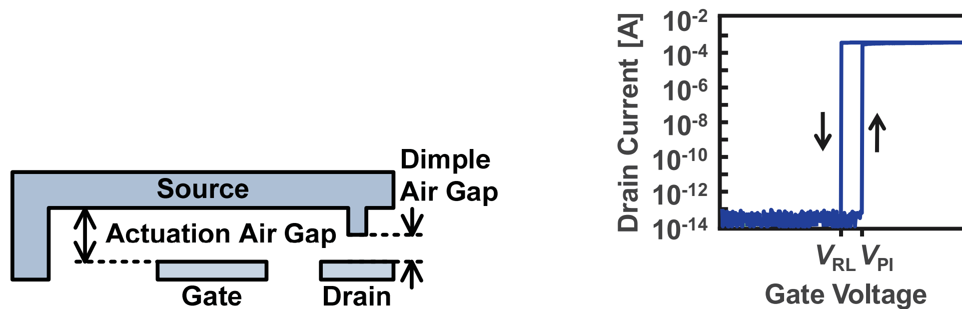


Figure 1.6: (a) Diagram of a three-terminal mechanical relay. The air gap between the source and drain electrodes is what leads to zero *off*-state leakage. The electrical contact between the source and drain is made by the electrostatic force induced between the gate and source electrodes. (b) Measured I_{DS}-V_{GS} characteristic of a relay fabricated in the Berkeley Microfabrication Laboratory. The relay turns *on* or *off* abruptly with a small change in the gate voltage.

confirmed by measured I_{DS}-V_G characteristic in Figure 1.6(b). The relay turns *on* abruptly (less than 0.1 mV/decade) when the electrostatic force induced between the gate and source electrodes brings the source into contact with the drain.

This Ph.D. dissertation aims to address challenges for achieving a NEM relay technology that can provide a potential pathway for achieving truly "greener" computing via reducing carbon footprint of information-processing systems. In particular, relay design and process technology improvements, which led to the successful demonstration of relay-based digital IC building blocks, are discussed from both device- and circuit-level perspectives. In addition, multi-electrode relays that can lead to more efficient design and implementation of ultra-low-power digital electronics are discussed.

In Chapter 2, improvements to the prototype four-terminal (4-T) relay design are investigated to resolve the parasitic issues inherent in the original design, namely a relatively weak body bias effect, undesirable influence of source/drain bias voltages on the gate switching voltages, and parasitic actuation by channel-to-body potential.

In Chapter 3, an electrostatically actuated 4-T relay is demonstrated to function as a memory cell. This indicates promise for embedding non-volatile memory with relay-based logic circuits, without incurring additional fabrication cost.

In Chapter 4, a see-saw relay design is proposed and demonstrated to provide for perfectly complementary switching behavior that is symmetric about V_{DD} / 2. Also, the experimental results demonstrate that the complementary see-saw relay design is versatile for mechanically implementing various logic functions and a latch function with a single structure.

In Chapter 5, multi-gate, multi-source/drain relays are proposed and demonstrated, in order to provide an evolutionary pathway for compact implementation of ultra-low-power digital electronics (*e.g.* table-top supercomputer).

In Chapter 6, the key results and contributions of this work are summarized, and future research directions are briefly suggested.

References

[1] J. G. Koomey, "Worldwide Electricity Used in Data Centers," *IOPscience Environmental Research Letters*, vol. 3, no. 3, pp. 034008, September 2008.

[2] "Samsung Green Memory," *Samsung Semiconductor Chief Information Officer Forum (SSCF ' 10)*, [Online] Available: http://www.samsung.com/global/business/semiconductor/Greenmemory/Downloads/Documents/Documents_PT.html, November 2010.

[3] Intel® Processors -- Specifications, [Online] Available: http://ark.intel.com/Default.aspx

[4] P. Packan, "Device and Circuit Interactions," in *IEEE International Electron Device Meeting (IEDM '07) Short Course: Performance Boosters for Advanced CMOS Devices*, December 2007.

[5] S. Tyagi, M. Alavi, R. Bigwood, T. Bramblett, J. Brandenburg, W. Chen, B. Crew, M. Hussein, P. Jacob, C. Kenyon, C. Lo, B. Mcintyre, Z. Ma, P. Moon, P. Nguyen, L. Rumaner, R. Schweinfurth, S. Sivakumar, M. Stettler, S. Thompson, B. Tufts, J. Xu, S. Yang, and M. Bohr, "A 130 nm Generation Logic Technology Featuring 70 nm Transistors, Dual V_t Transistors and 6 layers of Cu Interconnects," in *IEEE International Electron Device Meeting (IEDM '00) Technical Digest*, pp. 567-570, December 2000.

[6] P. Packan, S. Akbar, M. Armstrong, D. Bergstrom, M. Brazier, H. Deshpande, K. Dev, G. Ding, T. Ghani, O. Golonzka, W. Han, J. He, R. Heussner, R. James, J. Jopling, C. Kenyon, S-H. Lee, M. Liu, S. Lodha, B. Mattis, A. Murthy, L. Neiberg, J. Neirynck, S. Pae, C. Parker, L. Pipes, J. Sebastian, J. Seiple, B. Sell, A. Sharma, S. Sivakumar, B. Song, A. St. Amour, K. Tone, T. Troeger, C. Weber, K. Zhang, Y. Luo, and S. Natarajan, "High Performance 32 nm Logic Technology Featuring 2nd Generation High-k + Metal Gate Transistors," in *IEEE International Electron Device Meeting (IEDM '09) Technical Digest*, pp. 659-662, December 2009.

[7] B. Meyerson, *Semico Impact Conference*, Taiwan, January 2004.

[8] K. Asanović, R. Bodik, B. C. Catanzaro, J. J. Gebis, P. Husbands, K. Keutzer, D. A. Patterson, W. L. Plishker, J. Shalf, S. W. Williams, and K. A. Yelick, "The Landscape of Parallel Computing Research: A View from Berkeley," *Technical Report No. UCB/EECS-2006-183*, Electrical Engineering and Computer Sciences, University of California, Berkeley, December 2006.

[9] F. Chen, H. Kam, D. Marković, T.-J. King Liu, V. Stojanović, and E. Alon, "Integrated Circuit Design with NEM relays," in *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD '08)*, pp. 750-757, December 2008.

Chapter 2

Four-Terminal Relay Design Improvements

Abstract: The first prototype electrostatically actuated four-terminal (4-T) relays suffer from a relatively weak body bias effect, undesirable influence of source/drain bias voltages on the gate switching voltages, and parasitic actuation by channel-to-body potential. In this chapter, improvements to the 4-T relay design are investigated to resolve these issues.

2.1. Introduction

Transistor *off*-state leakage limits the degree to which the energy per operation can be lowered in Complementary-Metal-Oxide-Semiconductor (CMOS) digital integrated circuits (ICs) [1]. For this reason, nano-electro-mechanical (NEM) relays, which have zero *off*-state leakage, have been proposed for ultra-low-power digital IC applications [2]. Because a relay has abrupt switching characteristics (less than 0.1 mV/decade [3]), in principle it can be made to operate with lower supply voltage (V_{DD}) than CMOS transistors, for lower dynamic power consumption as well as zero static power consumption. Recently, a reliable micro-relay technology has been demonstrated [4], and a four-terminal (4-T) relay design has been introduced for digital logic applications [3]. The purpose of the fourth ("body") terminal is to provide a means to tune the gate switching voltages of the relay for low-voltage operation [3]. Basic digital building-block ICs, such as carry-generation circuit, oscillator, and adder, have been demonstrated with 4-T relays and passive circuit elements [5], [6], [7]. A scaled 4-T relay technology shows promise for breaking the lower energy limit for CMOS technology – by more than one order of magnitude – at throughputs up to 100 MOPS and comparable layout area [8].

2.2. 4-T Relay Design

The first prototype 4-T relay design reported in [3] (replicated in Figure 2.1) suffers from a relatively weak body bias effect and undesirable influence of source/drain bias voltages on the gate switching voltages, due to significant overlap between the actuated gate electrode and source/drain electrodes resulting in unwanted parasitic actuation effects. An improved 4-T relay design shown in Figure 2.2 is demonstrated to resolve these issues.

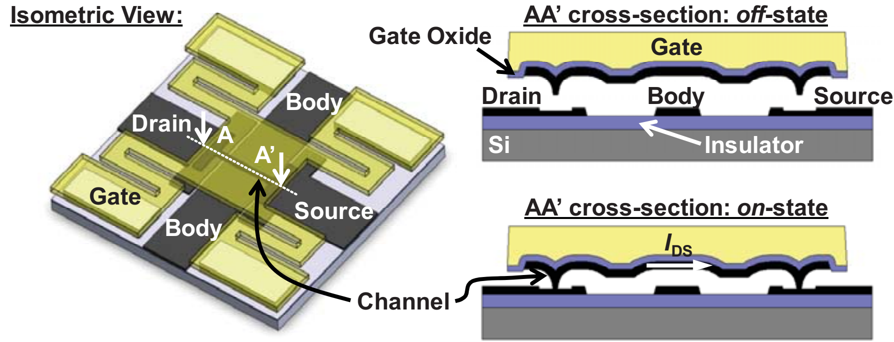


Figure 2.1: Isometric view of the prototype 4-T relay structure and schematic cross-sections in the *off*- and *on*-states [3]. In the *off*-state, no Direct-Current (DC) current can flow between the source and drain electrodes, because an air gap separates the channel from the source/drain electrodes. In the *on*-state, electrostatic force between the movable gate electrode and the underlying body electrode brings the channel into contact with the source and drain electrodes.

The device operating principle is the same as for the first prototype 4-T relay design described in [3]: As the voltage difference between the movable p+ polycrystalline-$Si_{0.4}Ge_{0.6}$ electrode and the underlying Tungsten (W) actuation electrode is increased above the turn-*on* threshold voltage, the electrostatic force becomes sufficiently large to bring the W channel layer into contact with the W source and drain electrodes (which are coplanar with the actuation electrode) so that current can flow between these two electrodes; the gate voltage corresponding to the turn-*on* threshold voltage is called the "pull-in" voltage, V_{PI}.

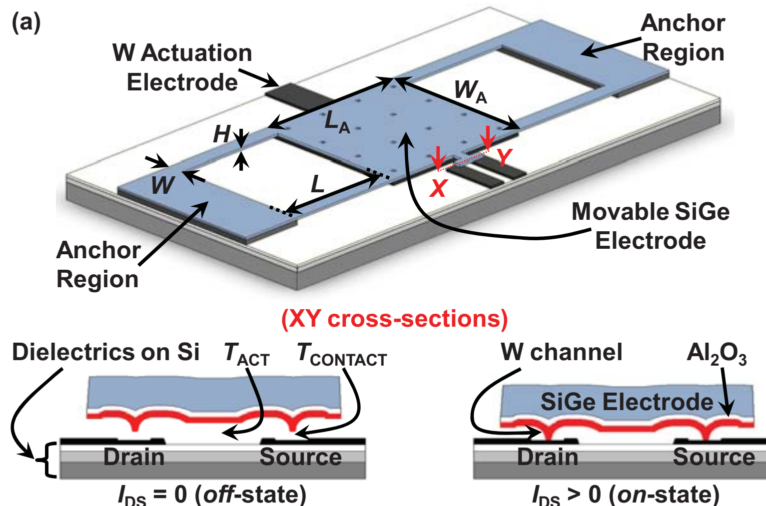
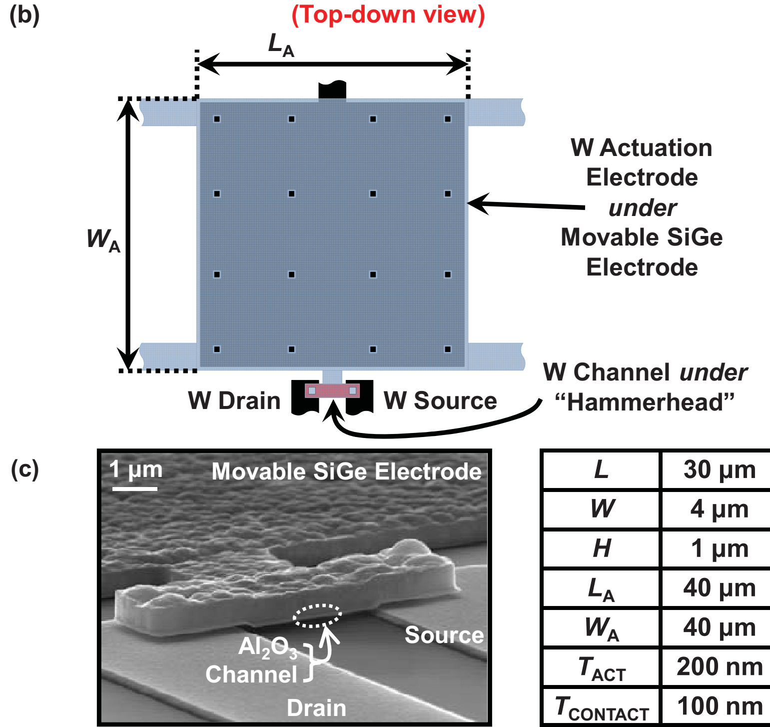


Figure 2.2: (a) Isometric view of the improved 4-T relay structure and schematic cross-sections of the channel region in the *off*- and *on*-states, showing the various relay design parameters. The W channel layer is attached underneath the poly-$Si_{0.4}Ge_{0.6}$ only in the "hammerhead" portion (7.5 μm by 2 μm). In the *off*-state, the channel layer is separated from the source/drain electrodes by an air gap, so that no current can flow between the source and drain electrodes. In the *on*-state, the moveable structure is actuated downward (by electrostatic force between the movable SiGe electrode and the underlying W actuation electrode) so that the channel contacts the source and drain electrodes.



L	30 µm
W	4 µm
H	1 µm
L_A	40 µm
W_A	40 µm
T_{ACT}	200 nm
$T_{CONTACT}$	100 nm

Figure 2.2: (b) Top-down view of the 4-T relay structure. Note that 1 µm^2 etch holes are patterned onto the gate stack, in order to ensure release of the structure by vapor-phase HF as well as to reduce residual stress of the gate stack [9]; Due to fringing electric field, the impact of reduced actuation area on relay operating voltages is minimal [9], [10]. (c) Scanning electron micrograph of the hammerhead region of a fabricated 4-T relay. The separation between the source and drain electrodes is 3.5 µm. The design parameter values used in this work are listed in the table on the right.

As the voltage difference between the poly-$Si_{0.4}Ge_{0.6}$ electrode and the underlying actuation electrode is decreased back down to be below the turn-*off* threshold voltage, the spring restoring force of the poly-$Si_{0.4}Ge_{0.6}$ folded-flexures becomes sufficiently large to break the contact between the channel and source/drain electrodes so that no current can flow; the gate voltage corresponding to the turn-*off* threshold voltage is called the "release" voltage, V_{RL}. Note that a gate-insulating layer (40 nm of Al_2O_3) electrically isolates the poly-$Si_{0.4}Ge_{0.6}$ electrode from the channel to prevent DC current flow between the poly-$Si_{0.4}Ge_{0.6}$ and the source/drain electrodes when the relay is in the *on*-state.

2.2.1. Relay Fabrication Process

Improved 4-T relays were fabricated in the Berkeley Microfabrication Laboratory using the four-mask process illustrated in Figure 2.3. Note that all photolithography steps were performed using a 5X Chrome reticle (6" X 6" X 0.25") with an anti-reflective coating in an ASML Deep-Ultra-Violet (DUV) 5500/90 stepper that can resolve 250 nm features.

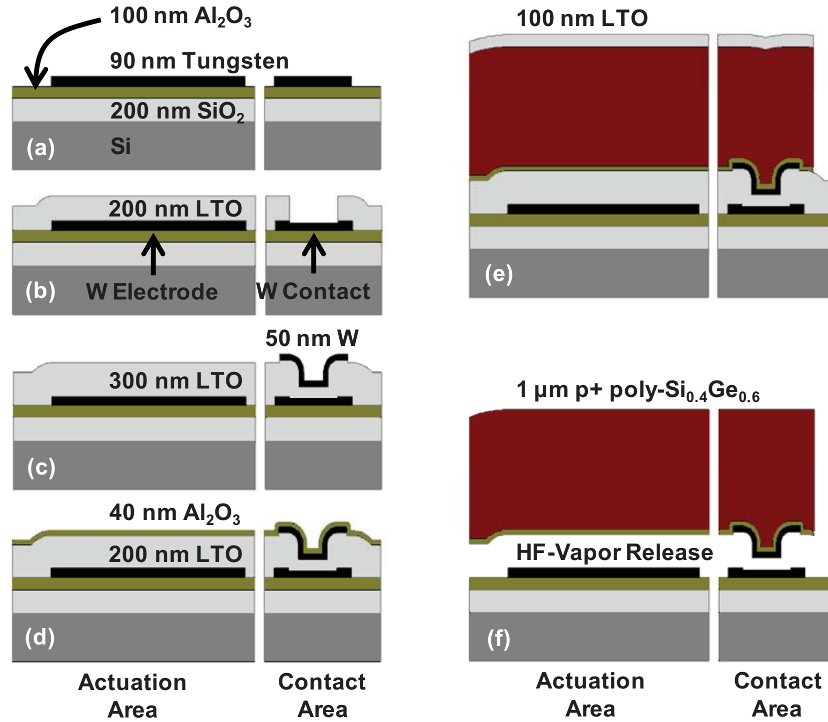


Figure 2.3: Four-mask fabrication process for the improved 4-T relay. (a) W body, drain, and source electrodes formed on the Al_2O_3-coated substrate. (b) 1st sacrificial low-temperature oxide (LTO) layer deposition, followed by contact area definition. (c) 2nd sacrificial LTO layer deposition, followed by W channel formation. (d) Gate oxide (Al_2O_3) deposition. (e) Gate layer (p+ poly-$Si_{0.4}Ge_{0.6}$) deposition followed by LTO deposition and gate patterning. (f) Release etching in vapor-HF.

A 200-nm-thick substrate-insulating layer of Silicon Dioxide (SiO_2) was thermally grown at 1000 ºC in steam ambient on 150-mm-diameter Boron-doped p-type (1-0-0) low-resistivity (about 20 Ω·cm) Silicon wafers. An Aluminum Oxide (Al_2O_3) layer deposited using an atomic layer deposition (ALD) reactor (Picosun SUNALE™ R-150) was used for the substrate-coating (100 nm) and gate-dielectric (40 nm) layers. The ALD processing temperature of 300 °C used to yield relatively low tensile stress (about 200 MPa) within the deposited film leads to a growth rate of approximately 0.9 nm per cycle [11], [12] and film thickness uniformity less than ± 1 % across the wafer and wafer-to-wafer. The pulsing and purging times of both Trimethylaluminum (TMA) and H_2O precursors were set to 0.1 and 4 seconds, respectively, in order to allow sufficient time for the self-terminating gas-solid reaction [13]. This Al_2O_3 layer can serve as an excellent substrate insulator, because it is highly resistant to HF-vapor release etching [14] and is known to provide good adhesion to W without any intermediate glue layer [15], [16]. In addition, ALD Al_2O_3 is a good gate-insulating material in that it shows reasonably low gate leakage and relatively high dielectric breakdown strength (larger than 5 MV/cm); For example, a 12-nm-thick ALD Al_2O_3 layer shows gate leakage currents less than 1 nA/cm^2, when an electric field of 2 MV/cm is applied [17].

The W material deposited by DC magnetron sputtering (base pressure ≈ $1 \cdot 10^{-7}$ Torr, process pressure ≈ 10 mTorr, and DC power = 2 kW) was used as the source, drain, body, and channel material. It has excellent resistance to physical wear due to its relatively high hardness [18]; Vickers hardness of Tungsten (approximately 3.42 GPa) is higher than that of most metals shown in the periodic table (*e.g.* ~0.17 GPa for Aluminum, ~0.87 GPa for Tantalum, and ~0.97 GPa for Titanium) [19], [20]. In addition, it has high resistance to vapor-phase HF release etching [21]. The SF_6 dry etch plasma chemistry (85 sccm of SF_6, chamber pressure = 10 mTorr, bias peak power = 100 W, and source peak power = 1000 W) used to pattern the W film provides an etch selectivity to the underlying Al_2O_3 substrate insulator (greater than 100:1) and to the sacrificial SiO_2 (about 3:1).

Undoped low-temperature-oxide (LTO deposited at 400 °C and 300 mTorr), about 200 nm thick in the actuation region and about 100 nm thick in the contact regions, was used as the sacrificial material. Note that the sacrificial oxide thickness translates to an *air* gap in the actuation and contact regions of the relay, upon removal of the sacrificial oxide. (This *air* gap formation is what leads to zero *off*-state leakage current.)

In-situ boron-doped p-type poly-$Si_{0.4}Ge_{0.6}$ deposited by low-pressure chemical vapor deposition (LPCVD) at 410 °C and 600 mTorr was used as the structural gate material. A thin layer (less than 5 nm) of amorphous Silicon (α-Si) was deposited in between the poly-$Si_{0.4}Ge_{0.6}$ and the sacrificial oxide (by flowing Disilane at 410 °C and 300 mTorr). This layer not only promotes the adhesion of the poly-$Si_{0.4}Ge_{0.6}$ to the underlying gate oxide, but also seeds film growth [22].

Prior to release etching, a Xenon Dichloride ($XeCl_2$) excimer laser (λ = 308 nm) pulse (400 mJ/cm^2, 38 ns) was applied to the poly-$Si_{0.4}Ge_{0.6}$ gate electrode through a 100-nm-thick SiO_2 hardmask layer, in order to reduce the impact of large negative stress gradient within the film

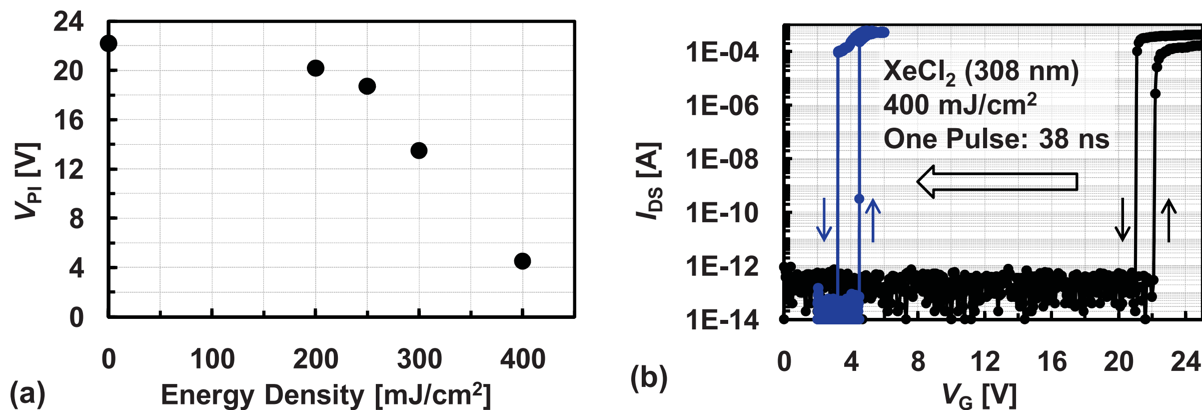


Figure 2.4: (a) The V_{PI} values (measured at room temperature under N_2 purge at atmospheric pressure) decrease with increasing energy density of a single laser pulse (38 ns). (b) Measured I_{DS}-V_G characteristics of the fabricated 4-T relay with *vs.* without the application of the laser pulse. V_D, V_B, and V_S were biased at 0.5 V, 0 V, and 0 V, respectively. Note that negative strain gradient within the poly-$Si_{0.4}Ge_{0.6}$ was reduced substantially, resulting in smaller T_{ACT} and hence lower actuation voltages.

[23], [24] and hence achieve lower operating voltages. This negative stress gradient causes the top portion of the poly-$Si_{0.4}Ge_{0.6}$ film to expand more than the less-compressive (or more-tensile) bottom portion, *i.e.* it induces negative strain gradient through the film thickness. The movable structure is gradually buckled upward from its anchors toward its geometric center (it appears like an open umbrella), and as a result, T_{ACT} is rendered non-uniform across the actuation area (being the largest at the center).

As shown in Figure. 2.4, measured V_{PI} and V_{RL} of the laser-annealed relays decrease with increasing energy density of the laser pulse, because of reduction in effective T_{ACT}. Note that the negative strain gradient effect is diminished, because compressive stress within the top portion is relieved [22], [23].

Finally, the dies were released in vapor-phase HF -- the surface temperature of the dies was kept to 50 ºC during the release etch step, so that the reaction by-product (H_2O) can vaporize more quickly, thereby preventing release stiction.

2.2.2. Dependence of Gate Switching Voltage on Body Bias

Measured I_{DS}-V_G characteristics of a 4-T relay are shown in Figure 2.5, for operation with the poly-$Si_{0.4}Ge_{0.6}$ electrode as the gate and the underlying W electrode as the body. Note that V_{PI} can be reduced very effectively by adjusting the body bias, in order to allow for operation with reduced gate-voltage swing.

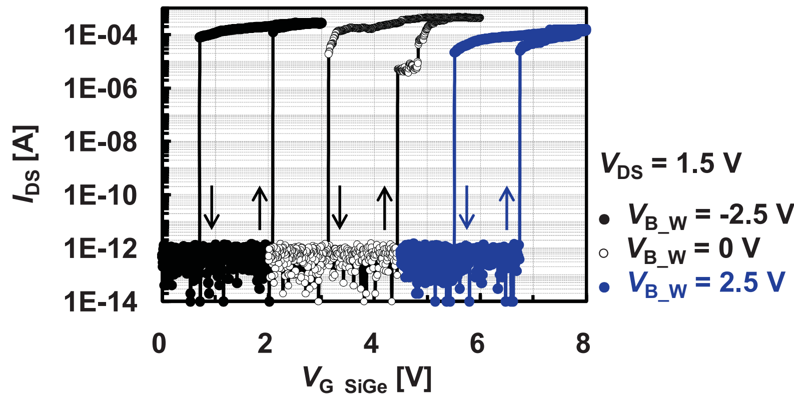


Figure 2.5: Measured I_{DS}-V_G characteristics of a fabricated 4-T relay at room temperature under N_2 purge at atmospheric pressure, for various body biases. Zero *off*-state leakage and abrupt switching behavior (less than 1 mV/decade) are observed. The *on*-state resistance (R_{ON}) values are 5.5 kΩ, 3.3 kΩ, and 9.1 kΩ at $V_{G_SiGe} = 1.2 \cdot V_{PI}$ for V_{B_W} of -2.5 V, 0 V, and 2.5 V, respectively. The hysteretic switching behavior is due to pull-in mode operation (because $T_{CONTACT} > T_{ACT} / 3$) as well as surface adhesion forces which increase with decreasing R_{ON}. The variation in R_{ON} is partly attributed to oxidation of the channel and source/drain electrode surfaces over time. The measurements were made firstly for V_{B_W} = 0 V, then for V_{B_W} = -2.5 V, finally for V_{B_W} = 2.5 V.

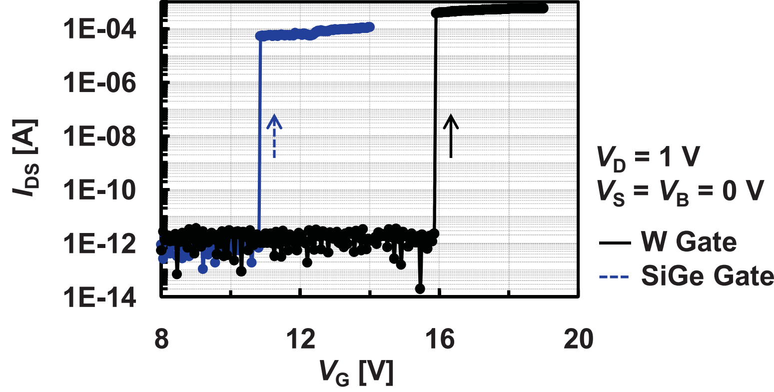


Figure 2.6: V_{PI} for operation with the W electrode as the gate (15.9 V) is higher (by ~47 %) than that for operation with the poly-$Si_{0.4}Ge_{0.6}$ electrode as the gate (10.85 V). The data were taken at room temperature under N_2 purge at atmospheric pressure.

Ideally, a change in the V_B should result in commensurate changes to the gate switching voltages V_{PI} and V_{RL}. Also, the V_{PI} and V_{RL} values for operation with the poly-$Si_{0.4}Ge_{0.6}$ electrode as the gate (and the W electrode as the body) should be the same as for operation with the W electrode as the gate (and the poly-$Si_{0.4}Ge_{0.6}$ as the body), to allow for versatility in digital circuit design [3]. Neither of these was found to be the case for the first prototype 4-T relay design: a 1 V change in V_B resulted in an about 0.5 V change in V_{PI} [3], and V_{PI} for operation with the W electrode as the gate is considerably higher (by approximately 50 %) than for operation with the poly-$Si_{0.4}Ge_{0.6}$ electrode as the gate (Figure 2.6), due to the relatively small percentage of the actuated region of the poly-$Si_{0.4}Ge_{0.6}$ electrode that overlaps with the underlying W electrode. To address these issues, the percentage overlap between the actuated region of the poly-$Si_{0.4}Ge_{0.6}$ electrode and the underlying W electrode is increased to be nearly 100 % in the improved 4-T relay design. Note that the area of the overlap between the poly-$Si_{0.4}Ge_{0.6}$ and the W source/drain electrodes is reduced to 8 μm^2, which is much smaller than the poly-$Si_{0.4}Ge_{0.6}$-to-W overlap area of about 1600 μm^2. Since the capacitance between the poly-$Si_{0.4}Ge_{0.6}$ and the source/drain

Table 2.1: Calculated parallel-plate capacitance values for the first prototype and improved 4-T relay designs. In the improved design, parasitic electrostatic force in the source/drain regions is relatively small, because the *on*-state capacitance between the poly-$Si_{0.4}Ge_{0.6}$ and source/drain electrodes is much smaller than that between the poly-$Si_{0.4}Ge_{0.6}$ and W electrodes. Note that the ratio of the *on*-state to *off*-state capacitances is approximately 2.49 and 55.58 for the prototype and improved designs, respectively.

Parallel-Plate Capacitance	Prototype 4-T Design		Improved 4-T Design	
	$Si_{0.4}Ge_{0.6}$ to W	$Si_{0.4}Ge_{0.6}$ to Source/Drain	$Si_{0.4}Ge_{0.6}$ to W	$Si_{0.4}Ge_{0.6}$ to Source/Drain
***off*-state**	**19.57 fF**	**13.39 fF**	**66.13 fF**	**0.43 fF**
***on*-state**	**38.45 fF**	**44.43 fF**	**3703.41 fF**	**5.38 fF**

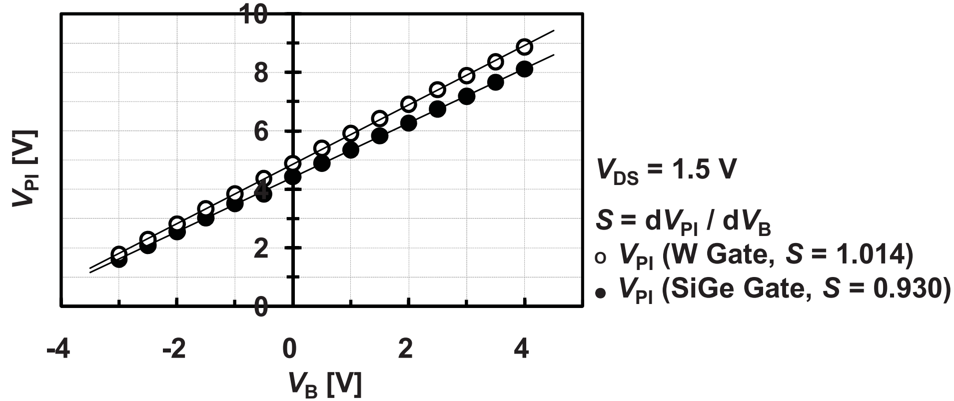


Figure 2.7: Measured gate pull-in voltages for operation with the $Si_{0.4}Ge_{0.6}$ electrode as the gate and for operation with the W electrode as the gate, as a function of the body bias voltage. V_{PI} can be linearly tuned by changing V_B.

electrodes is much smaller than that between the poly-$Si_{0.4}Ge_{0.6}$ and the W electrode (as shown in Table 2.1), parasitic electrostatic force in the source/drain regions is relatively small. As a result, the gate switching voltage changes directly with body bias (*i.e.* $dV_{PI} / dV_B \cong 1$), as shown in Figure 2.7.

The V_{PI} and dV_{PI} / dV_B values for operation with the poly-$Si_{0.4}Ge_{0.6}$ electrode as the gate are slightly lower than for operation with the W electrode as the gate, because of a small amount of parasitic actuation between the poly-$Si_{0.4}Ge_{0.6}$ electrode and the source/drain electrodes. Specifically, the effective actuation area for operation with the poly-$Si_{0.4}Ge_{0.6}$ electrode as the gate is slightly larger than for operation with the W electrode as the gate, so that V_{PI} is slightly lower for operation with the poly-$Si_{0.4}Ge_{0.6}$ electrode as the gate. However, this also means that the percentage overlap between the gate and body electrodes is slightly smaller for operation with the poly-$Si_{0.4}Ge_{0.6}$ electrode as the gate, so that dV_{PI} / dV_B is slightly smaller for operation with the poly-$Si_{0.4}Ge_{0.6}$ electrode as the gate.

2.2.3. Dependence of Gate Switching Voltage on Drain Bias

Ideally, a change in the V_D should not affect the gate switching voltages. This was not the case for the first prototype 4-T relay design [3]: a 1 V change in V_D resulted in an approximately 0.25 V change in V_{PI}, due to significant overlap (roughly 37 %) between the actuated region of the poly-$Si_{0.4}Ge_{0.6}$ electrode and the source/drain electrodes. Since this overlap area is dramatically reduced (to about 0.5 %) in the improved 4-T relay design, V_{PI} does not change significantly with V_D, as shown in Figure 2.8. $dV_{PI} / dV_D < 0$ for operation with the W electrode as the gate, because the drain electrode is located on the same side of the actuation gap as the gate in this case, so that an increase in V_D helps to actuate the poly-$Si_{0.4}Ge_{0.6}$ electrode downward. $dV_{PI} / dV_D > 0$ for operation with the poly-$Si_{0.4}Ge_{0.6}$ electrode as the gate, because the drain electrode is located on

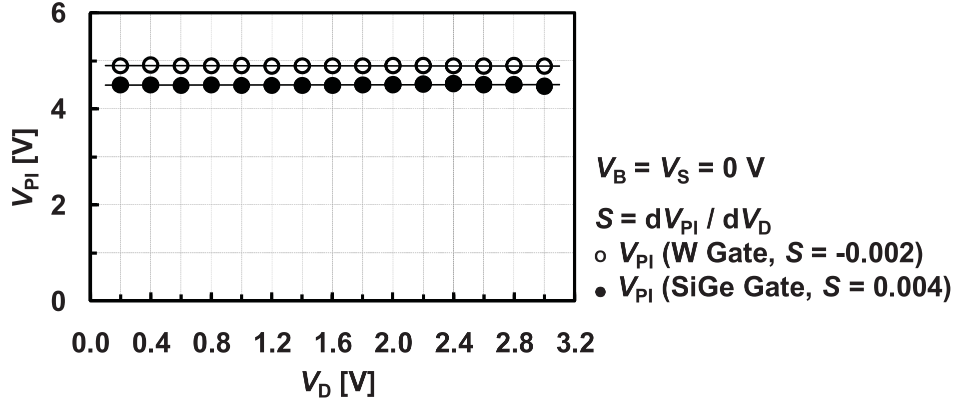


Figure 2.8: The measured V_{PI} values show negligible dependence on the drain. All measurements were made at room temperature under N_2 purge at atmospheric pressure.

the opposite side of the actuation gap as the gate in this case, so that an increase in V_D reduces the electrostatic force exerted on the poly-$Si_{0.4}Ge_{0.6}$ electrode.

2.2.4. Parasitic Actuation by Channel-to-Body Potential

In the prototype 4-T relay, the W channel interposed between the gate and body electrodes is unnecessarily large [3]; the area of overlap between the channel and body electrodes (170 μm^2) is comparable to that between the gate and body (450 μm^2). This is undesirable, because the parasitic electrostatic force is induced between the channel and body electrodes, when a voltage (V_{DD}) is applied between them. This force keeps the relay in the *on*-state (*i.e.* causing in-use stiction), even if the gate-to-body voltage is no longer applied. In the improved 4-T relay design, therefore, the W channel layer was removed from the gate-to-body overlap area to remove parasitic actuation. (as Figure 2.2(a) shows, it is attached underneath the poly-$Si_{0.4}Ge_{0.6}$ only in the hammerhead portion.)

2.3. Revised 4-T Relay Design

Experimental measurements of the improved 4-T relays are the basis for the revised 4-T relay layout shown in Figure 2.9. In particular, the improved 4-T relay design is mechanically-refined, such that the cantilever springs are now replaced with the symmetric flexure-beam springs to relieve growth stress build-up within the poly-$Si_{0.4}Ge_{0.6}$ microstructures [3], [4], [26]. To ensure good electrostatic integrity (*i.e.* symmetric actuation, no dependence of V_{PI} on source/drain bias, and a commensurate change to V_{PI} with a change in body-bias), the revised 4-T relay is designed such that the gate-to-source/drain capacitance is less than 2.5 % of the gate-to-body capacitance. Note that the channel is still interposed between the gate and body electrodes; This is to maintain

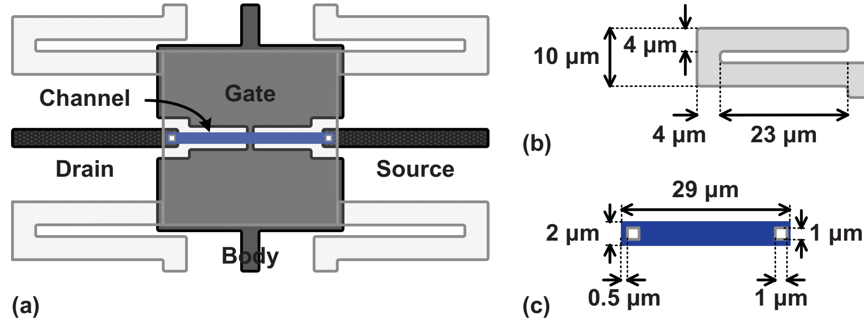


Figure 2.9: (a) Top-down view of the revised 4-T relay layout. Release etch holes are not shown. The minimum spacing between the source/drain and body electrodes is conservatively set to 2 μm to prevent surface leakage. The gate-to-body cut-out is set to 1 μm to allow for a sufficient photolithographic alignment tolerance. The W channel is enclosed by the gate electrode with minimum boundary dimension of 0.5 μm. The channel-to-body overlap is minimized to prevent in-use stiction, due to the unwanted electrostatic force between the channel and body. (b) Dimensions of the flexure-beam support. The effective spring constant of the flexure-beams (1-μm-thick poly-$Si_{0.4}Ge_{0.6}$ with Young's modulus of 140 GPa [22]) calculated using the flexure-beam stiffness equation in [25] is approximately 168 N/m. (c) Dimensions of the W channel. Each contact is enclosed by the channel with boundary dimension of 0.5 μm. The contact dimensions are limited to 1 μm; This is conservative, considering that the minimum feature size that the photolithography tool (ASML DUV 5500/90 Stepper) can resolve is 250 nm.

the original device footprint that was used for the demonstration of the relay-based IC building blocks [5]. To allow only minimal parasitic actuation, therefore, the channel-to-body overlap area is minimized to 2 μm^2 (*c.f.* gate-to-body overlap area is about 750 μm^2). In addition, the as-fabricated contact-gap-to-actuation-gap thickness ratio ($T_{CONTACT}$ to T_{ACT}) was chosen to be 0.5. This ratio should be smaller than one third in order to minimize the hysteric switching behavior -- by allowing the relay to operate in non-pull-in mode. Otherwise, V_{DD} scaling is limited by the release voltage of the relay. To achieve most energy-efficient relay operation, however, the relay should operate in pull-in mode (with a $T_{CONTACT}$-to-T_{ACT} ratio of 0.7 to 0.8) [25].

To validate the new relay design, Finite-Element-Method (FEM) analysis was performed using a multi-physics simulation software (CoventorWare™ Designer [27] and Analyzer [28]). Figures 2.10(a) shows the simulated V_{PI} value of the 4-T relay for operation with the poly-Si electrode as the gate and the underlying W electrode as the body. Note that poly-Si (Young's modulus of 158 GPa) was used to model the poly-$Si_{0.4}Ge_{0.6}$ layer, as the latter material is unavailable in the Material Properties Database in CoventorWare™. The simulated V_{PI} value of the relay for operation with the poly-Si electrode as the gate is consistent with that for operation with the underlying W electrode as the gate, as shown in Figure 2.10(b).

Figure 2.11(a) shows V_{PI} of the relay for operation with the poly-Si electrode as the gate. Figure 2.11(b) confirms minimal dependence of the gate switching voltage of the revised 4-T relay upon

drain bias (dV_{PI} / $dV_D \approx 0$), and Figure 2.11(c) verifies effective adjustment of V_{PI} of the relay with body-bias (*i.e.* a commensurate change to V_{PI} with a change in body-bias).

Figure 2.12(a) shows the simulated V_{PI} value of the relay for operation with the poly-Si electrode as the gate. Figure 2.12(b) confirms that the relay operation is negligibly affected by the parasitic electrostatic force induced by the channel-to-body potential.

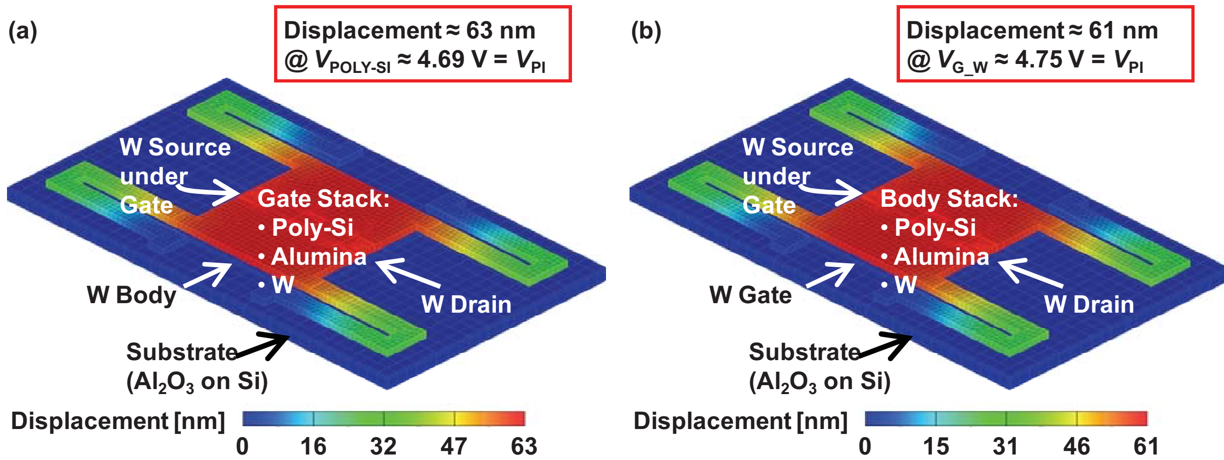


Figure 2.10: T_{ACT} = 200 nm, $T_{CONTACT}$ = 100 nm, T_{GOX} = 40 nm, $T_{ELECTRODE}$ = $T_{CHANNEL}$ = 50 nm, T_{SUB_OX} = 80 nm, and $H_{POLY-SI}$ = 1 µm are the vertical dimensions used for this work. The W channel and Si substrate were left floating. (a) The relay is actuated downward by ~63 nm (≈ $T_{CONTACT}$ / 3) at $V_{POLY-SI}$ of 4.69 V, right before it turns *on* abruptly. V_D, V_S, and V_{B_W} were set to 0 V. (b) When V_{G_W} is increased to about 4.75 V, the relay is on the verge of being pulled-in. $V_D = V_S = V_{POLY-SI}$ = 0 V.

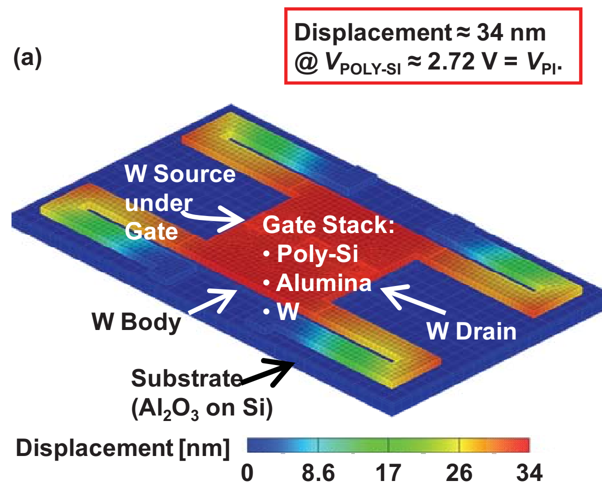


Figure 2.11: The vertical dimensions of the relay are T_{ACT} = 130 nm and $T_{CONTACT}$ = 90 nm; the other vertical dimensions are as described in Figure 2.9. Both the W channel and the Si substrate were left floating. The source was biased at 0 V. (a) The relay pulls-in at $V_{POLY-SI}$ = ~2.72 V. The relay is about to turn *on* at the downward displacement of ~34 nm (≈ T_{ACT} / 3) due to the pull-in mode operation. $V_D = V_{B_W}$ = 0 V.

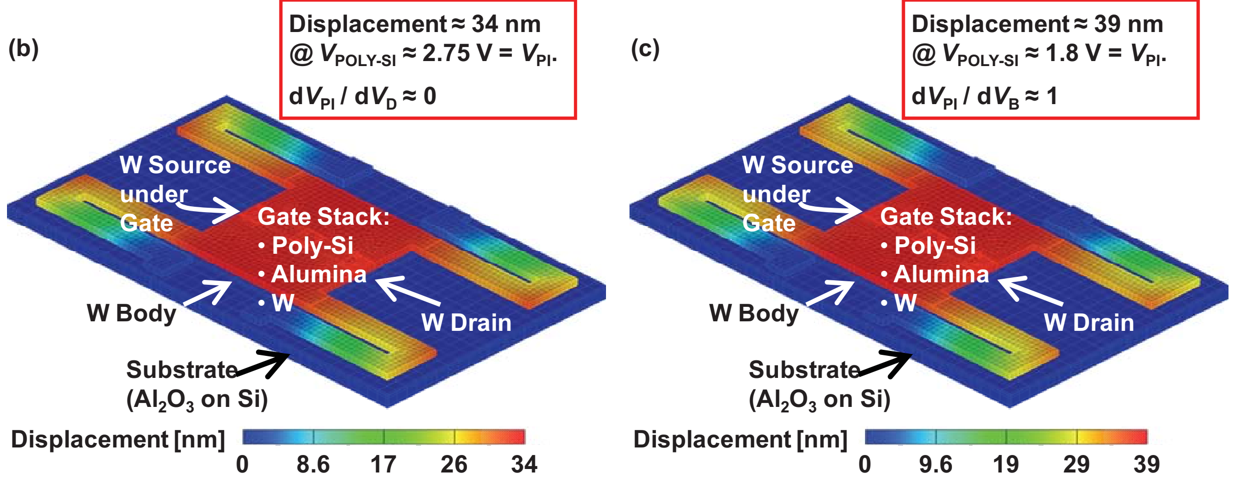


Figure 2.11: (b) V_D was increased to $V_{DD} = 1.2 \cdot V_{PI} \approx 3.4$ V. V_{B_W} remains at 0 V. The relay unconditionally snaps closed, when $V_{POLY-SI}$ is increased to ~2.75 V. (c) V_D was set equal to $V_{DD} = 1.2 \cdot V_{PI} \approx 2.3$ V. V_{B_W} = -1 V. The relay turns *on* abruptly at $V_{POLY-SI} \approx 1.8$ V.

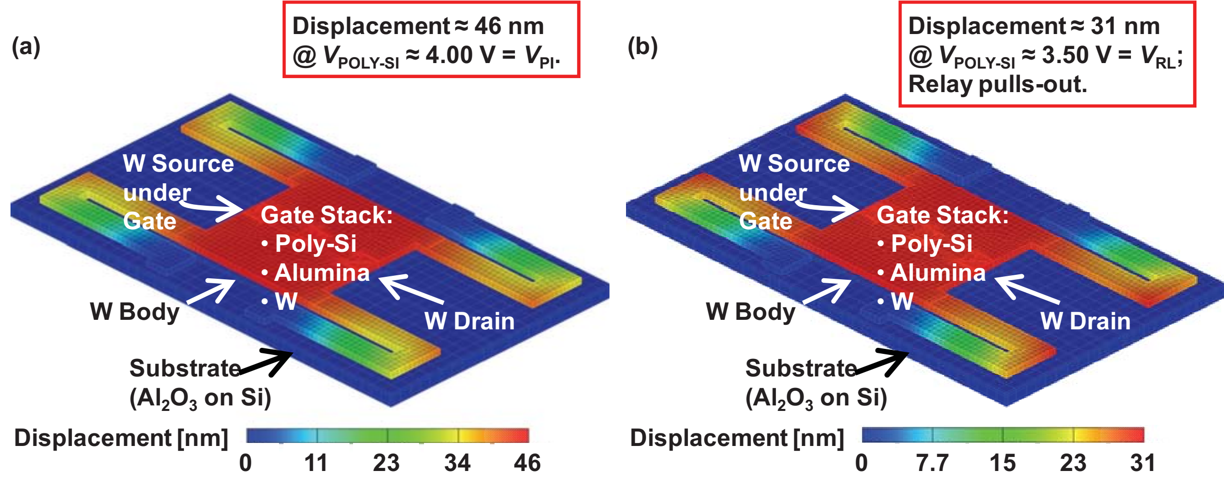


Figure 2.12: T_{ACT} = 170 nm and $T_{CONTACT}$ = 90 nm; the other vertical dimensions are as described in the Figure 2.9. The Si substrate was left floating. (a) V_{PI} of the relay is 4 V. The downward displacement of ~46 nm is close to T_{ACT} / 3, where pull-in occurs. $V_D = V_S = V_{B_W} = 0$ V. (b) V_D, V_S and $V_{CHANNEL}$ were set equal to $1.5 \cdot V_{PI} = 6$ V. $V_{B_W} = 0$ V. $V_{POLY-SI}$ was decreased from 6 V. The relay turns *off* at $V_{POLY-SI} = 3.5$ V.

The revised 4-T relays fabricated in the Berkeley Microfabrication Laboratory have been demonstrated to have good electrostatic integrity [29]; In fact, these relays were used to demonstrate the operation of an one-bit adder (12 relays) [29] as well as a full (7:3) compressor (98 relays) [30].

2.4. Summary

Improvements in 4-T relay design to maximize the percentage overlap between the actuated region of the movable electrode and the underlying actuation electrode, and to minimize the overlap between the movable electrode and the source/drain electrodes, are demonstrated to provide for a strong body bias effect with equally effective relay switching using either the movable electrode or the underlying actuation electrode as the gate, as well as dramatically reduced drain bias effect on the gate switching voltage. Although there is still room for further improvement, for example, to increase layout compactness and to prevent pull-in of the entire actuated portion of the poly-$Si_{0.4}Ge_{0.6}$ electrode, this work represents a significant step forward in the development of 4-T relays for ultra-low-power digital integrated circuit applications.

2.5. References

[1] R. H. Dennard, J. Cai, and A. Kumar, "A Perspective on Today's Scaling Challenges and Possible Future Directions," *Elsevier Solid-State Electronics*, vol. 51, no. 4, pp. 518-525, April 2007.

[2] K. Akarvardar, D. Elata, R. Parsa, G. C. Wan, K. Yoo, J. Provine, P. Peumans, R. T. Howe, and H.-S. P. Wong, "Design Considerations for Complementary Nano-Electro-Mechanical Logic Gates," in *IEEE International Electron Device Meeting (IEDM '07) Technical Digest*, pp. 299-302, December 2007.

[3] R. Nathanael, V. Pott, H. Kam, J. Jeon, and T.-J. King Liu, "4-Terminal Relay Technology for Complementary Logic," in *IEEE International Electron Device Meeting (IEDM '09) Technical Digest*, pp. 223-226, December 2009.

[4] H. Kam, V. Pott, R. Nathanael, J. Jeon, E. Alon, and T.-J. King Liu, "Design and Reliability of a Micro-Relay Technology for Zero-Standby-Power Digital Logic Applications," in *IEEE International Electron Device Meeting (IEDM '09) Technical Digest*, pp. 809-812, December 2009.

[5] F. Chen, M. Spencer, R. Nathanael, C. Wang, H. Fariborzi, A. Gupta, H. Kam, V. Pott, J. Jeon, T.-J. King Liu, D. Marković, V. Stojanović, and E. Alon "Demonstration of Integrated Micro-Electro-Mechanical (MEM) Switch Circuits for VLSI Applications," in *IEEE International Solid-State Circuits Conference (ISSCC '10) Technical Digest*, pp. 150-151, February 2010.

[6] H. Fariborzi, M. Spencer, V. Karkare, J. Jeon, R. Nathanael, C. Wang, F. Chen, H. Kam, V. Pott, E. Alon, T.-J. King Liu, V. Stojanović, and D. Marković, "Analysis and Demonstration of MEM-Relay Power Gating," *IEEE Custom Integrated Circuits Conference (CICC '10)*, San Jose, CA, September 2010.

[7] E. J. J. Kruglick, "Micro-Relay Design, Performance, and Systems," *Ph.D. Dissertation*, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Spring 1999.

[8] F. Chen, H. Kam, D. Markovic, T.-J. King Liu, V. Stojanović, and E. Alon, "Integrated Circuit Design with NEM relays," in *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD '08)*, pp. 750-757, December 2008.

[9] V. L. Rabinovich, R. K. Gupta, and S. D. Senturia, "The Effect of Release-Etch Holes on the Electromechanical Behavior of MEMS Structures," in *Proceedings of the IEEE International Conference on Solid-State Sensors and Actuators (Transducers '97)*, pp. 1125-1128, June 1997.

[10] G. M. Rebeiz, *RF MEMS Theory, Design, and Technology*, Hoboken, NJ: John-Wiley & Sons, Inc., 2003.

[11] R. L. Puurunen, J. Saarilahti, and H. Kattelus, “Implementing ALD Layers in MEMS processing,” *Electrochemical Society Transactions*, vol. 11, no. 7, pp. 3-14, October 2007.

[12] M. K. Tripp, C. Stampfer, D. C. Miller, T. Helbling, C. F. Herrmann, C. Hierold, K. Gall, S. M. George, and V. M. Bright, “The Mechanical Properties of Atomic Layer Deposited Alumina for Use in Micro- and Nano-Electro-Mechancal Systems,” *Elsevier Sensors and Actuators A: Physical*, vol. 130-131, pp. 419-429, August 2006.

[13] R. L. Puurunen, “Surface Chemistry of Atomic Layer Deposition: a Case Study for the Trimethylaluminum/Water Process,” *AIP Journal of Applied Physics*, vol. 97, no. 12, pp. 121301, June 2007.

[14] T. Bakke, J. Schmidt, M. Friedrichs, and B. Völker, “Etch Stop Materials for Release by Vapor HF Etching,” in *Proceedings of the 16th Workshop on Workshop on Micromachining, Micromechanics, and Microsystems (MME '05)*, pp. 103-106, September 2005.

[15] R. M. Costescu, D. G. Cahill, F. H. Fabreguette, Z. A. Sechrist, and S. M. George, “Ultra-Low Thermal Conductivity in W/Al_2O_3 Nanolaminates,” *Science*, vol. 303, pp. 989-990, February 2004.

[16] J. M. Jungk, N. R. Moody, T. M. Mayer, R. A. Wind, S. M. George, and W. W. Gerberich, “Properties and Fracture of Tungsten-Alumina Atomic Layer Deposited Nanolaminates,” in *Proceedings of the ICF XI: 11th International Conference on Fracture*, March 2005.

[17] M. D. Groner, J. W. Elam, F. H. Fabreguette, and S. M. George, “Electrical Characterization of Thin Al_2O_3 Films Grown by Atomic Layer Deposition on Silicon and Various Metal Substrates,” *Elsevier Thin Solid Films*, vol. 413, no. 1-2, pp. 186-197, June 2002.

[18] R. Holm, *Electric Contacts: Theory and Applications*, Berlin, NY: Springer-Verlag, 1967.

[19] *[Online] Available: http://periodictable.com/Properties/A/VickersHardness.html*.

[20] *Mathematica User's Manual [Online] Available: http://reference.wolfram.com/mathematica/ref/ElementData.html*, Wolfram Research, Inc., Champaign, IL, 2010.

[21] K. R. Williams, K. Gupta, and M. Wasilik, “Etch Rates for Micromachining Process - Part II,” *IEEE Journal of Microelectromechanical Systems*, vol. 12, no. 6, pp. 761-778, December 2003.

[22] C. W.-Z. Low, "Novel Processes for Modular Integration of Silicon-Germanium MEMS with CMOS Electronics," *Ph.D. Dissertation*, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, February 2007.

[23] S. Sedky, R. T. Howe, and T.-J. King Liu, "Pulsed-Laser Annealing, a Low-Thermal-Budget Technique for Eliminating Stress Gradient in poly-SiGe MEMS Structures," *IEEE Journal of Microelectromechanical Systems*, vol. 13, no. 4, pp. 669-675, August 2004.

[24] J. El-Rifai, A. Witvrouw, A. A. Aziz1, R. Puers, C. V. Hoof, and S. Sedky, "Selective Laser Annealing for Improved SiGe MEMS Structural Layers at 210 ºC," in *Proceedings of the IEEE International Conference of Microelectromechanical Systems (MEMS '10)*, pp. 324-327, January 2010.

[25] H. Kam, T.-J. King Liu, V. Stojanović, D. Marković, and E. Alon, "Design, Optimization, and Scaling of MEM Relays for Ultra-Low-Power Digital Logic," *IEEE Transactions on Electron Devices*, vol. 58, no. 1, pp. 236-250, January 2011.

[26] W. C. Tang, C. T.-C. Nguyen, and R. T. Howe, "Laterally Driven Poly-Silicon Resonant Microstructures," in *Proceedings of the IEEE International Conference of Microelectromechanical Systems (MEMS '89)*, pp. 53-59, February 1989.

[27] *CoventorWare™2008 Designer Reference Manual*, Coventor, Inc., Cary, NC, 2008.

[28] *CoventorWare™2008 Analyzer Reference Manual*, Coventor, Inc., Cary, NC, 2008.

[29] M. Spencer, F. Chen, C. Wang, F. Chen, R. Nathanael, H. Fariborzi, H. Kam, V. Pott, J. Jeon, T.-J. King Liu, D. Marković, E. Alon, and V. Stojanović, "Demonstration of Integrated Micro-Electro-Mechanical Switch Circuits for VLSI Applications," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 308-320, January 2011.

[30] H. Fariborzi, F. Chen, R. Nathanael, J. Jeon, T.-J. King Liu, and V. Stojanović, "Design and Demonstration of Micro-Electro-Mechanical Relay Multipliers," submitted to *37th European Solid-State Circuits Conference (ESSCIRC '11)*.

Chapter 3

Four-Terminal Relay Design and Operation for Embedded Memory Applications

Abstract: An electrostatically actuated four-terminal (4-T) micro-relay is demonstrated to function as a memory cell. This indicates promise for embedding non-volatile memory with relay-based logic circuits, without incurring additional fabrication cost.

3.1. Introduction

Complementary-Metal-Oxide-Semiconductor (CMOS) technology scaling is challenged by power density constraints, due to a slowdown in voltage scaling and limitations in cooling technology [1]. The root cause of this CMOS power crisis is transistor *off*-state leakage current (I_{OFF}), which increases exponentially with decreasing threshold voltage. This fundamental issue can be circumvented altogether by using mechanical switches which have zero I_{OFF} and abrupt *on/off* switching behavior [2], [3]. For efficient implementation of digital integrated circuits with nano-electro-mechanical (NEM) switches, a four-terminal (4-T) relay design is necessary [4], [5]. Recent developments to improve device yield and endurance (greater than 10^9 *on/off* switching cycles) [6] enabled the demonstration of functional monolithically integrated digital building blocks (logic, dynamic memory, and clocking structures) implemented with 4-T relays [7], [8], [9]. For many applications, *e.g.* system on chip, the capability to implement non-volatile memory (NVM) in a logic process (*i.e.* without significant additional fabrication cost) is desirable. Therefore, the capability of a 4-T relay to function as a NVM cell is investigated in this chapter.

3.2. Memory Relay Design

Figure 3.1(a) is a scanning electron micrograph of a programmable 4-T relay that can be fabricated using the same process as described in Chapter 2.2.1 for 4-T logic relays. If the voltage applied between the Tungsten (W) gate electrode and movable poly-$Si_{0.4}Ge_{0.6}$ body (V_{GB}) is greater than the "pull-in" voltage (V_{PI}), then the channel (attached underneath the poly-$Si_{0.4}Ge_{0.6}$ body in the "hammerhead" region via an intermediary Al_2O_3 dielectric layer) is brought into contact with the W source and drain electrodes so that direct current can flow between these electrodes, *i.e.*, the relay is turned *on*. If V_{GB} is less than the "release" voltage, V_{RL}, then an air gap separates the channel from the source/drain electrodes so that no direct current

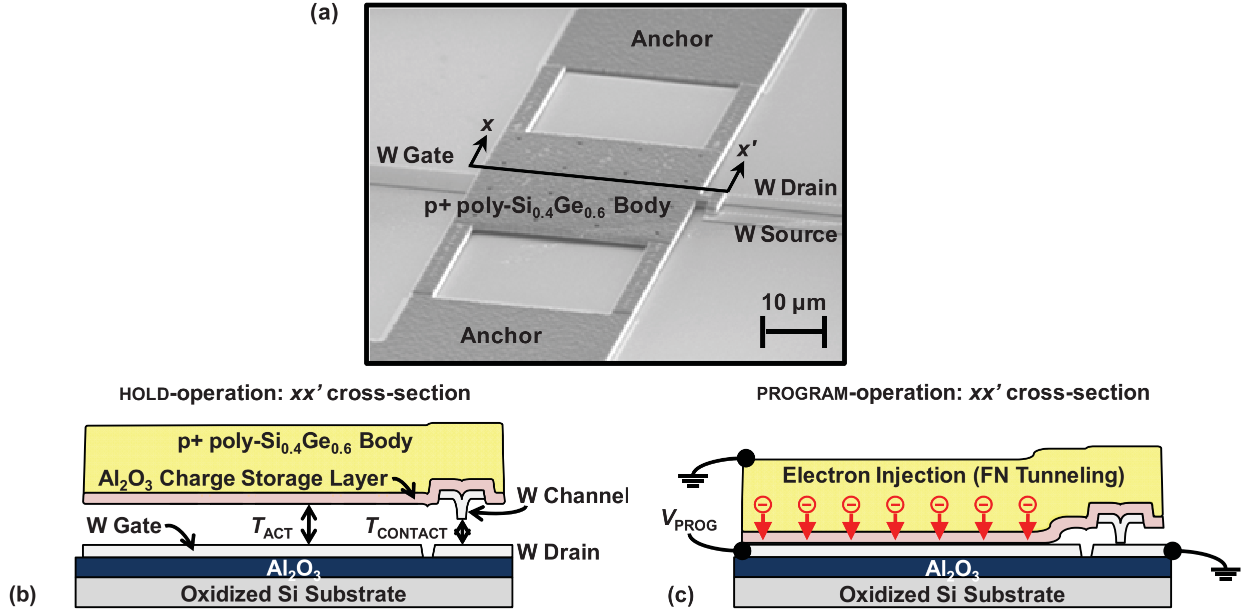


Figure 3.1: (a) Scanning electron micrograph of a fabricated 4-T relay. A Tungsten (W) channel (2 μm wide by 7.5 μm long) is attached underneath the movable body in the hammerhead region only. (b) Schematic cross-section of the relay in the **HOLD** state. T_{ACT} = 200 nm, $T_{CONTACT}$ = 100 nm, and the area of each contact dimple is 1 μm^2. The W electrodes are 50 nm thick and the poly-$Si_{0.4}Ge_{0.6}$ body electrode is 1 μm thick. (c) Illustration of **PROGRAM**-operation: If the applied gate-to-body voltage is sufficiently large, then the body is pulled in and electrons are injected into the Al_2O_3 layer, and hence, V_{PI} is reduced.

can flow, *i.e.*, the relay is turned *off*. As illustrated in Figure 3.1(b), the as-fabricated thickness of the gap in the W source/drain contact regions, $T_{CONTACT}$, is smaller than the as-fabricated thickness of the gap in the actuation region, T_{ACT}. This "dimpled" contact design is beneficial to precisely define the area of the contacting regions and to reduce the turn-*on* delay.

The relay switching voltages, V_{PI} and V_{RL}, can be affected by electronic charge within the dielectric layer [4]. Therefore, binary information can be stored in the form of trapped charge, similarly as in Silicon-Oxide-Nitride-Oxide-Silicon-type (SONOS-type) memory devices [10], [11]. In the **HOLD** state ($V_G = V_B = 0$ V), the poly-$Si_{0.4}Ge_{0.6}$ body electrode is suspended above the W gate, source, and drain electrodes, and stored charge is retained in the dielectric layer. To program the cell, a large positive gate voltage (V_{PROG}) is applied such that V_{GB} is much larger than V_{PIO} (where V_{PIO} is the native pull-in voltage), so that the poly-$Si_{0.4}Ge_{0.6}$ body electrode is pulled in to the W gate electrode, and electron injection into the dielectric layer via Fowler-Nordheim (FN) tunneling can occur, as shown in Figure 3.1(c), to reduce the relay switching voltages. To read out the binary information stored in the cell, a moderate gate voltage (V_{READ}) is applied such that V_{GB} is less than V_{PIO}, and then the cell current is sensed: if the cell is programmed, then the relay will be turned *on* so that current flows; if the cell is erased, then the relay will be turned *off* so that no current flows.

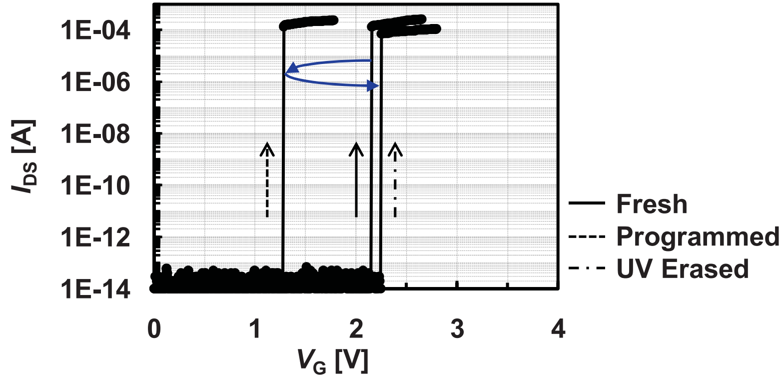


Figure 3.2: Measured I_{DS} *vs.* V_G characteristics of a fabricated 4-T relay (at room temperature, under N_2 purge) before programming, after programming (15 minutes at V_{GB} = 10 V), and after a subsequent 15-minute UV exposure. V_D = 1.5 V, V_S = 0 V, and V_B = -2.5 V. Note that the zero-leakage and abrupt-switching behavior of the relay provides for a large (essentially infinite) current-sensing margin.

3.3. Testing Results

Figure 3.2 compares the measured I_{DS} *vs.* V_G characteristics for a 4-T relay before and after programming and after ultraviolet (UV) exposure. A reduction in switching voltage is clearly seen for the programmed device. Note that the zero I_{OFF} provides for a very large current-sensing window. Together with the abrupt switching behavior (less than 1 mV/decade), this makes it possible to operate a relay-based memory cell with a relatively small program/erase voltage window, or as a multi-level cell. Of course, variation in the programmed-state and erased-state switching voltages must be minimized, for example, by established circuit techniques [12] to facilitate this.

UV exposure is effective for removing most of the stored charge; any remaining charge is likely trapped at the dielectric/body interface. The reduction in switching voltages increases with the duration of the programming voltage pulse (hence FN tunneling), as shown in Figure 3.3. The device also can be erased via FN tunneling by applying a large gate voltage of opposite polarity, as shown in Figure 3.4. For very long erase times, holes are injected into the dielectric layer, so that the switching voltages are increased beyond their native values.

The program/erase times used in this work are much longer than typical (10 ms) for flash memory devices [13]. This is because a single layer of dielectric material (40-nm-thick Al_2O_3) serves as both the tunneling material and charge storage material. The use of a multi-layer dielectric stack comprising at least a thin (less than 3-nm-thick [14], [15]) tunnel dielectric layer and a charge-storage layer with smaller energy band gap (to allow for direct tunneling of electrons into the conduction band) should provide for much faster program/erase operation, as well as improved retention and endurance characteristics. Note that endurance characteristics are not reported herein because the dielectric stack is sub-optimal.

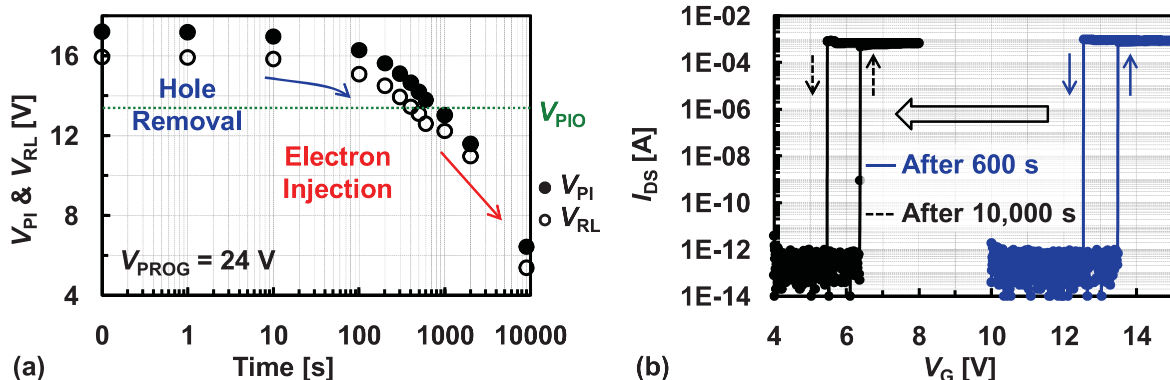


Figure 3.3: (a) Measured program behavior of a 4-T relay. V_D, V_S and V_B were biased at 0 V during the program operation. The FN tunneling current is on the order of pA. (b) Measured I_{DS}-V_G characteristics after 600 and 10,000 seconds programming voltage pulses were applied.

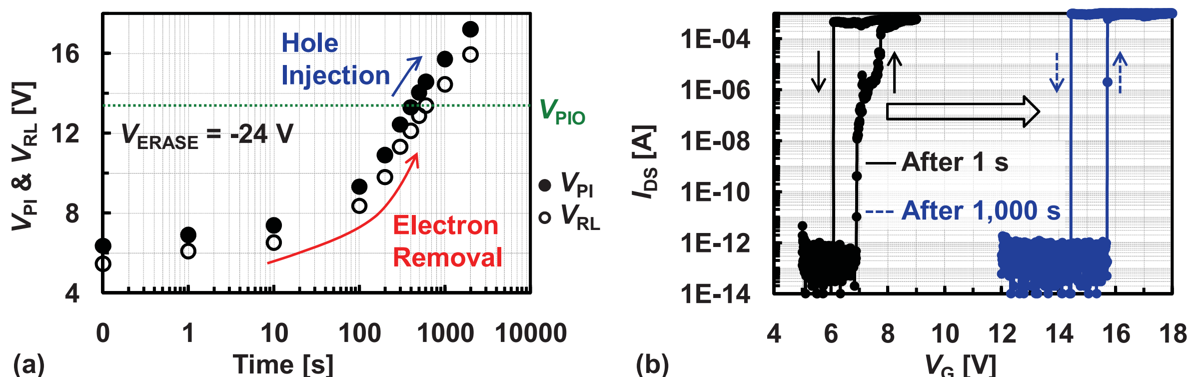


Figure 3.4: (a) Measured erase behavior of a 4-T relay. V_D, V_S and V_B were biased at 0 V during the erase operation. (b) Measured I_{DS}-V_G characteristics after 1 and 1,000 seconds erasing voltage pulses were applied.

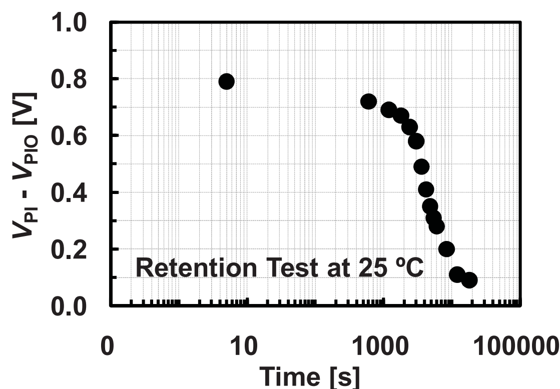


Figure 3.5: Measured hole retention behavior of an over-erased 4-T relay cell.

Figure 3.5 shows the measured hole retention behavior of an over-erased cell. A retention time of approximately 1 hour (corresponding to a minimum voltage window of 0.5 V) is seen. The use of a multi-layer dielectric stack, as shown in Figure 3.6, should provide for much longer retention time, as there would be a significant potential barrier to prevent charge loss into the body.

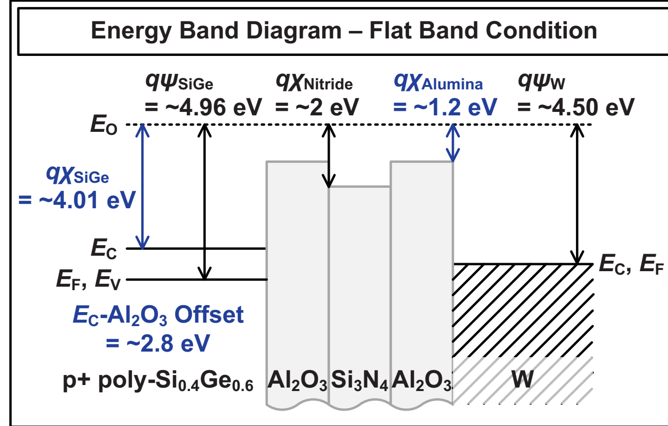


Figure 3.6: Energy band diagram for a SONOS-type memory 4-T relay under flat band condition. Si_3N_4 is used as a charge trap layer (less than 10 nm), and Al_2O_3 is used as tunneling (less than 5 nm, *c.f.* mean free path of electrons in the oxide = 3.2 nm [15]) and blocking (less than 15 nm) layers. Si_3N_4 is a good trap material, since it provides deep-level traps in which a net positive or negative charge can be stored [16]. A programming voltage applied across the dielectric stack leads to a thinner electrical tunneling oxide thickness, *i.e.* higher tunneling probability.

Storage density is a consideration for any memory device. As mentioned previously, a relay-based memory cell potentially can store more than one bit of information. To further increase storage density, the device layout area must be reduced. If the vertical dimensions, such as T_{ACT}, $T_{CONTACT}$, and the body electrode thickness, are scaled down in proportion with the lateral dimensions, the cell operating voltages will also scale down proportionately [3]. In order to avoid stiction issues in the contacting regions, a dry release-etch process (*e.g.* using a mixture of anhydrous HF and vapor-phase methanol at sub-atmospheric pressures) can be used [17], [18].

3.4. Summary

In this chapter, a 4-T relay is demonstrated to function as a memory cell. To achieve fast program/erase operation and long retention time, a multi-layer dielectric is necessary. Such a dielectric stack would not pose an issue for a logic relay, because it only serves to electrically isolate the channel. In other words, the same multi-layer dielectric stack can be used for logic *vs.* memory relays, and would not significantly affect the operation, endurance, or scalability of the logic relay. Note that for a given actuation gap (T_{ACT}), the relay switching voltages will increase with the equivalent air-gap thickness of the gate dielectric stack – which would be very small since the permittivities of the gate dielectric materials would be much larger than that of air. In contrast, the dielectric layer(s) in a Metal-Oxide-Semiconductor-Field-Oxide-Transistor (MOSFET) must also provide adequate capacitive coupling between the gate electrode and the channel region; therefore, the dielectric stacks for logic vs. memory devices in a CMOS technology typically are different. Thus, a 4-T relay logic technology can offer embedded memory capability without significant additional process cost.

3.5. References

[1] R. H. Dennard, J. Cai, and A. Kumar, "A Perspective on Today's Scaling Challenges and Possible Future Directions," *Elsevier Solid-State Electronics*, vol. 51, no. 4, pp. 518-525, April 2007.

[2] K. Akarvardar, D. Elata, R. Parsa, G. C. Wan, K. Yoo, J. Provine, P. Peumans, R. T. Howe, and H.-S. P. Wong, "Design Considerations for Complementary Nano-Electro-Mechanical Logic Gates," in *IEEE International Electron Device Meeting (IEDM' 07) Technical Digest*, pp. 299-302, December 2007.

[3] F. Chen, H. Kam, D. Marković, T.-J. King Liu, V. Stojanović, and E. Alon, "Integrated Circuit Design with NEM Relays," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD '08) Technical Digest*, pp. 750-757, December 2008.

[4] R. Nathanael, V. Pott, H. Kam, J. Jeon, and T.-J. King Liu, "4-Terminal Relay Technology for Complementary Logic," in *IEEE International Electron Device Meeting (IEDM '09) Technical Digest*, pp. 223-226, December 2009.

[5] E. J. J. Kruglick, "Micro-Relay Design, Performance, and Systems," *Ph.D. Dissertation*, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Spring 1999.

[6] H. Kam, V. Pott, R. Nathanael, J. Jeon, E. Alon, and T.-J. King Liu, "Design and Reliability of a Micro-Relay Technology for Zero-Standby-Power Digital Logic Applications," in *IEEE International Electron Device Meeting (IEDM '09) Technical Digest*, pp. 809-812, December 2009.

[7] F. Chen, M. Spencer, R. Nathanael, C. Wang, H. Fariborzi, A. Gupta, H. Kam, V. Pott, J. Jeon, T.-J. King Liu, D. Marković, V. Stojanović, and E. Alon "Demonstration of Integrated Micro-Electro-Mechanical (MEM) Switch Circuits for VLSI Applications," in *IEEE International Solid-State Circuits (ISSCC '10) Conference Technical Digest*, pp. 150-151, February 2010.

[8] M. Spencer, F. Chen, C. Wang, F. Chen, R. Nathanael, H. Fariborzi, H. Kam, V. Pott, J. Jeon, T.-J. King Liu, D. Marković, E. Alon, and V. Stojanović, "Demonstration of Integrated Micro-Electro-Mechanical Switch Circuits for VLSI Applications," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 308-320, January 2011.

[9] H. Fariborzi, M. Spencer, V. Karkare, J. Jeon, R. Nathanael, C. Wang, F. Chen, H. Kam, V. Pott, E. Alon, T.-J. King Liu, V. Stojanović, and D. Marković, "Analysis and Demonstration of MEM-Relay Power Gating," *IEEE Custom Integrated Circuits Conference (CICC '10)*, San Jose, CA, September 2010.

[10] B. V. Keshevan and H. C. Lin, "MONOS Memory Element," in *IEEE International Electron Device Meeting (IEDM '68) Technical Digest*, pp. 140, October 1968.

[11] P. C. Y. Chen, "Threshold-Alterable Si-Gate MOS Devices," *IEEE Transactions on Electron Devices*, vol. 24, no. 5, pp. 584-586, May 1977.

[12] K. Imamiya, H. Nakamura, T. Himeno, T. Yamamura, T. Ikehashi, K. Takeuchi, K. Kanda, K. Hosono, T. Futatsuyama, K. Kawai, R. Shirota, N. Arai, F. Arai, K. Hatakeyama, H. Hazama, M. Saito, H. Meguro, K. Conley, K. Quader, and J. J. Chen, "A 125-mm^2 1-Gb NAND Flash Memory with 10-MByte/s Program Speed," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 11, pp. 1493-1501, November 2002.

[13] S. Minami and Y. Kamigaki, "A Novel MONOS Nonvolatile Memory Device Ensuring 10-Year Data Retention after 10^7 Erase/Write Cycles," *IEEE Transactions on Electron Devices*, vol. 40, no. 11, pp. 2011-2017, November 1993.

[14] M. H. White, Y. Yang, A. Purwar, and M. L. French, "A Low Voltage SONOS Nonvolatile Semiconductor Memory Technology," *IEEE Transactions on Components, Packaging, and Manufacturing Technology - Part A*, vol. 20, no. 2, pp. 190-195, June 1997.

[15] M. Erdamar, K. Shanmugasundaram, P. Roman, P. Mumbauer, M. Klimkiewicz, and J. Ruzyllo, "Deep Lateral Anhydrous HF/Methanol Etching for MEMS Release Processes," *SPIE Journal of Micro/Nano Lithography, MEMS, and MOEMS*, vol. 7, no. 3, pp. 033014, July 2008.

[16] M. H. White, D. A. Adams, and J. Bu, "On the Go with SONOS," *IEEE Circuits and Devices Magazine*, vol. 16, no. 4, pp. 22-31, July 2000.

[17] J. Bu and M. H. White, "Design Considerations in Scaled SONOS Nonvolatile Memory Devices," *Elsevier Solid-State Electronics*, vol. 45, no. 1, pp. 113-120, January 2001.

[18] W. I. Jang, C. A. Choi, M. L. Lee, C. H. Jun, and Y. T. Kim, "Fabrication of MEMS Devices Using Anhydrous HF Gas-Phase Etching with Alcoholic Vapor," *IOPscience Journal of Micromechanics and Microengineering*, vol. 12, no. 3, pp. 297-306, April 2002.

Chapter 4

See-Saw Relay Technology for Complementary Logic and Memory Circuits

Abstract: A dual-ended "see-saw" relay design is proposed and demonstrated. Fabricated see-saw relays demonstrate perfectly complementary switching behavior that is symmetric about V_{DD} / 2, with extremely steep switching behavior (less than 1 mV/decade), very small switching asymmetry (approximately 300 mV), and low *on*-state resistance (less than 1 kΩ). The perfectly complementary and symmetric operation provides for zero crowbar current, as evidenced by an abrupt inverter voltage transfer characteristic. Various logic functions can be implemented by appropriately biasing a single see-saw relay. The see-saw relay also can be configured as a bi-stable latch, so that a memory cell can be implemented with one relay and one access transistor or relay. Measurements of see-saw-relay switching speed are well matched to lumped-parameter modeling results.

4.1. Introduction

Nano-electro-mechanical relays have been proposed for ultra-low-power digital integrated circuit applications [1], [2], because their ideal switching behavior (zero *off*-state leakage and abrupt *on/off* switching behavior) in principle allows the supply voltage (V_{DD}) to be scaled down further than for Complementary-Metal-Oxide-Semiconductor (CMOS) transistors [3]. Micrometer-scale single-ended relays have recently been demonstrated to operate reliably with more than 10^9 *on/off* cycles [4] and with relatively low gate-voltage swing (less than 2 V), when a body electrode is used for threshold-voltage (V_{TH}) adjustment [5]. Operation of relays as "pull-up" and "pull-down" devices is desirable for implementing complementary logic and was demonstrated in [5], [6]. Various functional digital-IC building blocks, such as carry-generation circuit, oscillator, and **XOR**, have been demonstrated using these single-ended relays and passive circuit components [6], [7], [8], [9]. The V_{TH} values should be set such that switching is symmetric about V_{DD} / 2, in order to maximize operating voltage margin and to minimize crowbar current. This is challenging to achieve, however, with single-ended relays, due to process-induced V_{TH} variations and their passive pull-out mechanism.

A single structure is also more desirable for more compact mechanical implementation of simple digital logic functions. For example, a single micro-electro-mechanical structure operated in

bending mode has been demonstrated to function as either an **OR** or **AND** gate [11]. A compound structure operated in torsional (see-saw) mode has been proposed to function as either a **NAND** or **NOR** gate [12]; however, its functionality was not verified electrically, due to very high contact resistance.

A single structure operated in see-saw mode has also been investigated for radio-frequency (RF) [13], [14] and projection display applications [15], [16]. For example, a large array of torsional structures (more than 500,000 elements) that can rotate through a large angle (± 10 degrees) has been demonstrated to operate reliably with more than 10^{12} cycles without mechanical failures, such as fatigue, stress rupture, and creep [17], [18]. A torsional RF micro-electro-mechanical-systems (MEMS) switch has been shown to operate with very low contact resistance less than 1 Ω (and hence low insertion loss) [19], [20].

In this chapter, a different see-saw relay design is proposed and demonstrated to achieve perfectly complementary switching behavior that is symmetric about V_{DD} / 2 to provide for minimal switching hysteresis, maximum noise margin, and zero "crow-bar current". This "see-saw" design inherently provides for more reliable operation, because electrostatic force as well as spring restoring force is used for *off*-switching as well as for *on*-switching, thereby overcoming surface adhesion forces in the contacting regions [19], [20]. Due to this active pull-out mechanism (rather than passive pull-out that relies solely on the spring restoring force), the see-saw relay design inherently shows minimal hysteric behavior during switching operations. Since the see-saw relay achieves the functionality of two single-ended relays with a single movable structure, it is also advantageous for compact implementation of relay-based digital logic integrated circuits; functional seesaw-relay-based logic and memory circuits are also described in this chapter.

4.2. Prototype See-Saw Relay

4.2.1. Device Design and Fabrication

Figure 4.1(a) shows the see-saw relay design that utilizes a single suspended gate electrode (the see-saw beam) anchored by two torsional beams which allow the ends of the gate to move up and down in a perfectly complementary fashion. The two Tungsten (W) channel electrodes attached via an intermediary gate-dielectric (40-nm-thick ALD Al_2O_3) layer to the underside of the p+ polycrystalline-$Si_{0.4}Ge_{0.6}$ gate at either end are used to connect a pair of W source and drain electrodes upon physical contact. Figure 4.1(b) illustrates how the two W body electrodes (co-planar with the source/drain electrode pairs) located underneath the gate electrode, one on either side of the torsional axis, are used to apply voltages across the actuation air gap (with as-fabricated thickness T_{ACT}) to electrostatically actuate the see-saw beam downward on either side to close the contact air gaps (with as-fabricated thickness $T_{CONTACT}$), turning *on* either only the

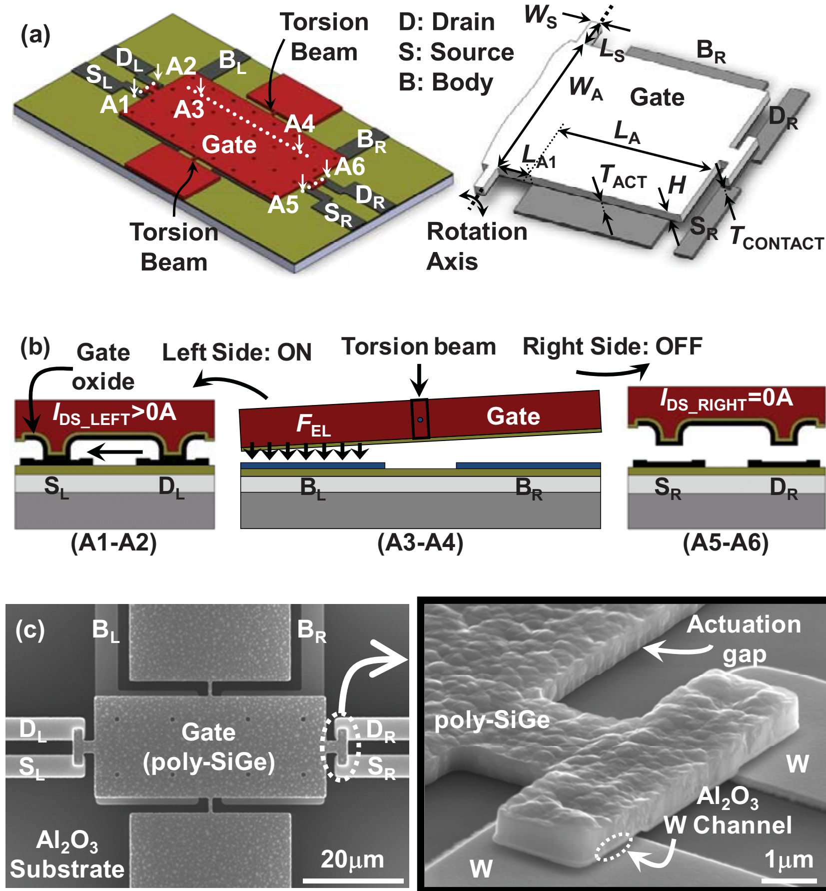


Figure 4.1: (a) Three-dimensional schematic of the see-saw relay structure and definition of design parameters. Parameter values used in this work are L = 3 µm, W = H = 1 µm, L_C = 7.5 µm, W_C = 2 µm, L_A = 11 µm, 14 µm, 30 µm, or 33 µm, L_{A1} = 9 µm or 12 µm, W_A = 20 µm or 40 µm, T_{ACT} = 0.2 µm, and $T_{CONTACT}$ = 0.1 µm. (b) Schematic cross-sectional views from (a). The mechanical gate, anchored by two torsion beams, is torqued by electrostatic force induced by applying a voltage between the gate and a body electrode. The two steady states are *on-off* (illustrated here) and *off-on*, for perfectly complementary operation. (c) Scanning electron micrograph (SEM) of a fabricated see-saw relay and zoomed-in view of the right channel region.

left side or only the right side. Details of the seesaw-relay fabrication process are described in Chapter 2. Scanning electron micrographs of fabricated see-saw relays are shown in Figures 4.1(c).

4.2.2. Testing Results from Device Perspective

I_{DS}-V_G and I_{DS}-V_{DS} characteristics of see-saw relays measured in N_2 atmosphere are plotted in Figures 4.2(a) and 4.2(b), respectively, for right body bias V_{BR} = V_{DD} = 10 V and left body bias V_{BL} = 0 V. As the gate voltage (V_G) is increased from 0 V to V_{DD} (10 V), the left side turns *on*

abruptly, and the right side turns *off* abruptly at V_{ON_L} = V_{OFF_R}, when the see-saw beam is actuated to bring the left channel into contact with the left source/drain electrodes. The *on*-state resistances are approximately 926 Ω and 800 Ω, when V_G = 8 V and 10 V, respectively. As V_G is reduced back to 0 V, the left side turns *off* abruptly and the right side turns *on* abruptly at V_{OFF_L} = V_{ON_R}, when the see-saw beam is actuated to bring the right channel into contact with the right source/drain electrodes. Note that V_{ON_L} and V_{ON_R} are symmetric about V_{DD} / 2, and that the switching asymmetry, which can be defined by the difference in pull-in voltage that makes the right side turn *on* (7.16 V) and that makes the left side turn *on* (6.84 V = 10 V – 3.16 V), is very small (300 mV). This small difference could be caused by measurement variations. On the other hand, it might be caused by strain gradient of the p+ poly-$Si_{0.4}Ge_{0.6}$ layer, which was not completely removed via excimer laser annealing; As a result, T_{ACT} on either side of the relay would be slightly different.

The complementary operation of the see-saw relay can be used to achieve a very steep inverter voltage transfer characteristic, as shown in Figure 4.3. The observed asymmetric shifts in V_{ON_LEFT} and V_{ON_RIGHT} values are caused by strain gradient of the poly-$Si_{0.4}Ge_{0.6}$ gate electrode, which makes T_{ACT} on the left and right sides inconsistent from the as-fabricated T_{ACT}. On the other hand, the asymmetric changes might be caused by gate-oxide charge trapping after 400 *on/off* switching cycles. This reliability issue, along with the mismatch between input and output voltage swings, can be mitigated by reducing the magnitude of the gate actuation voltage, for example, by reducing V_{DD} as shown in Figure 4.4 or by scaling down T_{ACT} and/or $T_{CONTACT}$. In this work, the output voltage range was kept conservatively low in order to minimize the possibility of micro-welding-induced failure. If the W electrodes are coated with a ultra-thin layer of TiO_2, the output voltage range can be increased to approximately 4 V without reliability issues because the work of adhesion in the contacting regions can be reduced [4] and current hence Joule heating is reduced. For the formation of small gaps (below 100 nm), stiction issues

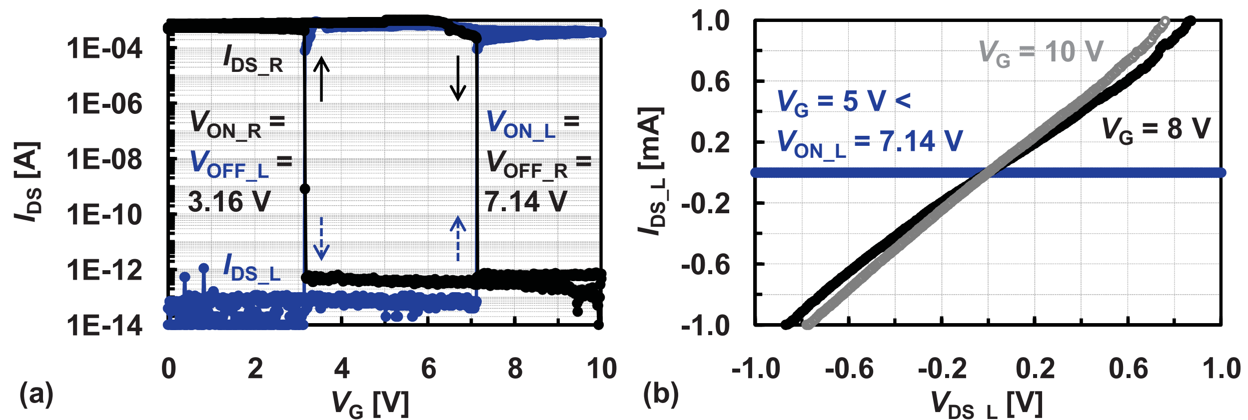


Figure 4.2: (a) Measured I_{DS}-V_G characteristics with V_{BR} = V_{DD} = 10 V, V_{BL} = 0 V, and V_{DS_R} = V_{DS_L} = 0.8 V. Note that turn-*on* and turn-*off* are abrupt (< 0.1 mV/decade) and that V_{ON_R} and V_{ON_L} are symmetric about 5 V. L_A = 30 µm, L_{A1} = 12 µm, and W_A = 40 µm. (b) Measured I_{DS_L}-V_{DS_L} characteristics for the relay in (a). The *on*-state resistances are approximately 926 Ω and 800 Ω when V_G = 8 V and 10 V, respectively.

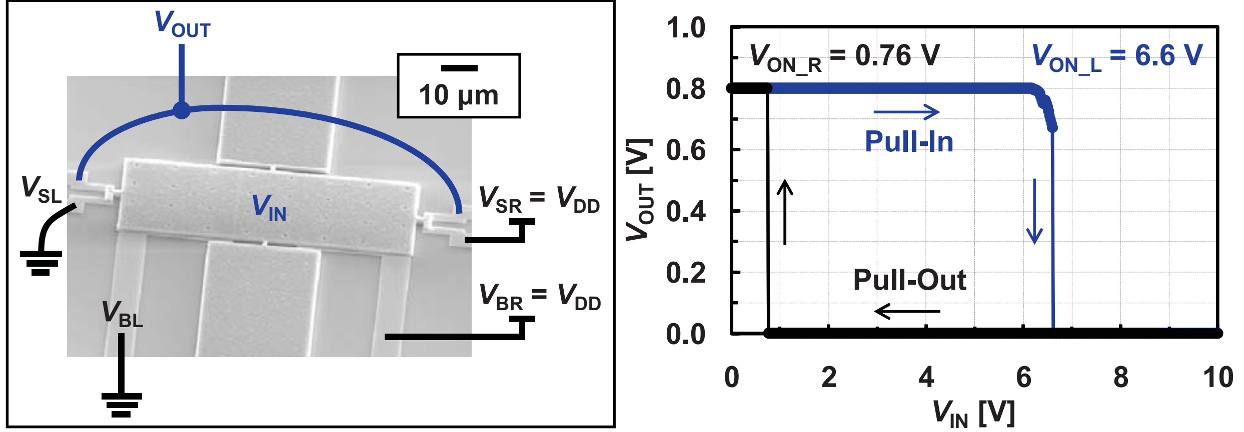


Figure 4.3: Measured voltage transfer curve of the see-saw relay in Figure 4.2, after approximately 400 switching cycles. The V_{ON_L} and V_{ON_R} values are shifted due to the strain gradient of the poly-$Si_{0.4}Ge_{0.6}$ gate electrode and/or symmetric charging of the Al_2O_3 gate oxide.

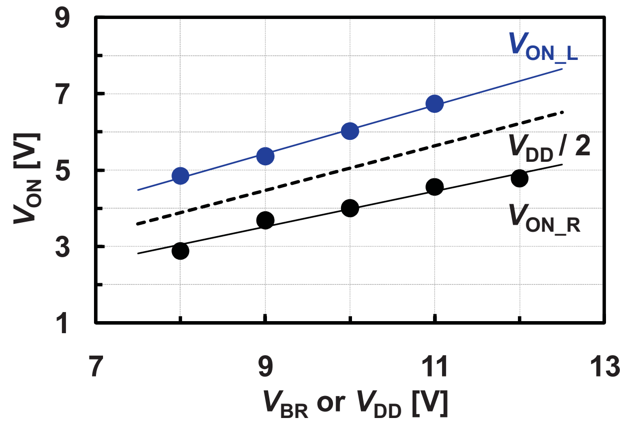


Figure 4.4: Measured values of V_{ON_L} and V_{ON_R} *vs.* V_{BR} for a different see-saw relay. Note the symmetry of switching about $V_{BR} / 2 = V_{DD} / 2$, and the reduction in switching voltages with decreasing V_{DD}.

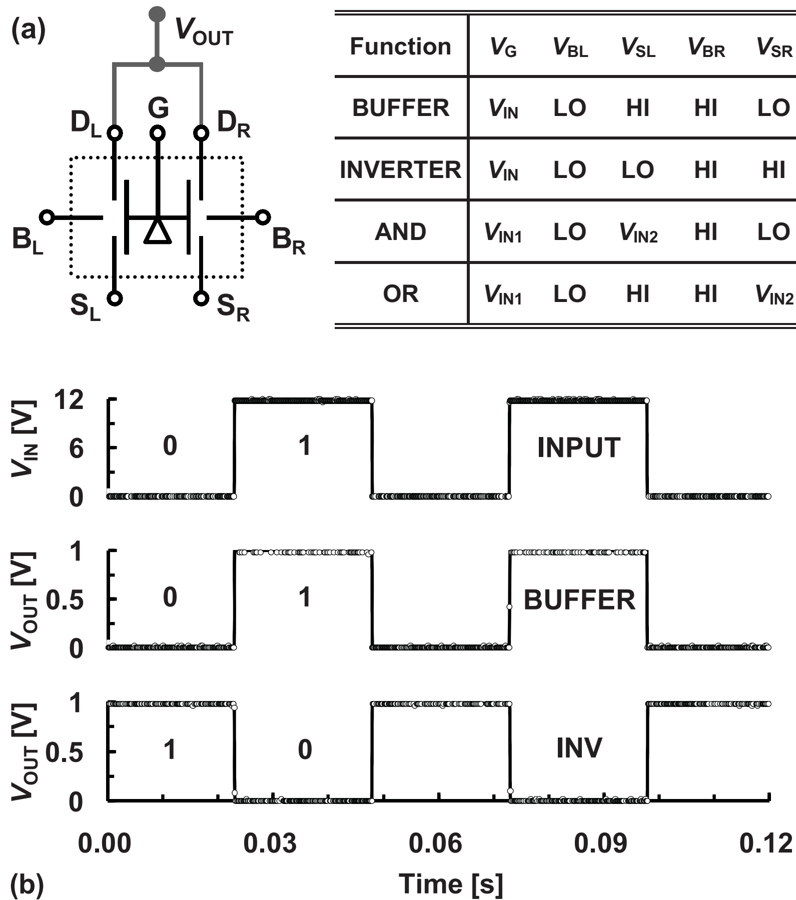
can be avoided by using an advanced dry release etch process, for example, with a mixture of vapor-phase methanol and anhydrous HF at reduced (sub-atmospheric) pressures [21], [22].

In order to reduce the device layout area, the vertical dimensions of the see-saw relay (*i.e.* T_{ACT}, $T_{CONTACT}$, and the thickness of the gate layer) should be scaled down to allow for smaller actuation area without increased gate actuation voltage. It should be noted that the optimal circuit topology for a relay-based integrated circuit is significantly different than that for a CMOS integrated circuit, so that fewer relays can often be used to implement a digital logic function [2], [7], [9]; thus the layout area efficiency of a relay-based circuit can be comparable to or better than that of a functionally equivalent CMOS circuit.

4.2.3. Testing Results from Circuit Perspective

Digital logic functions such as **BUFFER**, **NOT** (**INVERTER**), **AND**, and **OR** are easily implemented using a single see-saw relay with the drain electrodes tied together to form the output node and with appropriate bias configurations as shown in Figure 4.5(a). The measured logic waveforms (timing diagrams) for single-input functions in Figure 4.5(b) and for dual-input functions in Figure 4.5(c) confirm the versatile functionality of the see-saw relay. Note that a lower '**HI**' voltage (1 V) is used to bias the source electrodes *vs.* the gate electrode, in order to avoid potential reliability issues caused by micro-welding.

A single see-saw relay can also be configured as shown in Figure 4.6(a) to form a bi-stable latch. With an access transistor added, a Static-Random-Access-Memory (SRAM) cell is formed. In this work, a discrete n-channel Metal-Oxide-Semiconductor-Field-Effect-Transistor (MOSFET) is used as the access transistor to connect the data-storage node to a bit-line node. Another discrete n-channel MOSFET is used to write voltage data (generated by a function generator, Agilent 33120A 1-channel 15 MHz Function Generator) to the bit-line node. Full functionality of



Function	V_G	V_{BL}	V_{SL}	V_{BR}	V_{SR}
BUFFER	V_{IN}	LO	HI	HI	LO
INVERTER	V_{IN}	LO	LO	HI	HI
AND	V_{IN1}	LO	V_{IN2}	HI	LO
OR	V_{IN1}	LO	HI	HI	V_{IN2}

Figure 4.5: (a) Drain electrodes are connected to form the logic gate output node (left). A see-saw relay can be properly configured for implementing logic functions (right). (b) Measured **BUFFER/INVERTER** timing diagrams: (top) **INPUT** signal, (middle) **BUFFER** output signal, (bottom) **INVERTER** output signal.

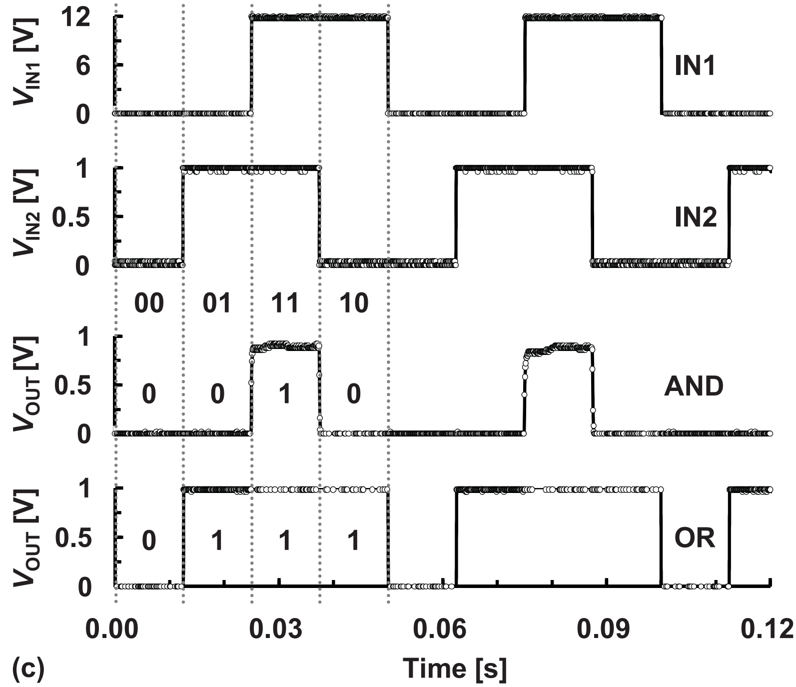
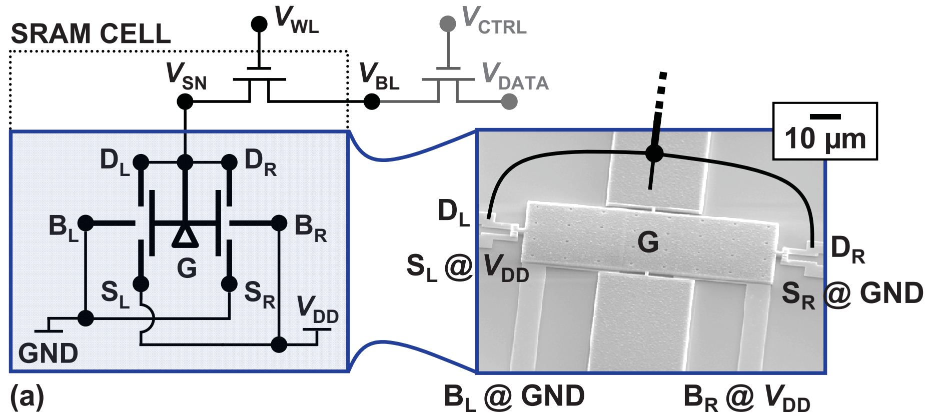


Figure 4.5: (c) Measured logic gate timing diagrams: (top) **`INPUT 1`** signal, (middle top) **`INPUT 2`** signal, (middle bottom) **`AND`** output signal, (bottom) **`OR`** output signal.

the hybrid SRAM cell is demonstrated by the measured timing diagrams shown in Figure 4.6(b). Note that a 100 kΩ resistor is inserted between the gate and each drain electrode for this demonstration, in order to allow for lower V_{DS} than V_{GB} to avoid reliability issues caused by micro-welding.

The measurement setup, illustrated in Figure 4.7(a), is used to estimate the switching delay of a see-saw relay under N_2 ambient conditions. The switching delay is estimated by measuring the time difference between a change in the input signal (from a function generator, Agilent 81150A 2-channel 120 MHz Function Generator) and a change in the output signal (from the see-saw relay). The measured switching delay is observed to decrease with increasing V_{DD}, and is approximately within 10 % of the value predicted by a lumped parameter model [23], as shown in Figure 4.7(b). In order to increase SRAM storage density and to reduce V_{DD}, the dimensions of a see-saw relay should be scaled down.



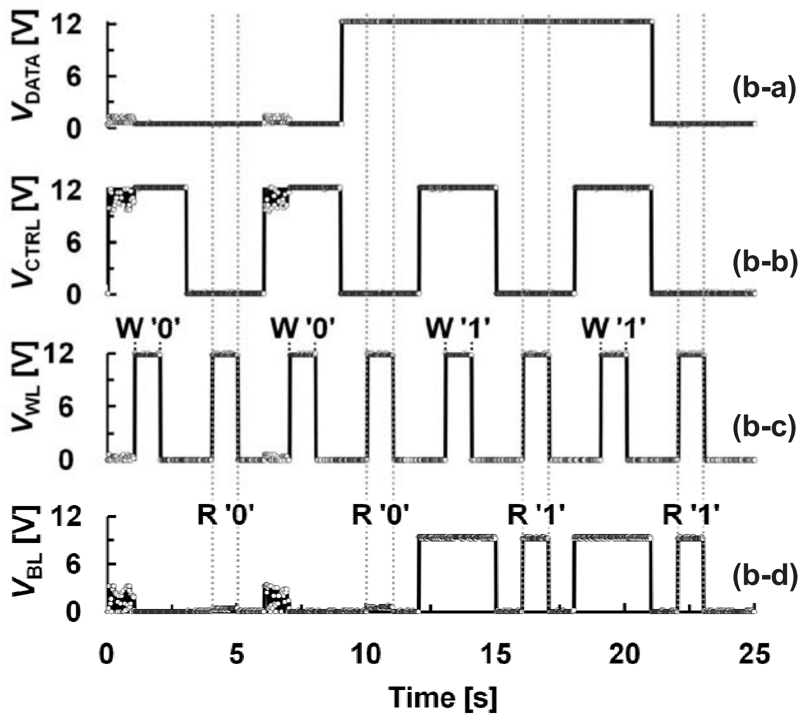


Figure 4.6: (a) Schematic circuit diagram of a relay one transistor (1R-1T) SRAM cell. (b) SRAM cell timing diagrams which demonstrate see-saw relay operation as a bi-stable latch: (b-a) voltage data signal V_{DATA} (supplied by a function generator) used for write operations, (b-b) control signal V_{CTRL} (supplied by a problematic function generator – hence the "noisy" signal for the first two pulses) used to write voltage data to the bit-line node, (b-c) word-line signal V_{WL} (supplied by a function generator) used to access the storage node, and (b-d) measured bit-line voltage (V_{BL}) waveform. The noise in the measured bit-line voltage is due to capacitive coupling to the noisy control signal V_{CTRL}.

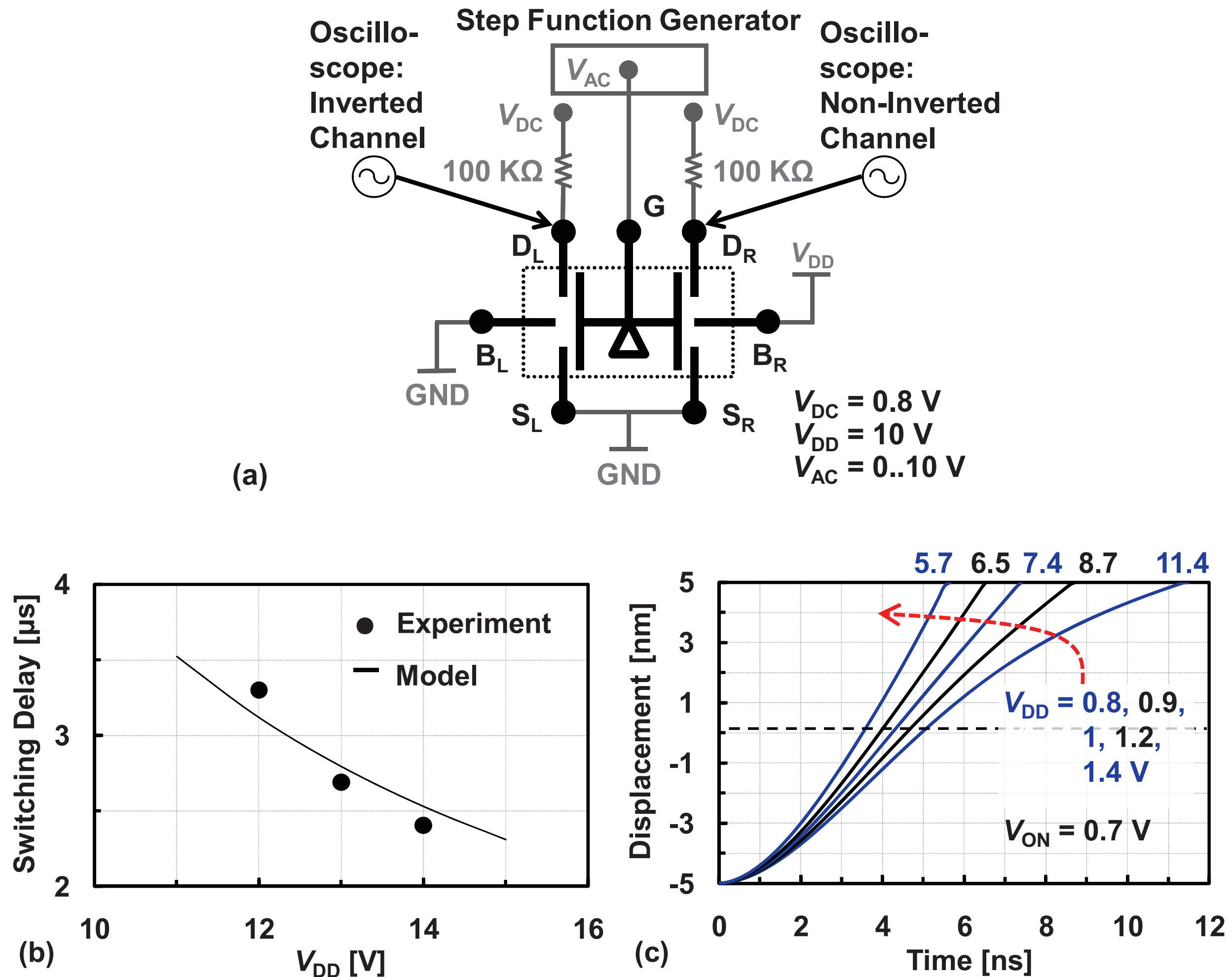


Figure 4.7: (a) Switching delay measurement setup. (b) Measured and modeled switching delay as a function of operating voltage. The measurement was performed under N_2 ambient conditions. (c) Simulated right gate-tip displacement *vs.* time, for a see-saw relay scaled to 32 nm technology. T_{ACT} = 10 nm, $T_{CONTACT}$ = 5 nm, and H = 16 nm (reference Figure 4.1).

Using a calibrated relay model [23], the switching performance of a see-saw relay for a 32-nm-equivalent technology is simulated. The simulation results indicate that sub-10-ns switching delay at V_{DD} less than 1 V is achievable for V_{DD} / V_{ON} > 1.3, as shown in Figure 4.7(c). Note that the quality factor is assumed to be 1, which is conservative assuming vacuum packaging, for example, via microshell encapsulation, for operation in the milliTorr (mT) range [24]. To achieve a given delay performance requirement, V_{DD} must be selected such that V_{DD} / V_{ON} will exceed a minimum required value, accounting for variations in V_{ON}, due to process-induced variations in device dimensions and environmental variations.

4.3. See-Saw Relay Design Optimization

The testing results of the prototype see-saw relays indicate that there is significant room for improvement and optimization. In this section, monotonicity analysis is conducted, in order to find the optimal device dimensions that yield low operating voltages and hence improve the energy performance of the device. From this monotonicity analysis, design intuitions about the limiting set of cases are derived, as are an optimal set of device dimensions. The results of the monotonicity analysis are validated by comparing against those produced by constrained numerical optimization in **MATLAB®**. In addition, the accuracy of the values of the objective function obtained from these analyses is compared against the value obtained via Finite-Element-Method (FEM) simulation of the see-saw relay device.

4.3.1. Monotonicity Analysis

Monotonicity analysis is a pre-optimization tool that can be used to solve an optimization problem with a differentiable objective function while satisfying a particular set of constraints [25], [26]. This can be achieved by enumerating various cases, in which groups of constraints are identified as redundant, active, or inactive, and by systematically checking for boundedness of all variable parameters. By doing so, algebraic solutions to constraint-bound cases can be identified, and thus, numerically-based optimization is no longer needed; because a combinatorially large solution space is now simplified to a manageable size, computational loads can be reduced significantly. The formal rules of monotonicity analysis are described as follows [26]:

1. Monotonicity is defined only for differentiable functions, $f(\boldsymbol{x})$, with respect to x_i.
2. $df(\boldsymbol{x}) / dx_i > 0$ means that $f(\boldsymbol{x})$ is strictly monotonically increasing.
3. A variable, x_i, is defined as "bounded below by a constraint," when $g_j(\boldsymbol{x}) \leq 0$, which is monotonic in x_i, if the variable achieves its minimum value at $g_j(\boldsymbol{x}) = 0$, *e.g.* $g(x) = 1 - x \leq 0$.
4. A variable, x_i, is defined as "bounded above by a constraint," when $g_j(\boldsymbol{x}) \leq 0$, which is monotonic in x_i, if the variable achieves its maximum value at $g_j(\boldsymbol{x}) = 0$, *e.g.* $g(x) = x - 1 \leq 0$.
5. An inequality constraint, $g_j(\boldsymbol{x})$, is defined as "active at $\boldsymbol{x}^*$," if $g_j(\boldsymbol{x}^*) = 0$.
6. An inequality constraint, $g_j(\boldsymbol{x})$, is defined as "inactive at $\boldsymbol{x}^*$," if $g_j(\boldsymbol{x}^*) < 0$.
7. An equality constraint, $h_k(\boldsymbol{x}) = 0$, is defined as "irrelevant at $\boldsymbol{x}^*$," if the optimal solution is not affected by eliminating the constraint.
8. If the objective function, $f(\boldsymbol{x})$, is ± monotonic with respect to x_i, then there must exist at least

one active constraint that is of opposite monotonicity in x_i, *i.e.* all values must be bounded.

9. A variable, x_i, not explicitly represented in the objective function, must either be contained only in inactive constraints or be contained in two active constraints of opposite monotonicity.

10. Degenerate or unbounded problems are violations of the rules 8 and 9.

4.3.1.1. *Assumptions*

Unless otherwise specified, the design optimization studies shown in this section are performed based on the following assumptions:

1. Non-linearities are ignored.

2. Infinitely stiff plates and trusses.

3. Perfectly rectangular torsional beams.

4. Mechanical deformation is elastic.

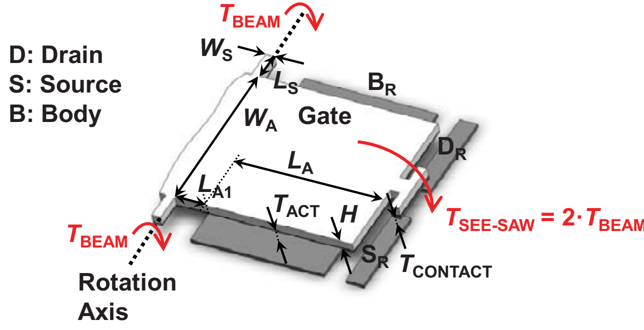
5. Purely torsional mode operation, *i.e.* minimal out-of-axis motion or forces, including gravity.

6. Ideal processing conditions, *i.e.* devices are fabricated as designed, with no manufacturing variations.

7. Ideal materials that are free of any mechanical reliability issues, such as stress/strain gradient, fatigue, and creep.

8. Isotropic materials are used, which have constant Young's modulus (E) and moment of inertia (I) throughout the structure.

9. The structure is well encapsulated such that the impact of temperature, humidity, and particle contamination on device performance can be ignored.



Parameter	Value	Parameter	Value
T_{ACT}	200 nm	E_{SiGe}	140 GPa
$T_{CONTACT}$	100 nm	G_{SiGe}	57.4 GPa
H	1 µm	R_S	500
λ	1 µm	$F_{ADHESION}$	0.2 µN [27]
β	0.141 (for $W_S = H$) [29]	L_{MAX}	100 µm
(b)		W_{MAX}	30 µm

Figure 4.8: (a) Definition of design parameters of the see-saw relay. (b) Values of constant parameters used for the monotonicity analysis.

4.3.1.2. Problem Statement

To optimize the see-saw relay design, it is first necessary to state the problem in terms of an objective function and design constraints. Because one ultimate goal of the see-saw relay design is to achieve a low power operation, an intuitive objective function is the pull-in voltage (V_{PI}) -- the voltage at which either side of a see-saw relay unconditionally snaps closed -- which is derived by comparing the spring restoring force (F_{SPRING}) to the electrostatic actuation force (F_{ELEC}).

Specifically, F_{ELEC} increases quadratically as this force makes one particular side of the movable see-saw structure move downward, whereas F_{SPRING} increases linearly with increasing downward displacement. Therefore, there is a critical displacement beyond which F_{ELEC} is always larger than F_{SPRING}, so that the contact gap will eventually close abruptly, as the applied voltage across the actuation gap increases. This phenomenon, referred to as "pull-in", is dependent on relay design parameters and can be quantified by the voltage at which it occurs [13], [28]:

$$\text{Objective: } V_{PI} = \sqrt{\frac{8}{27} \cdot \frac{k_T T_{ACT}^3}{\varepsilon_O L_A W_A}},$$

where ε_O is the permittivity of free space, T_{ACT} represents the distance between the movable gate electrode and the fixed body electrode on each side, as illustrated in Figure 4.8(a), L_A and W_A represent the corresponding labeled dimensions in the same figure, and k_T represents the torsional spring constant that characterizes the restoring force generated in the structure's anchors, as they rotate.

Of note is that k_T can be derived, to the first degree approximation, as follows. In general, the torque, T_{BEAM}, applied at one end of a torsional beam with rectangular cross-section is given by [29]

$$T_{\mathrm{BEAM}} = k_{\mathrm{T}}''\cdot\phi = \beta\cdot W_{\mathrm{S}}\cdot H^{3}\cdot\frac{G}{L_{\mathrm{S}}}\cdot\phi\ ,$$

where k_{T}'' is the torsional spring constant in N·m/rad, ϕ is the rotation angle that the beam makes, as it rotates, β is a dimensionless parameter that is related to the aspect ratio of the cross-section of the rectangular torsional beam, W_{S} and L_{S} are the width and length of the torsional beam, respectively, as shown in Figure 4.8(a), H is the thickness of the beam as shown in the same figure, and G is the shear modulus of the beam material.

Assuming that the torque applied to make the see-saw structure rotate ($T_{\mathrm{SEE\text{-}SAW}}$) is equally distributed on both of the torsional beams, as illustrated in Figure 4.8(a),

$$T_{\mathrm{SEE\text{-}SAW}} = 2\cdot T_{\mathrm{BEAM}} = 2\beta\cdot W_{\mathrm{S}}\cdot H^{3}\cdot\frac{G}{L_{\mathrm{S}}}\cdot\phi\ .$$

Because the small angle approximation is valid (*i.e.* $\phi \ll 10^{\circ}$; $\phi \approx 0.14^{\circ}$ for the prototype see-saw design with $L_{\mathrm{A}} + L_{\mathrm{A1}} = 40$ μm), assuming that a concentrated load (F) is applied at the distance (L_{L}) from the torsional axis,

$$T_{\mathrm{SEE\text{-}SAW}} = F\cdot L_{\mathrm{L}} = (k_{\mathrm{T}}\cdot T_{\mathrm{ACT}})\cdot L_{\mathrm{L}} = \big(k_{\mathrm{T}}\cdot L_{\mathrm{L}}\cdot\tan(\phi)\big)\cdot L_{\mathrm{L}} = (k_{\mathrm{T}}\cdot L_{\mathrm{L}}\cdot\phi)\cdot L_{\mathrm{L}} = k_{\mathrm{T}}\cdot L_{\mathrm{L}}^{2}\cdot\phi\ ,$$

where k_{T} is the torsional spring constant in N/m and L_{L} is the distance between the torsional axis and the point at which the distributed force from the electrostatic actuation is modeled by a point load.

Next, equating the two analytical equations for T_{PLATE},

$$2\beta\cdot W_{\mathrm{S}}\cdot H^{3}\cdot\frac{G}{L_{\mathrm{S}}}\cdot\phi = k_{\mathrm{T}}\cdot L_{\mathrm{L}}^{2}\cdot\phi\ ,$$

which in turn leads to the first equality constraint (h_1), as follows:

$$h_1\colon k_{\mathrm{T}} = \frac{2\beta\cdot W_{\mathrm{S}}\cdot H^{3}\cdot G}{L_{\mathrm{S}}\cdot L_{\mathrm{L}}^{2}}\ ,$$

Additionally, equating L_{L} in terms of the relevant design parameters yields the second equality constraint (h_2):

$$h_2\colon L_{\mathrm{L}} = 0.5\cdot L_{\mathrm{A}} + L_{\mathrm{A1}},$$

where L_{A1} is the offset between each body electrode and the torsional axis, as shown in Figure 4.8(a).

In order to ensure that the torsional mode operation of the see-saw relay dominates over the flexural mode operation while it is switching, it is necessary to include a constraint on the torsional spring constant relative to the flexural spring constant (k_F). Note that k_F represents the restoring force of the anchors, as they deflect vertically; if the device tends to deflect too far vertically during switching operations, the entire movable structure can be stuck down onto the substrate. As assumed in Chapter 4.3.1.1, since the see-saw plate is a rigid body, k_F of the two torsional beams is given by [29], [30]

$$h_3\text{: } k_F = 2E \cdot W_S \cdot \left(\frac{H}{L_S}\right)^3 ,$$

where E is the Young's modulus of the structural material. The spring ratio constraint (g_1) is expressed as a function of the constant, R_S:

$$g_1\text{: } \frac{k_F}{k_T} \geq R_S \,.$$

The other constraint that is required for proper operation as a component of a logic circuit is that the device be able to restore itself to its initial position after being actuated. To ensure this, the restoring force of the torsional spring needs to be greater than the surface adhesion force:

$$g_2\text{: } k_T \cdot T_{ACT} \geq F_{ADHESION} \,.$$

In addition, constraints are introduced to ensure that the see-saw relay can be fabricated in the Berkeley Microfabrication Laboratory. In particular, no device dimension is allowed to be less than a minimum feature size, λ:

$$g_3\text{: } L_A \geq \lambda ,$$

$$g_4\text{: } W_A \geq \lambda ,$$

$$g_5\text{: } L_S \geq \lambda ,$$

$$g_6\text{: } W_S \geq \lambda ,$$

and the ratio of the length and width dimensions is constrained to confine the strain gradient effect to one dimension:

$$g_7\text{: } L_A + L_{A1} \geq 2W_A \,.$$

In addition, practical design rules (which had been verified by the prior experiment) are introduced to address the potential problems of buckling or fracture under cycling stress. The most relevant of these concerns is a constraint on the minimum possible actuation offset:

$$g_8 : L_{A1} \geq 2\lambda \, .$$

Finally, the device needs to be confined to a finite area:

$$g_9 : L_{MAX} \geq 2L_A + 2L_{A1} + W_S \, ,$$

$$g_{10} : W_{MAX} \geq W_A + 2L_S \, ,$$

where L_{MAX} and W_{MAX} are constants that specify the maximum allowable footprint of the see-saw relay.

These constraints are compiled into a form amenable to monotonicity analysis, as follows:

$$g_1 : k_T \cdot R_S - k_F \leq 0 \, ,$$

$$g_2 : F_{ADHESION} - k_T \cdot T_{ACT} \leq 0 \, ,$$

$$g_3 : \lambda - L_A \leq 0 \, ,$$

$$g_4 : \lambda - W_A \leq 0 \, ,$$

$$g_5 : \lambda - L_S \leq 0 \, ,$$

$$g_6 : \lambda - W_S \leq 0 \, ,$$

$$g_7 : 2W_A - L_A + L_{A1} \leq 0 \, ,$$

$$g_8 : 2\lambda - L_{A1} \leq 0 \, ,$$

$$g_9 : 2L_A + 2L_{A1} + W_S - L_{MAX} \leq 0 \, ,$$

$$\mathrm{g}_{10} : W_A + 2L_S - W_{MAX} \leq 0 \, ,$$

$$h_1 : k_T \cdot L_S \cdot L_L^2 = 2\beta \cdot W_S \cdot H^2 \cdot G \, ,$$

$$h_2 : L_L = 0.5 \cdot L_A + L_{A1},$$

$$h_3 : k_F \cdot L_S^3 - 2E \cdot W_S \cdot H^3 = 0 \, ,$$

and the results are compiled into an always true monotonicity table, Table 4.1.

Note that many of the line width constraints (g_3, g_4, and g_5) are irrelevant, because other constraints already limit these dimensions. For instance, if g_9 and g_7 or g_{10} constraints are active,

Table 4.1: Always-true monotonicity table.

	k_T	L_A	W_A	k_F	L_S	W_S	L_L	L_{A1}	Activity
Obj.: $(8k_T \cdot T_{ACT}^3 / (27\varepsilon_O \cdot L_A \cdot W_A))^{0.5}$	+	–	–						Always active
g_1: $k_T \cdot R_S - k_F$	+			–					Activated if h_3 is active
g_2: $F_{ADHESION} - k_T \cdot T_{ACT}$	–								
g_3: $\lambda - L_A$		–							Irrelevant
g_4: $\lambda - W_A$			–						Irrelevant
g_5: $\lambda - L_S$					–				Irrelevant
g_6: $\lambda - W_S$						–			Always active for realistic $F_{ADHESION}$
g_7: $2W_A - L_A - L_{A1}$		–	+					–	
g_8: $2\lambda - L_{A1}$								–	
g_9: $2L_A + 2L_{A1} + W_S - L_{MAX}$		+				+		+	
g_{10}: $W_A + 2L_S - W_{MAX}$			+		+				
h_1: $L_S \cdot L_L^2 \cdot k_T = 2\beta \cdot W_S \cdot H^2 \cdot G$	–/+				–/+	+/–	–/+		Activated if h_2 is active
h_2: $L_L = 0.5 \cdot L_A + L_{A1}$		–/+					+/–	–/+	Activated if h_1 is active
h_3: $k_F \cdot L_S^3 - 2E \cdot H^3 \cdot W_S = 0$				+	+	–			Activated if g_1 is active

the constraints g_3, g_4, and g_5 become irrelevant. The constraint, g_6, however, must be always active, because a minimized W_S reduces the objective function, but the ratio of spring constants is not negatively penalized. In the case of very strong surface adhesion, it might be necessary to increase W_S beyond this minimum feature size to generate a high enough restoring force in the structure.

In addition, there are two pairs of linked constraints: h_1 and h_2 must always be activated together, as must g_1 and h_3. This is because the values L_L and K_F appear only in these pairs, respectively, and providing a positive and negative bound on L_L and K_F requires that both of the constraints be active. For this reason, h_3 can only be activated with one set of signs, even though it is an equality constraint.

As shown in Table 4.1, the always-true table reveals a large number of possible cases and sub-cases, because almost all of the variable columns have multiple possible positive or negative bounds. This makes pure monotonicity analysis of this structure somewhat challenging. By building from simple cases to more complex ones, however, it is possible to make an understandable logic flow that addresses a set of realistic design cases. A logic flow designed for the monotonicity analysis is featured in Figure 4.9.

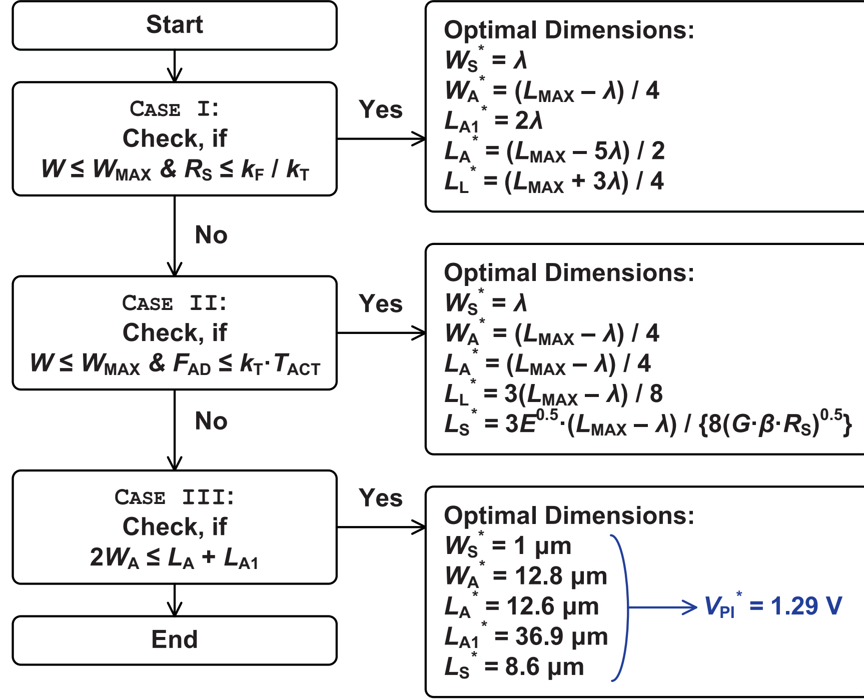


Figure 4.9: Logic flow that indicates the three cases of the monotonicity analysis.

4.3.1.3. Case Study I

The first case uses the minimum number of constraints to frame the monotonicity problem, as shown in Table 4.2. The objective function activates a positive constraint on k_T and negative constraints on L_A and W_A. In this case, the k_T constraint is bounded by the surface adhesion force constraint (g_2) to prevent the activation of the variable-rich equality constraints, h_1 and hence h_2. L_A is limited by the L_{MAX} constraint (g_9). W_A is limited by the strain confinement constraint (g_7), which prevents the introduction of L_S as a variable. Lastly, W_S and L_{A1} are also limited by the L_{MAX} constraint (g_9).

A set of optimal solutions can be found from **CASE I**. Specifically, first consider the values that can be pinned by the active constraints, which are distinguished from the other variables with a star superscript:

$$W_S^* = \lambda \, ,$$

$$k_T^* = \frac{F_{\text{ADHESION}}}{T_{\text{ACT}}} \, .$$

Table 4.2: Minimal constraint monotonicity table.

	k_T	L_A	W_A	W_S	L_{A1}	Activity
Obj.: $(8k_T \cdot T_{ACT}^3 / (27\varepsilon_O \cdot L_A \cdot W_A))^{0.5}$	+	–	–			Always active
g_2: $F_{ADHESION} - k_T \cdot T_{ACT}$	–					Active to bound k_T in Obj.
g_6: $\lambda - W_S$				–		Active to bound W_S in g_9.
g_7: $2W_A - L_A - L_{A1}$		–	+		–	Active to bound W_A in Obj.
g_8: $2\lambda - L_{A1}$					–	Active to bound L_{A1} in g_9.
g_9: $2L_A + 2L_{A1} + W_S - L_{MAX}$		+		+	+	Active to bound L_A in Obj.

Next, the constraints (g_7 and g_9) that contain L_{A1} and L_A give rise to a master equality:

$$L_A + L_{A1} = 2W_A = \frac{L_{MAX} - W_S}{2} ,$$

which leads to an optimal design point for W_A:

$$W_A^* = \frac{L_{MAX} - \lambda}{4} .$$

Since the offset constraint (g_8) is active to minimize L_{A1}, the rest of the dimensions are found to be

$$L_{A1}^* = 2\lambda ,$$

$$L_A^* = \frac{L_{MAX} - 5\lambda}{2} .$$

In addition, even though the variable, L_L, is inactive (*i.e.* initially unconstrained), it can be computed from the unconstrained L_L constraint (h_2), because L_{A1}^* and L_A^* are already pinned by the active constraints:

$$L_L = 0.5 \cdot L_A + L_{A1} \rightarrow$$

$$L_L^* = 0.5 \cdot L_A^* + L_{A1}^* = \frac{L_{MAX} + 3\lambda}{4} .$$

Of particular interest in **`CASE I`** is the decision to minimize L_{A1}. This is because the objective function is not dependent on L_{A1}, as long as g_2 is constrained, *i.e.* k_T is tied to $F_{ADHESION} / T_{ACT}$. That is to say, while g_2 is active, the objective function is only affected by changes in L_A and W_A. Also of note is that L_S can be computed from the inactive k_T constraint (h_1), as follows:

$$L_S = \frac{32\beta\cdot\lambda\cdot G\cdot H^3\cdot T_{ACT}}{(L_{MAX} + 3\lambda)^2\cdot F_{ADHESION}} .$$

Because this variable, L_S, is not explicitly constrained by **CASE I**, it forms the basis for both of the following tests, which indicate if **CASE I** is indeed valid: a check to see if the total width constraint is obeyed and a test to confirm that the spring constant ratio (a function of L_S) is valid. These are noted in the logic flow in Figure 4.9, but are more precisely stated as follows:

$$W_A + 2L_S = \frac{L_{MAX} - \lambda}{4} + \frac{64\beta\cdot\lambda\cdot G\cdot H^3\cdot T_{ACT}}{(L_{MAX} + 3\lambda)^2\cdot F_{ADHESION}} \leq W_{MAX} ,$$

$$\frac{k_F}{k_T} = \frac{2E\cdot F^2_{ADHESION}}{T^2_{ACT}\cdot\lambda^2}\cdot\left(\frac{(L_{MAX} + 3\lambda)^2}{32\beta\cdot G\cdot H}\right)^3 \geq R_S .$$

Physically, the set of the optimal solutions (the variables with a star superscript) found in **CASE I** indicates a somewhat ideal design, where the see-saw device design is not initially constrained by W_{MAX}, and where structural material properties are assumed such that, even with a minimized L_L, both a smallest possible k_T and a sufficiently large k_F that obey the R_S constraint (g_1) can be achieved. Therefore, **CASE I** represents a good alignment of the design constants to allow the highest degree of optimization.

4.3.1.4. Case Study II

The next interesting case arises when the initial constraints in **CASE I** fail due to a high R_S and more limited area. These constraints result in a set of active constraints that are complementary to the active g_2 in **CASE I**; by unpinning k_T from its minimum value (*i.e.* remove the g_2 constraint from the monotonicity table) and by enforcing the stiffness ratio constraint (g_1), the opposite set of functionality constraints can be addressed. Note that L_{A1} is also unpinned from its minimum value. The new set of active constraints is summarized in Table 4.3 and gives rise to a new set of optimal solutions. Again, starting with the quantities that are directly defined by their active constraints:

$$W^*_S = \lambda .$$

The objective function can be expressed as a function of only one variable, L_L, by recognizing that

$$W^*_A = \frac{L_{MAX} - \lambda}{4} \text{ (from } g_7 \text{ and } g_9) ,$$

Table 4.3: Monotonicity table for **CASE II**.

	k_T	L_A	W_A	k_F	L_S	W_S	L_L	L_{A1}	Activity
Obj.: $(8k_T \cdot T_{ACT}{}^3 / (27\varepsilon_O \cdot L_A \cdot W_A))^{0.5}$	+	–	–						Always active
g_1: $k_T \cdot R_S – k_F$	+			–					Active to bound k_F in h_3.
g_6: $\lambda – W_S$						–			Active to bound W_S in h_1.
g_7: $2W_A – L_A – L_{A1}$		–	+					–	Active to bound W_A in Obj.
g_9: $2L_A + 2L_{A1} + W_S – L_{MAX}$		+				+		+	Active to bound L_A in Obj.
h_1: $2\beta \cdot W_S \cdot H^2 \cdot G – L_S \cdot L_L{}^2 \cdot k_T = 0$	–				–	+	–		Active to bound k_T in Obj.
h_2: $L_L – 0.5 \cdot L_A – L_{A1} = 0$		–					+	–	Active to bound L_L in h_1 & L_{A1} in g_9
h_3: $k_F \cdot L_S{}^3 – 2E \cdot H^3 \cdot W_S = 0$				+	+	–			Active to bound L_S in h_1.

$$L_A = L_{MAX} - \lambda - 2L_L \text{ (from } g_9 \text{ and } h_2\text{)} ,$$

$$L_S^2 = \frac{E}{G} \cdot \frac{L_L^2}{\beta \cdot R_S} \text{ (from } g_1, h_1, \text{ and } h_3\text{)} ,$$

which leads to an objective function of the following form:

$$V_{PI}^2 = \frac{64\beta \cdot H^3 \cdot G \cdot T_{ACT}^3 \cdot \lambda \cdot \sqrt{\beta \cdot R_S \cdot G/E}}{27\epsilon_O \cdot (L_{MAX} - \lambda) \cdot (L_{MAX} - \lambda - 2L_L) \cdot L_L^3} \triangleq \frac{\text{Constant}}{(L_{MAX} - \lambda - 2L_L) \cdot L_L^3} .$$

The optimal L_L can be found by differentiating the objective function with respect to L_L and equating the resulting expression to zero:

$$L_L^* = \frac{3}{8}(L_{MAX} - \lambda) ,$$

and this expression is used to solve for the other critical dimensions:

$$L_A^* = \frac{L_{MAX} - \lambda}{4} ,$$

$$L_S^* = \frac{3}{8}\sqrt{\frac{E}{G} \cdot \frac{1}{\beta \cdot R_S}}(L_{MAX} - \lambda) .$$

In **CASE II**, since the spring restoring force constrain (g_2) is not enforced, it is necessary to check the values found for $k_T \cdot T_{ACT}$ against $F_{ADHESION}$ to determine if **CASE II** is the valid case. In addition, the W_{MAX} constraint (g_{10}) is also inactive, and hence, it is necessary to check if the

device indeed fits in the appropriate footprint. These checks are summarized in Figure 4.9, but are given more specifically as follows:

$$\frac{F_{\mathrm{ADHESION}}}{T_{\mathrm{ACT}}} \leq k_{\mathrm{T}} = \left(\frac{8}{3}\right)^{3} \sqrt{\frac{G}{E} \cdot \beta \cdot R_{\mathrm{S}}} \frac{2\beta \cdot \lambda \cdot H^{3} \cdot G}{(L_{\mathrm{MAX}} - \lambda)^{3}} ,$$

$$\left(\frac{1}{4} + \frac{3}{4}\sqrt{\frac{E}{G} \cdot \frac{1}{\beta \cdot R_{\mathrm{S}}}}\right) \cdot (L_{\mathrm{MAX}} - \lambda) \leq W_{\mathrm{MAX}} .$$

CASE II represents one design scenario, in which the minimum torsional spring constant (k_{T}) might be unattainable, and hence makes the relay device harder to achieve lower operating voltages, if limited device footprint is given (for example, L_{MAX} is required to be very small). In addition, the flexural spring constant (k_{F}) is might be on the verge of being large enough, because the stiffness ratio constraint (g_1) is forced to be active. Therefore, **CASE II** is in some ways the opposite of the first case, perhaps reflecting some of the worst possible design conditions.

4.3.1.5. Case Study III

To find the case most representative of reality, a balance between the first two cases is investigated by enforcing both the stiffness ratio (g_1) and the surface adhesion limit on k_{T} (g_2). The constraint on the maximum allowable width (g_{10}) is chosen to be active, in order to limit the device width with a given footprint. This latter addition represents a departure from the assumptions in the first two cases, where both allowed an inactive W_{MAX}. This is a justifiable change, because it makes **CASE III** closer to the prototype see-saw relays that are the basis for this optimization. Even though it is possible to represent a version of this case where W_{A} is limited by the g_7 constraint (and to represent the previous two cases with the g_{10}-based constraints), doing so would be repetitive because, fundamentally, only one of the g_7 or g_{10} constraints can limit W_{A} and changing which constraint is active only affects the device width without changing other variables. These changes are summarized in Table 4.4.

A set of optimal solutions for **CASE III** can be found in the same way as the previous two cases, noticing that a series of the active constraints reveals the design variable values as follows:

$$W_{\mathrm{S}}^{*} = \lambda ,$$

$$k_{\mathrm{T}}^{*} = \frac{F_{\mathrm{ADHESION}}}{T_{\mathrm{ACT}}} ,$$

Table 4.4: Monotonicity table for **CASE III**.

	k_T	L_A	W_A	k_F	L_S	W_S	L_L	L_{A1}	Activity
Obj.: $(8k_T \cdot T_{ACT}^3 / (27\varepsilon_O \cdot L_A \cdot W_A))^{0.5}$	+	–	–						**Always active**
g_1: $k_T \cdot R_S - K_F$	+			–					**Active to bound k_F in h_3.**
g_2: $F_{ADHESION} - k_T \cdot T_{ACT}$	–								**Active to bound k_T in Obj.**
g_6: $\lambda - W_S$						–			**Active to bound W_S in h_1.**
g_9: $2L_A + 2L_{A1} + W_S - L_{MAX}$		+				+		+	**Active to bound L_A in Obj.**
g_{10}: $W_A + 2L_S - W_{MAX}$			+		+				**Active to bound W_A in Obj.**
h_1: $2\beta \cdot W_S \cdot H^2 \cdot G - L_S \cdot L_L^2 \cdot k_T = 0$	–				–	+	–		**Active to bound W_S in g_6**
h_2: $L_L - 0.5 \cdot L_A - L_{A1} = 0$		–					+	–	**Active to bound L_L in h_1 & L_{A1} in g_9**
h_3: $k_F \cdot L_S^3 - 2E \cdot H^3 \cdot W_S = 0$				+	+	–			**Active to bound L_S in h_1.**

$$k_F^* = k_T^* \cdot R_S = \frac{F_{ADHESION} \cdot R_S}{T_{ACT}} \,,$$

$$k_F = \frac{2E \cdot W_S \cdot H^3}{L_S^3} \rightarrow L_S^* = \sqrt[3]{\frac{2E \cdot \lambda \cdot T_{ACT}}{R_S \cdot F_{ADHESION}}} \cdot H \,,$$

$$k_T = \frac{2\beta \cdot W_S \cdot H^3 \cdot G}{L_S \cdot L_L^2} \rightarrow L_L^* = \sqrt{\beta \cdot G \cdot \left(\frac{R_S}{E}\right)^{1/3} \cdot \left(\frac{2\lambda \cdot T_{ACT}}{F_{ADHESION}}\right)^{2/3}} \cdot H \,,$$

$$L_A^* = L_{MAX} - \lambda - 2 \cdot \sqrt{\beta \cdot G \cdot \left(\frac{R_S}{E}\right)^{1/3} \cdot \left(\frac{2\lambda \cdot T_{ACT}}{F_{ADHESION}}\right)^{2/3}} \cdot H \,,$$

$$L_{A1}^* = 2\sqrt{\beta \cdot G \cdot \left(\frac{R_S}{E}\right)^{1/3} \cdot \left(\frac{2\lambda \cdot T_{ACT}}{F_{ADHESION}}\right)^{2/3}} \cdot H - \frac{L_{MAX} - \lambda}{2} \,,$$

$$W_A^* = W_{MAX} - 2 \cdot \sqrt[3]{\frac{2E \cdot \lambda \cdot T_{ACT}}{R_S \cdot F_{ADHESION}}} \cdot H \,.$$

To determine that **CASE III** is indeed a valid case, the values found for $L_A + L_{A1}$ need to be checked against $2W_A$ (note that the g_7 constraint was set to be inactive). This is shown in Figure 4.9, and the specific check is given as:

$$W_{\mathrm{MAX}} \leq \frac{L_{\mathrm{MAX}} - \lambda}{4} + 2 \cdot \sqrt[3]{\frac{2E \cdot \lambda \cdot T_{\mathrm{ACT}}}{R_{\mathrm{S}} \cdot F_{\mathrm{ADHESION}}}} \cdot H \,.$$

Because almost all of the constraints are enforced in this case, it physically represents a design where both spring constants are at their minimum possible value (while maintaining the stiffness ratio), which results in the lowest V_{PI} in a given footprint.

4.3.1.6. Fundamental Insights from the Case Studies

In summary, two key lessons are embedded in the three case studies. First, the width of the see-saw relay (W_{A}) is set by one of the g_7 or g_{10} constraints and has little bearing on other aspects of the design, such as L_{S}, L_{L}, and the spring constants. Second, there are three possible approaches that make the spring constants be limited: k_{T} is at its minimum value and k_{F} is not, k_{F} is at its minimum value and k_{T} is not, or both k_{T} and k_{F} are at their minimum values. The latter of the three cases is closest to the ideal and yields the best V_{PI} performance per unit footprint. Also, in order to minimize V_{PI} with the smallest possible footprint overhead, W_{MAX} should be set such that W_{A} is precisely at the limit allowed by g_7, which would suggest that both of the W_{A} constraints (g_7 and g_{10}) are effectively active for the most optimally selected set of design constants.

4.3.2. Validation of Monotonicity Analysis

The monotonicity analysis is verified against constrained numerical optimization in **MATLAB®** [31] as well as FEM analysis in **CoventorWare™** [32], [33]. Each of the methods well predicts the value of the objective function, which agrees with each other within approximately 10 %.

4.3.2.1. Monotonicity Analysis vs. Nonlinear Constrained Optimization

The closed-form analytical solutions obtained from **CASE III** of the monotonicity analysis were evaluated with the constant values (summarized in Figure 4.8(b)) used in the prototype see-saw relay design. The results of the monotonicity analysis are plotted in Figure 4.10(a); the lowest V_{PI} of the see-saw relay was found to be 1.29 V, when L_{S} = 8.6 μm, W_{S} = 1 μm, L_{A1} = 36.9 μm, L_{A} = 12.6 μm, and W_{A} = 12.8 μm. To validate the analytical results of **CASE III**, the Rosenbrock Test Function ("**fmincon**" function in **MATLAB®** built-in for nonlinear constrained minimization [31]) was used to find a global optimal V_{PI} for the see-saw relay. All of the inequality and equality constraints were entered into **MATLAB®** with the same constant values

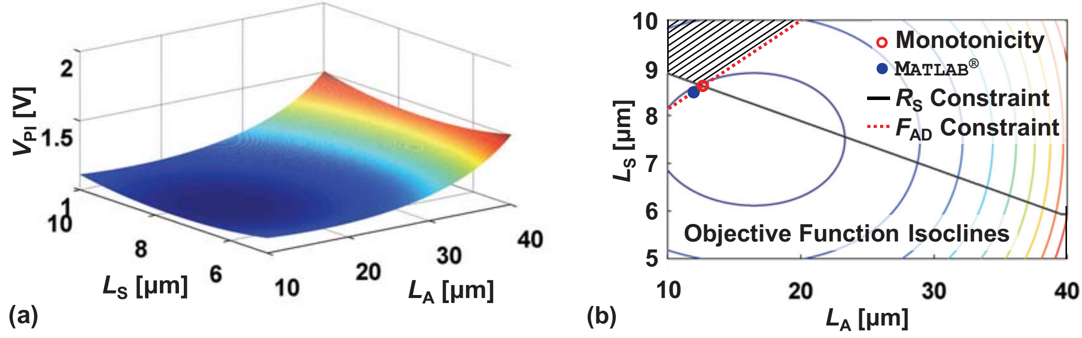


Figure 4.10: Analytical results of **`CASE III`** of the monotonicity analysis. (a) Three-dimensional plot of V_{PI} against L_A and L_S. (b) Contour plot of V_{PI} with a shaded region representing valid design space that satisfies the **`CASE III`** constraints. The V_{PI} of the monotonicity analysis (open dot) is close to that of the **`MATLAB`**® analysis (closed dot). The slight difference comes from the convergence issue that necessitated the use of a low tolerance with the numerical optimizer.

(Figure 4.8(b)). This numerical analysis yielded the optimal V_{PI} of 1.49 V, as indicated in Figure 4.10(b).

The results of both analyses match closely. This is non-trivial, especially when considering the difficulty of forcing an eight-variable constrained optimization to converge; the tolerances of the optimization were loose and account for the difference between the monotonicity and numerical optimum points. Note that convergence aids were used in finding the numerical solution, *e.g.*, a lower bound of zero on the search space was used, instead of negative infinity, which might have resulted in slight inconsistency (the closed dot is outside the valid region in Figure 4.10(b)).

4.3.2.2. Monotonicity Analysis vs. Finite-Element-Method Analysis

FEM analysis was performed on the see-saw relay with the optimal set of the device dimensions found from the monotonicity analysis. The result of the FEM simulation is shown in Figure 4.11. The values of the objective function predicted by the analyses are summarized in Table 4.5; The close agreement of the values of the objective function between the methods verifies that the monotonicity analysis indeed produced a valid, optimal design solution.

4.3.3. Summary of Design Optimization Study

The results of the analyses lead to an insight into the proper optimizations for the see-saw relay: the torsional spring constant (k_T) should be minimized, as should the flexural spring constant (k_F),

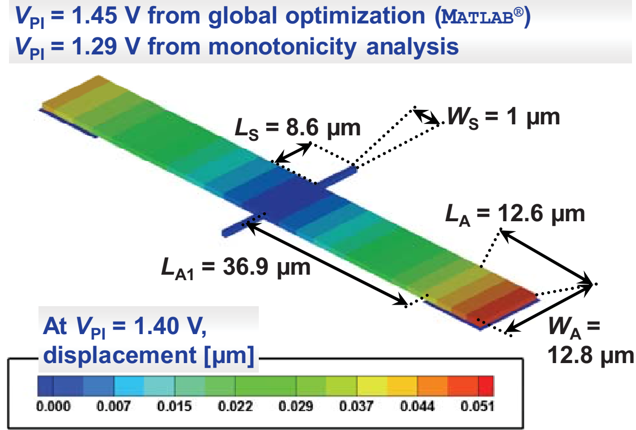


Figure 4.11: FEM simulation of the see-saw relay. V_{PI} matches both of the predictions to within 10%.

Table 4.5: The values of the objective function from the three optimization methods are compared.

Optimization Method	V_{PI} [V]	Percentage Error
Monotonocity	1.28	11.7
MATLAB® `fmincon` Function	1.49	2.8
Finite Element Method	1.45	N/A

the actuation area should have a reasonable offset from the torsional axis, and the actuation width should be maximized under the constraints of the strain gradient and footprint limitations.

Monotonicity analysis also revealed some underlying facts about the design cases. The constraint on device width has little effect on other variable parameters, and hence, the ideal footprint should be assigned to be reasonably lengthy enough (in the direction perpendicular to the torsional axis) such that the device width (W_F) is limited by the stress confinement constraint. In addition, there are three possible routines in monotonicity analysis that the torsional and flexural spring constants can decide device dimensions: only the k_T constraint can be active, only the k_F constraint can be active, or both spring constant constraints can be activated.

Finding these design insights with monotonicity analysis was somewhat challenging, because of multiple design constraints that affected the same variables. This suggests that the formulation of monotonicity design problems is vastly simplified by finding the dominant constraints and focusing on those, while ignoring corner cases that make the scale of the monotonicity table explode. This was accomplished by identifying the three dominant spring limitations and two dominant width limitations mentioned previously. It is also interesting to note that the monotonicity analysis revealed the tolerance and convergence limitations of numerical optimization, nevertheless, it is particularly useful in its initially proscribed role as a pre-optimization tool.

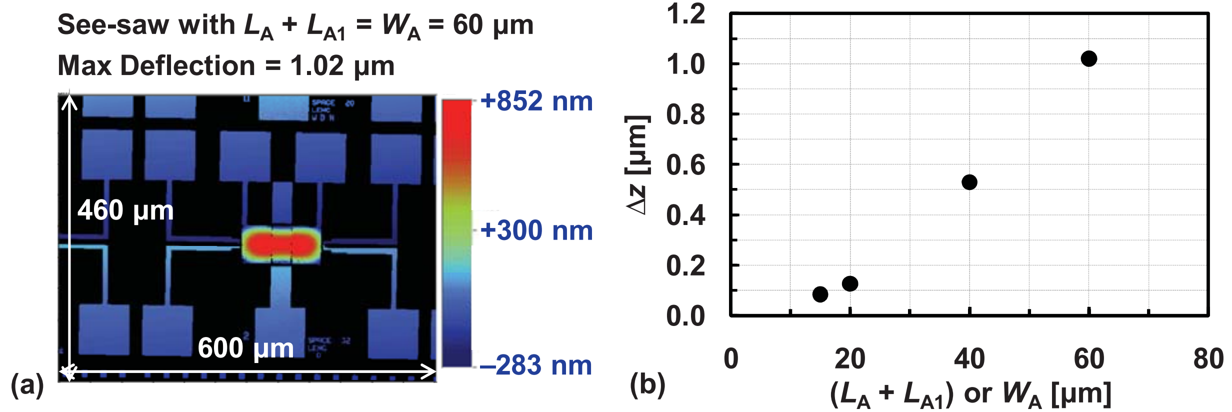


Figure 4.12: (a) Negative strain gradient within the poly-$Si_{0.4}Ge_{0.6}$ layer produces severe out-of-plane warpage. (b) The parasitic effect becomes more apparent, as the size of the structure becomes larger.

4.4. Corrugated See-Saw Relay Design

The case studies in the previous section were performed, based on the assumption that the structural material of the see-saw relay shows zero strain gradient effect, *i.e.* there is no out-of-plane curvature of the released structure. This is difficult to realize in practice, because intrinsic stress (and stress gradient) of a material -- which results in dimensional change and out-of-plane deflection upon removal of the sacrificial oxide underneath the pre-released structure -- depends on deposition conditions as well as film micro-grain structure. For example, undoped polycrystalline-Silicon (poly-Si) films are amorphous in nature when they are deposited at 550 ºC or below; The internal microstructure and residual stress of this material is dependent upon processing temperature and partial pressure of Silane (SiH_4), and thus, a low temperature anneal that leads to a fine-grained poly-Si is necessary to achieve low tensile stress and smooth surface texture [34].

When the gate stack of the see-saw relay (*e.g.* 1-µm-thick p+ poly-$Si_{0.4}Ge_{0.6}$ on top of 40-nm-thick Al_2O_3) has positive strain gradient, its bottom portion is more inclined to stretch out than its top portion upon release, because it is increasingly compressed toward its bottom. On the other hand, if the gate stack has negative strain gradient, the top portion of the gate stack tends to stretch out more than the bottom portion upon release, because the gate stack is increasingly compressed toward its top portion.

The negative strain gradient was certainly the case for the pre-laser-annealed see-saw relays, as confirmed using a white-light interferometer (Veeco Wyko NT3300 3-D Optical Profiling System) in Figure 4.12(a); Since the see-saw structure is anchored by the two torsional beams, it is increasingly buckled upward at its center line along the torsional axis as well as along the direction perpendicular to the torsional axis (*i.e.* it appears like an open umbrella). Figure 4.12(b) shows measured parasitic out-of-plane displacement of the see-saw relays, which increases with increasing device dimensions, because of more severe negative strain gradient.

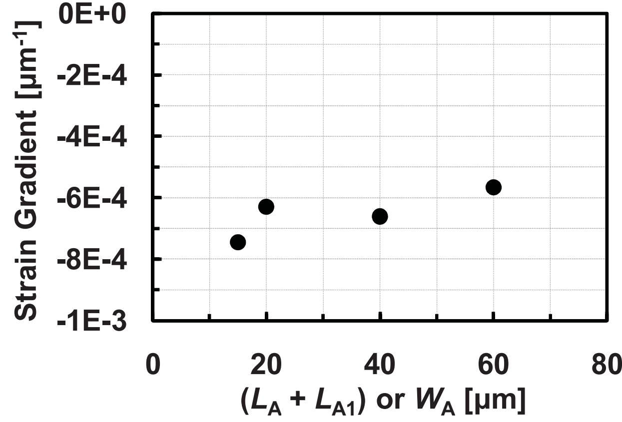


Figure 4.13: The values of negative strain gradient within the poly-$Si_{0.4}Ge_{0.6}$ layer are on the order of high 10^{-4} /μm. Ideally, strain gradient on the order of 10^{-5} /μm is desirable for the minimal impact of strain gradient on device performance.

Quantitatively, strain gradient is calculated -- to the first degree approximation -- from [35]

$$\frac{1}{\rho} = \frac{2\Delta z}{L_A + L_{A1}} ,$$

where ρ is the radius of the curvature of the see-saw structure, Δz is the out-of-plane deflection of the see-saw structure, L_A represents the length of the see-saw structure, and L_{A1} is the offset between each body electrode and the torsional axis, as Figure 4.8(a) shows. Using this equation, the strain gradient values of the see-saw structures were approximated, as shown in Figure 4.13. The calculated values on the order of high 10^{-4} per μm are relatively large, considering that inertial sensor applications typically require strain gradient of 10^{-5} per μm, which translates into 1.25 μm tip deflection of a 500-μm-long cantilever beam [35].

Clearly, minimization of the strain gradient effect is desirable, to achieve low operating voltages, since these require the formation of very small actuation and contact gaps with good uniformity.

In fact, various processing techniques have been investigated, in order to control the stress (and stress gradient) by engineering the structural and morphological properties of a film. For example, intrinsic stress can be minimized by properly tuning material deposition conditions. The hinge material of the Texas Instruments Digital Mirror Display devices -- the amorphous $TiAl_3$ sputtered with 4 % Oxygen -- was found to yield the lowest stress relaxation [36].

Annealing techniques have been proven to be effective in inducing changes in film grain microstructures and hence alleviating the strain gradient effect. For instance, positive stress gradient of an as-deposited poly-SiGe film can be changed to negative stress gradient by various post-deposition processing methods, such as furnace annealing, rapid thermal annealing, or flash lamp annealing; This is possible, because voids and defects within the film are removed, as the amorphous region of the film (which is needed initially to provide crystal seeds) is crystallized while the upper crystalline portion of the film is recrystallized [35], [37], [38].

In contrast, argon implantation or excimer laser annealing with Krypton Fluoride (KrF) or Xenon Dichloride ($XeCl_2$) laser sources produces the reverse effect within the deposited film, *i.e.* negative stress gradient is turned into positive stress gradient because compressive stress within the portion is relieved [35], [39], [40].

These processing approaches, however, might be inadequate for high-yield fabrication of the see-saw relays, because it is difficult to precisely control the stress (and stress gradient) within the gate stack, especially in academic laboratory settings, due to process-induced run-to-run and wafer-to-wafer variations.

Therefore, of particular interest would be the approach that can reduce out-of-plane deflection of the released structure via mechanical design. Lateral poly-Si resonator designs that incorporate symmetric folded flexures have been shown to relieve residual stress within the structural microstructures; however, they are still susceptible to out-of-plane warpage, due to the intrinsic stress gradient through the thickness of the film [41], [42].

Another intriguing approach would be to take advantage of the corrugated structural design [43], [44] -- which was invented for bulk micromachined sensors [45] -- to reduce intrinsic stress within the deposited film. Corrugated single-crystal Si membranes have much lower intrinsic stress (by a factor of 1,000 to 10,000), compared to non-corrugated membranes [46]. Corrugated cantilever beams made out of Chromium (Cr) or poly-SiGe show 20 to 50 times reduction in parasitic out-of-plane deflection because of higher effective moment of inertia that enhances stiffness of the structure [47], [48].

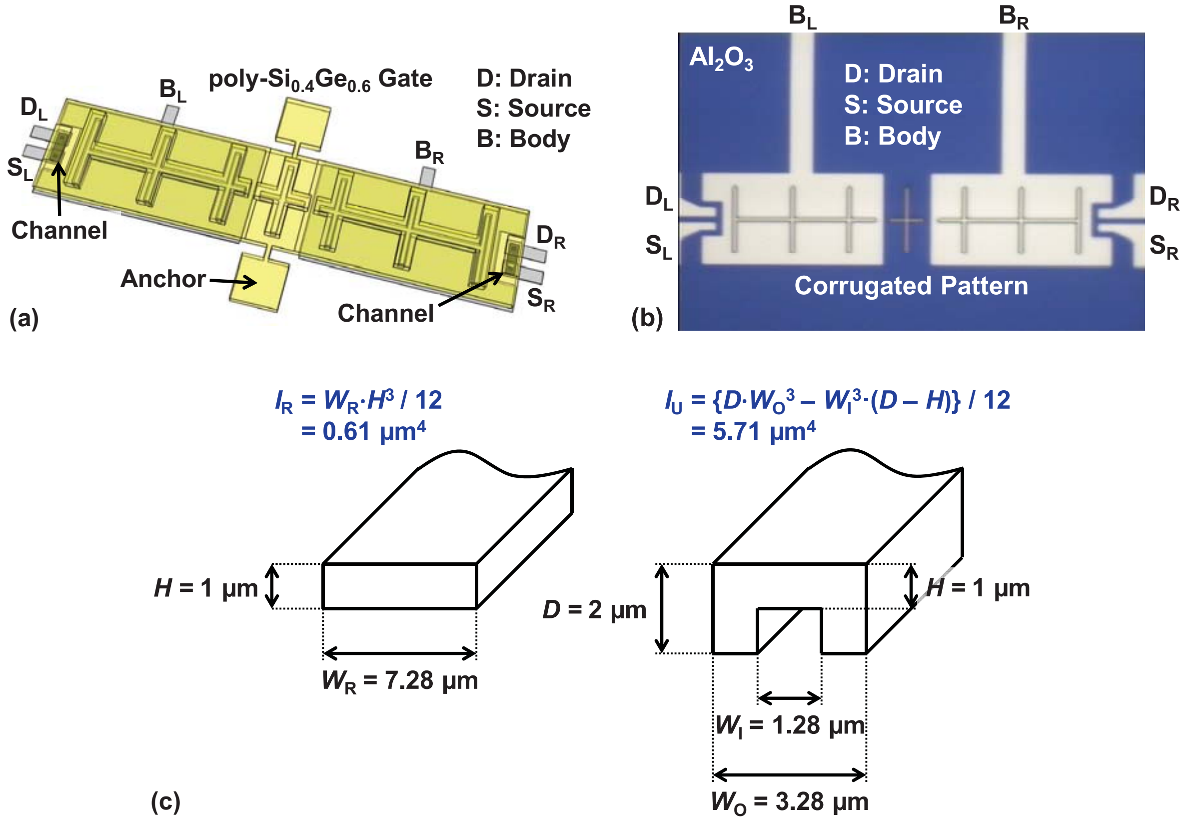


Figure 4.14: (a) Isometric view of the corrugated see-saw relay design. Design parameters are as defined in Figure 4.9 (except the hammerhead region that is removed in the new design), *i.e.* $T_{ACT} = 0.2$ µm, $T_{CONTACT} = 0.1$ µm, $L_S = 3$µm, $W_S = H = 1$ µm, $L_A = 45$ µm, $L_{A1} = 6$ µm, and $W_A = 21.5$ µm. The two stable states for perfectly complementary operation are *off-on* and *on-off*. (b) Optical micrograph of plan-view of a corrugated see-saw relay that was processed up to the source/drain and body electrode definition. (c) Area moment of inertia (I) of the rectangular (left) *vs.* U-shaped (right) cross-sections of a beam.

4.4.1. Device Design and Fabrication

The corrugated see-saw relay design shown in Figure 4.14(a) uses a single moveable plate (which serves as the gate electrode) anchored by two torsional beams which allow the ends of the plate to be displaced up and down in a perfectly complementary fashion. Two W channel electrodes are attached to the p+ poly-$Si_{0.4}Ge_{0.6}$ gate electrode at either end via an intermediary Al_2O_3 gate-oxide film. Two W body electrodes are located underneath the gate electrode, one on either side of the torsional axis. The device operating principle is the same as for the prototype see-saw relay design described earlier in this chapter: Differential bias voltages applied to the body electrodes cause the gate electrode to be electrostatically actuated downward on one side at a time. As a result, one of the channel electrodes makes contact with a pair of W source/drain electrodes (co-planar with the body electrodes), *i.e.* either only the left end or only the right end of the see-saw relay is turned *on* at a time.

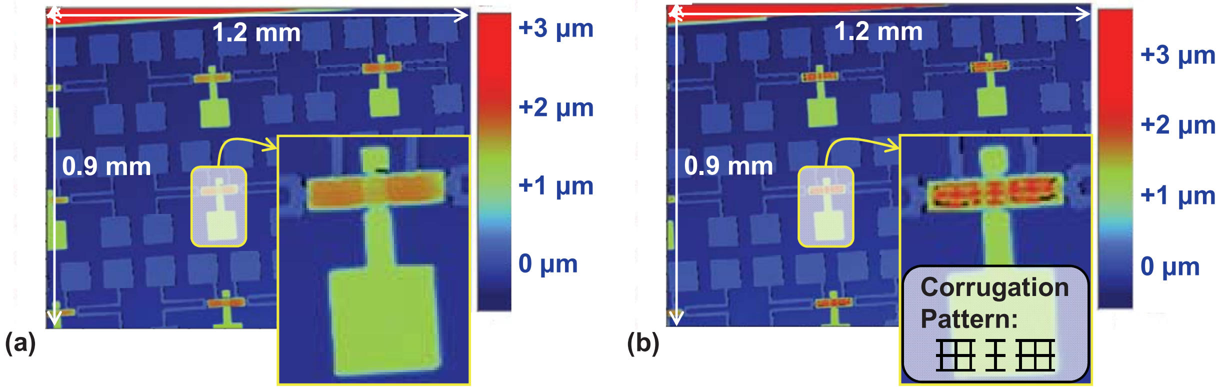


Figure 4.15: Measured out-of-plane deflection of the (a) non-corrugated see-saw relay *vs.* (b) corrugated see-saw relay. The negative strain is still present within the corrugated poly-$Si_{0.4}Ge_{0.6}$ layer.

Note that the corrugated see-saw structure consists of multiple U-shaped beam(s). As described earlier, this U-profile makes the structure stiffer by increasing the moment of inertia (*i.e.* increases the ability/capacity of the structure cross-section to resist bending). The area moment of inertia of the rectangular *vs.* U-shaped cross-sections of a beam are given by [49]

$$I_R \cong \frac{W_R \cdot H^3}{12} \ ,$$

$$I_U \cong \frac{D \cdot W_O^3 - W_I^3 \cdot (D - H)}{12} \ ,$$

respectively, where W_R, H, D, W_O, and W_I are correspond to the labeled dimensions in Figure 4.14(c). The calculated values of I_R *vs.* I_U are summarized in Figure 4.14(c) -- The moment of inertia of the U-shaped beam is higher by approximately 9.4 times than that of the rectangular beam. Note that H is chosen to be 1 μm due to the limited depth of focus (~1 μm) of the photolithography tool that was used to fabricated the relays.

Corrugated see-saw relays were fabricated in the Berkeley Microfabrication and Marvell Nanofabrication Laboratories. The original process flow described in Chapter 2 was slightly modified. The specific changes are as follows: (1) At the beginning of the process, a 1-μm-thick thermal SiO_2 was grown on the Si substrate. This layer was defined using the mask with various corrugation features. A 100 nm Al_2O_3 was then deposited to insulate the corrugated substrate. (2) The timed-etching step used to thin down the sacrificial oxide after the W channel formation was no longer used. To remove this step, a thinner 1st sacrificial oxide (100 nm instead of 200 nm) was deposited and patterned, prior to the 2nd sacrificial oxide deposition.

Fabricated see-saw relays were inspected using the white-light interferometer, as shown in Figure 4.15, to observe the impact of corrugation on strain gradient. The negative strain gradient of the corrugated relay seemed to be slightly lower than that of the non-corrugated one, and yet

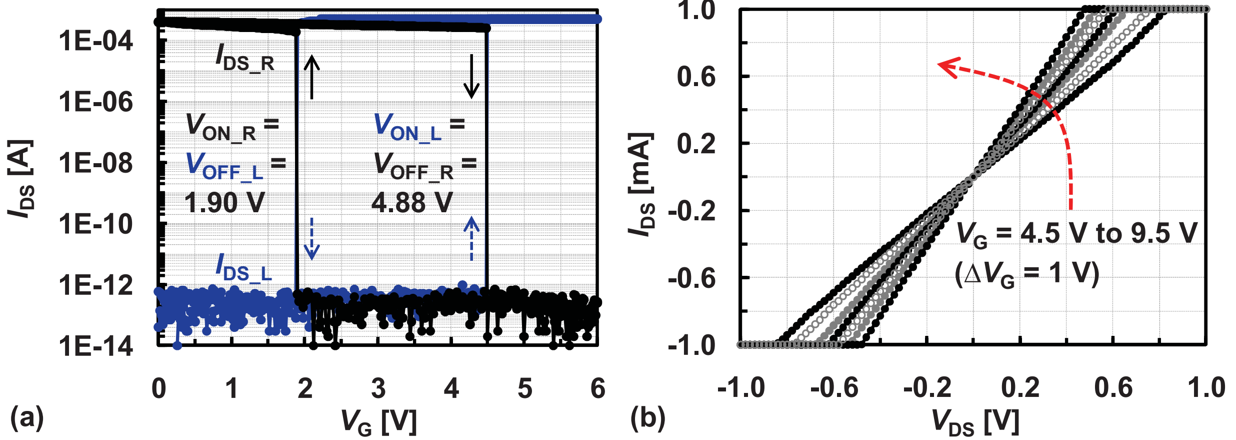


Figure 4.16: (a) Measured I_{DS}-V_{GS} characteristic. V_{BL} = 0 V and V_{BR} = V_{DD} = 6 V. V_{ON_R} and V_{ON_L} are slightly asymmetric about V_{DD} / 2, due to the negative strain gradient within poly-$Si_{0.4}Ge_{0.6}$ that makes T_{ACT} on either side different. (b) Measured I_{DS}-V_{DS} characteristic for the relay in (a). The *on*-state resistance decreases with increasing gate voltage (2.2 kΩ at V_G = 9.5 V).

not sufficiently low enough to yield low operating voltages and high production yield. Therefore, to reduce the impact of negative strain gradient within the poly-$Si_{0.4}Ge_{0.6}$ layer, a KrF excimer laser (λ = 248 nm) pulse (200 mJ/cm^2, 10 ns per shot, 10 shots) was applied to the gate electrode through a SiO_2 hardmask layer.

4.4.2. Testing Results

Fabricated see-saw relays were tested under N_2 purge at room temperature and atmospheric pressure. Unless otherwise specified, the drain-to-source bias voltage (V_{DS}) was constrained to be no greater than 2 V, in order to avoid reliability issues due to possible micro-welding between contacting electrodes. Figure 4.16(a) represents measured I_{DS}-V_{GS} characteristic of the corrugated see-saw relay. Note that the right and left body bias voltages were set to V_{BL} = 0 V and V_{BR} = V_{DD} = 6 V, respectively. As the gate voltage (V_G) is increased from 0 to V_{DD}, V_G needed to turn *on* the left side (V_{ON_L}) turns *off* the right side abruptly at approximately 4.88 V. Likewise, as V_G is reduced back to 0 V, the right side turns *on* abruptly and the left side turns *off* abruptly at about 1.9 V -- thus, the voltage needed to turn *on* the right side (V_{ON_R}) is $V_{DD} - V_G$ = 6 V – 1.9 V = 4.1 V. Notice that V_{ON_R} and V_{ON_L} are slightly asymmetric about V_{DD} / 2, which yields switching hysteresis (difference between V_{ON_L} and V_{ON_R}) of about 0.78 V. This hysteresis is caused by different T_{ACT} on the left and right sides of the relay. This indicates that the impact of laser-annealing was not uniform across the device. The turn-*on* voltages can be lowered by scaling down the as-fabricated T_{ACT} and $T_{CONTACT}$ and/or by further design optimization (for example, reducing the torsional spring constant of the see-saw relay).

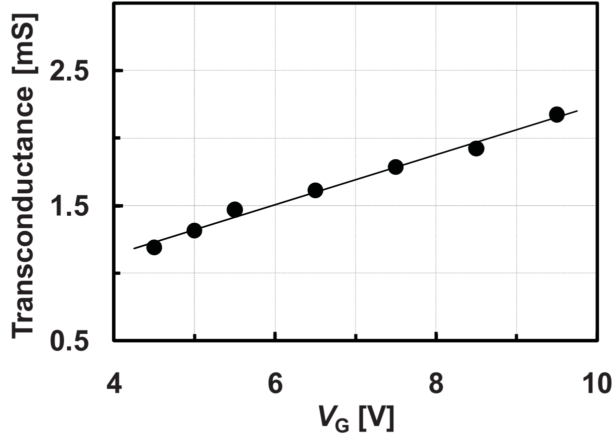


Figure 4.17: Measured values of g_m for the relay in Figure 4.16(a): g_m increases only slightly with increasing V_G.

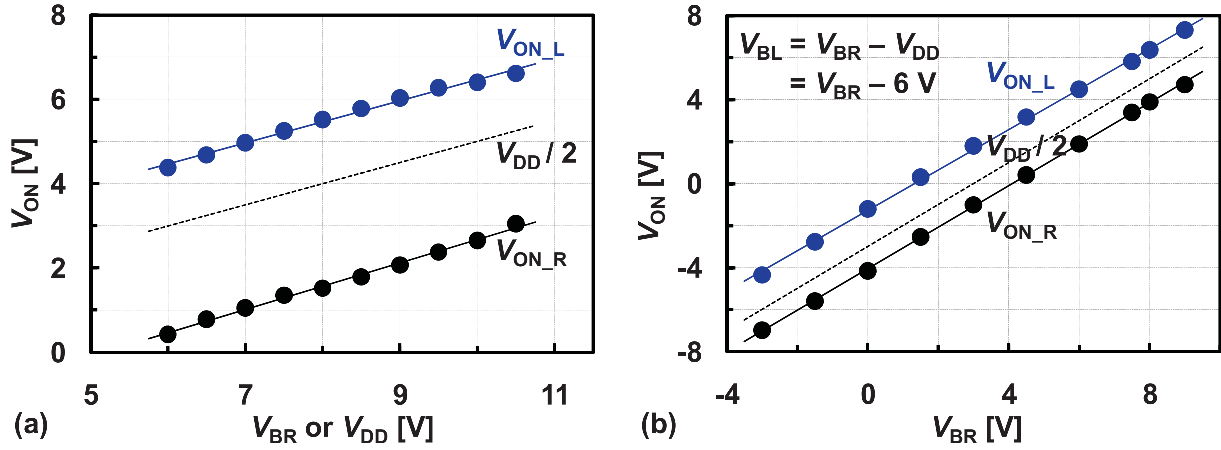


Figure 4.18: (a) Measured V_{ON_L} and V_{ON_R} with decreasing V_{BR} for a different see-saw relay. The switching voltages are reduced with decreasing $V_{BR} = V_{DD}$. (b) Measured values of V_{ON_L} and V_{ON_R} with change in body-bias for the relay shown in Figure 4.16(a). For each pair of V_{ON_L} and V_{ON_R}, V_G was swept from V_{BL} to V_{BR}. The difference between V_{BR} and V_{BL} was kept constant at $V_{DD} = 6$ V, *e,g,*, V_G was swept from –3 V to 3 V with $V_{BL} = -3$ V and $V_{BR} = 3$ V.

As V_G is increased beyond V_{ON_L} of the see-saw relay, the *on*-state resistance of the W to W contact decreases from ~4.0 kΩ to ~2.2 kΩ, as shown in Figure 4.16(b). This is because the effective contact area (A_R) increases with increasing F_{ELEC}. Note that $R_{CONTACT}$ at each side of the channel on either side is dependent upon the conditions under which the contact is formed as well as the properties of the contacting electrode material [9], [50]:

$$R_{CONTACT} = \frac{4 \cdot \rho \cdot \lambda}{3A_R} = \frac{4 \cdot \rho \cdot \lambda \cdot H \cdot \xi}{3F_{ELEC}(T_{ACT})} ,$$

where ρ and λ are the resistivity and the mean free path of electrons of the contacting electrode material, respectively, A_R is the effective contact area, H is the hardness of the contacting electrode material, and ξ is the deformation constant.

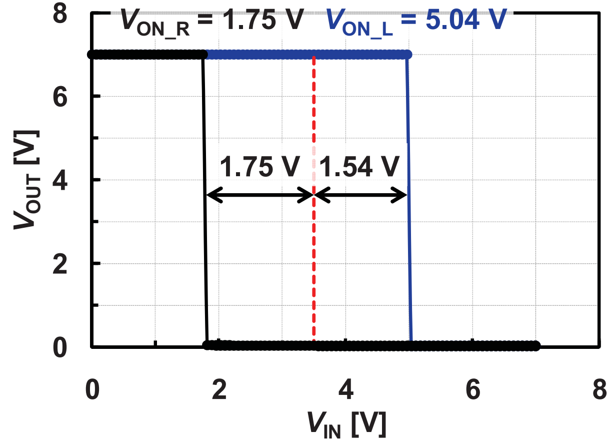


Figure 4.19: Measured VTC of a different see-saw relay. $V_{BR} = V_{DD} = 7$ V and $V_{BL} = 0$ V. The measurement setup was shown in Figure 4.3. For the equal input and output voltage swings, $V_{SR} = V_{DD} = 7$ V and $V_{SL} = 0$ V.

Figure 4.17 shows that measured values of g_m for the corrugated see-saw relay are relatively low (less than about 0.4 mS/μm of channel width) to be used for small-signal current amplification. The values of g_m for a small-signal MOSFET are typically in the range between 10 and 30 mS/μm. For the 65 nm technology node ($I_D \approx 1$ mA/μm), a g_m of 11 to 32 mS/μm is achieved with the overdrive voltage in the range between 70 to 200 mV [51].

Figures 4.18(a) and 4.18(b) show that the turn-*on* voltages of the see-saw relay can be adjusted via body-bias. In Figure 4.18(a), both V_{BR} and V_{DD} were decreased while V_{BL} was kept at 0 V, whereas in Figure 4.18(b), V_{BR} and V_{BL} were reduced by the same amount while V_{DD} was kept at $V_{BR} - V_{BL} = 6$ V. The asymmetry of switching about V_{DD} / 2 in both figures is caused by different T_{ACT} on the left and right sides of the relay (due to the effect of negative strain gradient).

Figures 4.19 shows measured VTC of a different corrugated see-saw relay. This relay shows complementary switching behavior that is almost perfectly symmetric about V_{DD} / 2. The slight asymmetry of switching indicates that the strain gradient effect was not completely removed.

4.5. Summary

A see-saw relay design is proposed and demonstrated to provide for perfectly complementary switching behavior that is symmetric about V_{DD} / 2. The main advantages of this design are that it can maximize the operating voltage margin and minimize the crowbar current and that by allowing active turn-*off*, it can also minimize the asymmetric switching behavior and improve the reliability of a relay-based digital logic technology. The complementary see-saw relay design is versatile for mechanically implementing various logic functions and a latch function with a single structure. It is a viable option for the implementation of compact, relay-based ultra-low-power integrated circuits in the future.

4.6. References

[1] K. Akarvardar, D. Elata, R. Parsa, G. C. Wan, K. Yoo, J. Provine, P. Peumans, R. T. Howe, and H.-S. P. Wong, "Design Considerations for Complementary Nanoelectromechanical Logic Gates," in *IEEE International Electron Device Meeting (IEDM '07) Technical Digest*, pp. 299-302, December 2007.

[2] F. Chen, H. Kam, D. Marković, T.-J. King Liu, V. Stojanović, and E. Alon, "Integrated Circuit Design with NEM Relays," in *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design (ICCAD '08)*, pp. 750-757, November 2008.

[3] R. H. Dennard, J. Cai, and A. Kumar, "A Perspective on Today's Scaling Challenges and Possible Future Directions," *Elsevier Solid-State Electronics*, vol. 51, no. 4, pp. 518-525, April 2007.

[4] H. Kam, V. Pott, R. Nathanael, J. Jeon, E. Alon, and T.-J. King Liu, "Design and Reliability of a Micro-Relay Technology for Zero-Standby-Power Digital Logic Applications," in *IEEE International Electron Device Meeting (IEDM '09) Technical Digest*, pp. 809-812, December 2009.

[5] R. Nathanael, V. Pott, H. Kam, J. Jeon, and T.-J. King Liu, "4-Terminal Relay Technology for Complementary Logic," in *IEEE International Electron Device Meeting (IEDM '09) Technical Digest*, pp.223-226, December 2009.

[6] R. Nathanael, V. Pott, H. Kam, J. Jeon, E. Alon, and T.-J. King Liu, "Four-Terminal Relay Body-Biasing Schemes for Complementary Digital Circuits," *IEEE Electron Device Letters*, vol. 31, no. 8, pp. 890-892, August 2010.

[7] F. Chen, M. Spencer, R. Nathanael, C. Wang, H. Fariborzi, A. Gupta, H. Kam, V. Pott, J. Jeon, T.-J. King Liu, D. Marković, V. Stojanović, and E. Alon, "Demonstration of Integrated Micro-Electro-Mechanical (MEM) Switch Circuits for VLSI Applications," in *IEEE International Solid-State Circuits Conference (ISSCC '10) Tech. Dig.*, pp. 150-151, February 2011.

[8] H. Fariborzi, M. Spencer, V. Karkare, J. Jeon, R. Nathanael, C. Wang, F. Chen, H. Kam, V. Pott, E. Alon, T.-J. King Liu, V. Stojanović, and D. Marković, "Analysis and Demonstration of MEM-Relay Power Gating," *IEEE Custom Integrated Circuits Conference (CICC '10)*, September 2010.

[9] M. Spencer, F. Chen, C. Wang, F. Chen, R. Nathanael, H. Fariborzi, H. Kam, V. Pott, J. Jeon, T.-J. King Liu, D. Marković, E. Alon, and V. Stojanović, "Demonstration of Integrated Micro-Electro-Mechanical Switch Circuits for VLSI Applications," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 308-320, January 2011.

[10] E. J. J. Kruglick, "Micro-Relay Design, Performance, and Systems," *Ph.D. Dissertation*, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Spring 1999.

[11] A. Hirata, K. Machida, H. Kyuragi, and M. Maeda, "An Electrostatic Micromechanical Switch for Logic Operation in Multichip Modules on Si," *Elsevier Sensors and Actuators A: Physical*, vol. 80, no. 2, pp. 119-125, February 2000.

[12] C.-Y. Tsai, W.-T. Kuo, C.-B. Lin, and T.-L. Chen, "Design and Fabrication of MEMS Logic Gates," *IOP Journal of Micromechanics and Microengineering*, vol. 18, no.4, pp. 045001.1-045001.10, April 2008.

[13] G. M. Rebeiz, *RF MEMS Theory, Design, and Technology*, Hoboken, NJ: John-Wiley & Sons, Inc., 2003.

[14] O. Degani, E. Socher, A. Lipson, T. Leitner, D. J. Setter, S. Kaldor, and Y. Nemirovsky, "Pull-In Study of an Electrostatic Torsion Microactuator," *IEEE Journal of Microelectromechanical Systems*, vol. 7, no. 4, pp. 373-379, April 1998.

[15] K. E. Peterson, "Silicon Scanning Torsional Mirror," *IBM Journal of Research and Development*, vol. 24, no. 5, pp. 631-637, September 1980.

[16] P. F. V. Kessel, L. J. Hornbeck, R. E. Meier, and M. R. Douglass, "A MEMS-Based Projection Display," *Proceedings of the IEEE*, vol. 86, no. 8, pp. 1687-1704, August 1998.

[17] M. R. Douglas, "Lifetime Estimates and Unique Failure Mechanisms of the Digital Micromirror Device (DMD)," in *Proceedings of the IEEE Annual International Reliability Physics Symposium (IRPS '98)*, pp. 9-16, March 1998.

[18] A. B. Sontheimer, "Digital Micromirror Device (DMD) Hinge Memory Lifetime Reliability Modeling," in *Proceedings of the IEEE Annual International Reliability Physics Symposium (IRPS '02)*, pp. 118-121, April 2002.

[19] J. Kim, S. Kwon, H. Jeong, Y. Hong, S. Lee, I. Song, and B. Ju, "A Stiff and Flat Membrane Operated DC Contact Type RF MEMS Switch with Low Actuation Voltage," *Elsevier Sensors and Actuators A: Physical*, vol. 153, pp. 114-119, April 2009.

[20] D. Hah, E. Yoon, and S. Hong, "A Low-Voltage Actuated Micromachined Microwave Switch Using Torsion Springs and Leverage," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, no. 12, pp. 2540-2545, December 2000.

[21] M. Erdamar, K. Shanmugasundaram, P. Roman, P. Mumbauer, M. Klimkiewicz, and J. Ruzyllo, “Deep Lateral Anhydrous HF/Methanol Etching for MEMS Release Processes,” *SPIE Journal of Micro/Nano Lithography, MEMS, and MOEMS*, vol. 7, no. 3, pp. 033014, March, 2008.

[22] W. I. Jang, C. A. Choi, M. L. Lee, C. H. Jun, and Y. T. Kim, “Fabrication of MEMS Devices Using Anhydrous HF Gas-Phase Etching with Alcoholic Vapor,” *IOPscience Journal of Micromachining and Microengineering*, vol. 12, no. 3, pp. 297-306, March, 2002.

[23] G. C. Wetsel, Jr. and K. J. Strozewski, “Dynamic Model of Microscale Electromechanical Spatial Light Modulator,” *AIP Journal of Applied Physics*, vol. 73, no. 11, pp. 7120-7124, November 1993.

[24] K. S. Lebouitz, A. Mazaheri, R. T. Howe, and A. P. Pisano, “Vacuum Encapsulation of Resonant Devices Using Permeable Poly-Silicon,” in *Proceedings of the IEEE 12th International Conference of Microelectromechanical Systems (MEMS '99)*, pp. 470-475, January 1999.

[25] P.Y. Papalambros and D. J. Wilde, *Principles of Optimal Design: Modeling and Computation 2nd Edition*, New York, NY: Cambridge University Press, 2000.

[26] L. Lin and A. Pisano, *ME C219 Lecture Notes: Parametric and Optimal Design of MEMS*, Department of Mechanical Engineering, University of California at Berkeley, February, 2009.

[27] D. Lee, V. Pott, H. Kam, R. Nathanael, and T.-J. King Liu, “AFM CHacracterization of adhesion force in micro-relays,” in *Proceedings of the IEEE 23rd International Conference of Microelectromechanical Systems (MEMS '10)*, pp. 232-235, January 2010.

[28] T.-J. King Liu, J. Jeon, R. Nathanael, H. Kam, V. Pott, and E. Alon, “Prospects for MEM Logic Switch Technology,” in *IEEE International Electron Device Meeting (IEDM '10) Technical Digest*, pp. 424-427, December 2010.

[29] E. P. Popov, *Engineering Mechanics of Solids 2nd Edition*, Upper Saddle River, NJ: Pearson Prentice Hall, 1998.

[30] G. K. Fedder, “Simulation of Microelectromechanical Systems,” *Ph.D. Dissertation*, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, September 1994.

[31] *MATLAB® R2011a User's Guide [Online] Available: http://www.mathworks.com/help/toolbox/optim/ug/fmincon.html*, The MathWorks, Inc., Natick, MA, 2010.

[32] *CoventorWare^{TM}2008 Designer Reference Manual*, Coventor, Inc., Cary, NC, 2008.

[33] *CoventorWare^{TM}2008 Analyzer Reference Manual*, Coventor, Inc., Cary, NC, 2008.

[34] M. J. Madou, *Fundamentals of Microfabrication 2nd Edition*, Boca Raton, FL: CRC Press, 2002.

[35] C. W.-Z. Low, "Novel Processes for Modular Integration of Silicon-Germanium MEMS with CMOS Electronics," *Ph.D. Dissertation*, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, February 2007.

[36] J. Tregilgas, "How We Developed an Amorphous Hinge Material," *ASM International Advanced Materials and Processes Magazine*, pp. 46-49, January 2005.

[37] C. W.-Z. Low, T.-J. King Liu, and R. T. Howe, "Characterization of Polycrystalline Silicon-Germanium Film Deposition for Modularly Integrated MEMS Applications," *IEEE Journal of Microelectromechanical Systems*, vol. 16, no. 1, pp. 68-77, January 2007.

[38] S. A. Bhave, B. L. Bircumshaw, C. W.-Z. Low, Y.-S. Kim, A. P. Pisano, T.-J. King Liu, and R. T. Howe, "Poly-SiGe: A High-*Q* Structural Material for Integrated RF MEMS," in *Proceedings of the IEEE Solid-State Sensor, Actuator, and Microsystems Workshop (Hilton Head '02)*, pp. 34-37, June 2002.

[39] S. Sedky, R. T. Howe, and T.-J. King Liu, "Pulsed-Laser Annealing, a Low-Thermal-Budget Technique for Eliminating Stress Gradient in poly-SiGe MEMS Structures," *IEEE Journal of Microelectromechanical Systems*, vol. 13, no. 4, pp. 669-675, August 2004.

[40] J. El-Rifai, A. Witvrouw, A. A. Aziz1, R. Puers, C. V. Hoof, and S. Sedky, "Selective Laser Annealing for Improved SiGe MEMS Structural Layers at 210 ºC," in *Proceedings of the IEEE 23rd International Conference of Microelectromechanical Systems (MEMS '10)*, pp. 324-327, January 2010.

[41] W. C. Tang, C. T.-C. Nguyen, and R. T. Howe, "Laterally Driven Poly-Silicon Resonant Microstructures," in *Proceedings of the IEEE 2nd International Conference of Microelectromechanical Systems (MEMS '89)*, pp. 53-59, February 1989.

[42] R. T. Howe, "Recent Advances in Surface Micromachining," *Technical Digest of the 13th Sensor Symposium*, pp. 1-8, June 1993.

[43] J. Brugger, G. Beljakovic, M. Despont, N. F. de Rooij, and P. Vettiger, "Silicon Micro/Nanomechanical Device Fabrication Based on Focused Ion Beam Surface Modification and KOH Etching," *Elsevier Microelectronic Engineering*, vol. 35, no. 1-4, pp. 401-404, February 1997.

[44] F. Ke, J. Miao, and J. Oberhammer, “A Ruthenium-Based Multimetal-Contact RF MEMS Switch with a Corrugated Diaphragm,” *IEEE Journal of Microelectromechanical Systems*, vol. 17, no. 6, pp. 1447-1459, December 2008.

[45] V. L. Spiering, S. Bouwstra, R. M. E. J. Spiering, and M. Elwenspoek, “On-Chip Decoupling Zone for Package-Stress Reduction,” in *Proceedings of the IEEE International Conference on Solid-State Sensors and Actuators (Transducers '91)*, pp. 982-985, June 1991.

[46] J. H. Jerman, “The Fabrication and Use of Micromachined Corrugated Silicon Diaphragm,” *Elsevier Sensors and Actuators A: Physical*, vol. 23, no. 1-3, pp. 982-992, April 1990.

[47] M. A. F. van den Boogaart, M. Lishchynska, L. M. Doeswijk, J. C. Greer, and J. Brugger, “Corrugated Membranes for Improved Pattern Definition with Micro/Nanostencil Lithography,” *Elsevier Sensors and Actuators A: Physical*, vol. 130-131, pp. 568-574, August 2006.

[48] P. D. Moor, J. John, S. Sedky, and C. V. Hoof, “Linear Arrays of Fast Uncooled Poly-SiGe Microbolometers for IR Detection,” *Proceedings of SPIE*, vol. 4028, pp. 27-34, April 2000.

[49] W.C. Young, *Roark's formulas for stress and strain*, New York, NY: McGraw-Hill, 1989.

[50] R. Holm, *Electric Contacts: Theory and Applications*, Berlin, NY: Springer-Verlag, 1967.

[51] R. J. Baker, *CMOS: Circuit Design, Layout, and Simulation 3rd Edition*, Hoboken, NJ: John-Wiley & Sons, Inc., 2003.

Chapter 5

Multi-Electrode Relays for Compact Implementation of Digital Integrated Circuits

Abstract: In this chapter, multi-gate, multi-source/drain relays are proposed and demonstrated to enable compact implementation of zero-leakage digital electronics at no incremental process cost. Basic multi-input logic functions (*e.g.* **AND**, **NAND**, **OR**, **MAJORITY**, and **WEIGHTED**) can be implemented by tuning the body bias voltage of a single multi-input relay. The gate electrodes of the multi-input relay can be properly sized to decode an analog input signal level, so that flash-type analog-to-digital-converters can be achieved via multi-gate operation. Performance improvement with scaling is also discussed.

5.1. Introduction

Nano-electro-mechanical relay technology has been proposed as a potential solution to overcome the fundamental energy-efficiency limit of Complementary-Metal-Oxide-Semiconductor (CMOS) technology, because it offers the ideal characteristics of zero *off*-state leakage current and abrupt *on/off* switching behavior which provide for zero static power consumption and aggressive supply voltage (V_{DD}) scaling [1]. Various digital integrated-circuit (IC) building blocks (logic, memory, and clocking structures) implemented purely with micro-electro-mechanical relays have been demonstrated recently [2], [3]. New relay designs that incorporate multiple input (gate) electrodes and multiple pairs of output (source/drain) electrodes would enable more compact implementation of digital ICs (Figure 5.1) at no incremental process cost. In this chapter, the first functional prototype multi-electrode relays are presented and performance improvement with scaling is discussed.

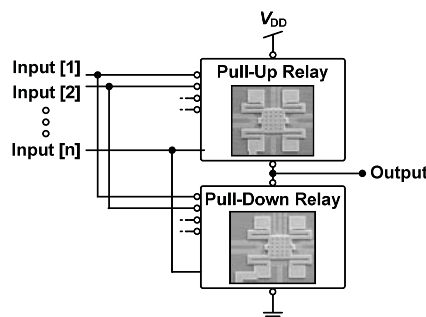
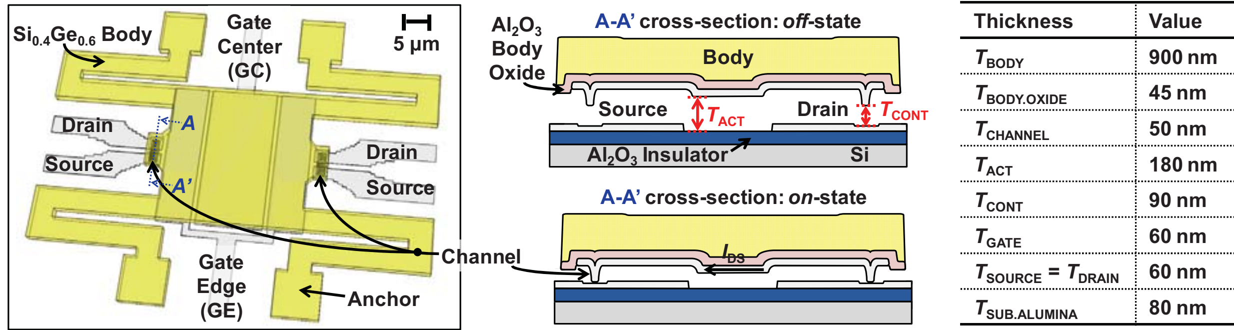


Figure 5.1: A generic multi-input relay combinational logic circuit.



Thickness	Value
T_{BODY}	900 nm
$T_{BODY.OXIDE}$	45 nm
$T_{CHANNEL}$	50 nm
T_{ACT}	180 nm
T_{CONT}	90 nm
T_{GATE}	60 nm
$T_{SOURCE} = T_{DRAIN}$	60 nm
$T_{SUB.ALUMINA}$	80 nm

Figure 5.2: Perspective view of a 2-input, dual-source/drain relay structure; In the *off*-state, an air gap prevents DC current to flow between the source and drain electrodes on each side. In the *on*-state, electrostatic force between the gate(s) and the body brings both channels into contact with the pairs of source/drain electrodes. Both gate electrodes are equally sized (472 μm^2 each).

5.2. 2-Input, Dual-Source/Drain Relay

Figure 5.2 illustrates the structure and operation of a 2-input, dual-source/drain relay: In the *off*-state, immeasurably low leakage current flows because an air gap (of thickness T_{CONT}) separates the channels from the pairs of source/drain electrodes on each side. As voltage applied to the gate(s) increase(s) beyond a threshold "pull-in" voltage (V_{PI}), the relay turns *on* abruptly because the attractive electrostatic force between the gate electrode(s) and the movable body electrode brings the channels (attached underneath the body electrode via an insulating oxide layer) into contact with the source/drain electrodes. A CMOS-compatible (peak substrate temperature < 410 °C) surface-micromachining process (Figure 5.3) was used to fabricate all of the multi-gate, dual-source/drain relays in this work. Scanning electron microscopy (SEM) images of a double-gate, dual-source/drain relay are shown in Figure 5.4.

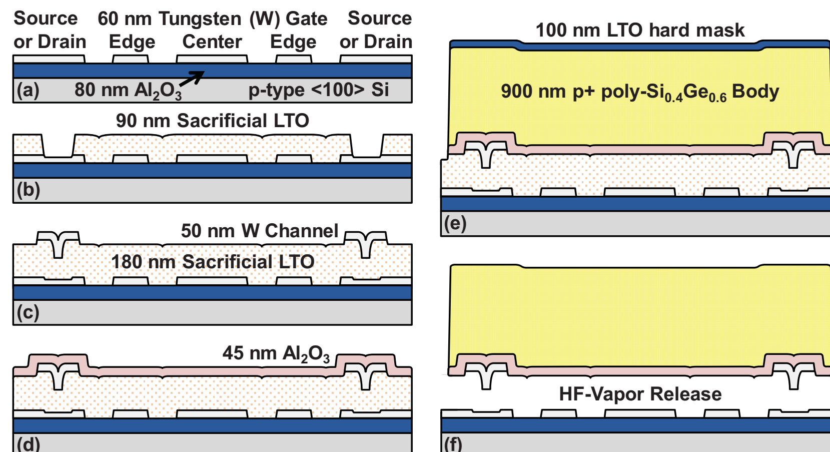


Figure 5.3: Low-thermal-budget process used to fabricate multi-input, dual-source/drain relays: **(a)** Tungsten (W) gate, source, and drain electrodes defined on the Al_2O_3-coated substrate. **(b)** 1st sacrificial Low-Temperature-Oxide (LTO) layer deposition at 400 °C, followed by contact region definition. **(c)** 2nd sacrificial LTO layer deposition, followed by W channel formation. **(d)** Al_2O_3 insulator ALD at 300 °C. **(e)** p+ poly-$Si_{0.4}Ge_{0.6}$ body layer deposition at 410 °C, followed by LTO deposition and body patterning. **(f)** Release etch in vapor-HF at 50 °C to form an air gap.

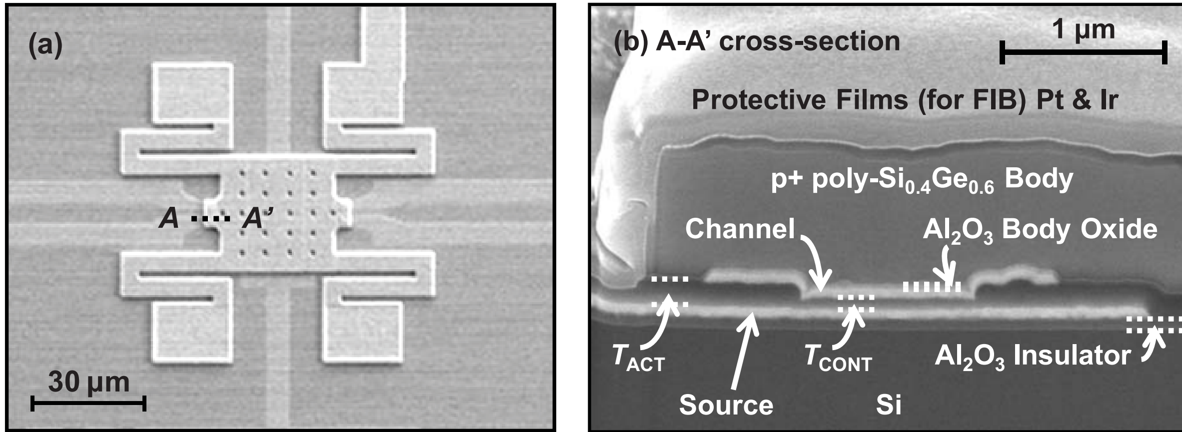


Figure 5.4: (a) Plan view scanning electron micrograph (SEM) of the 2-input relay. **(b)** Cross-sectional SEM (X-SEM) from focused ion beam (FIB) cut through A-A' in (a).

Figure 5.5 shows measured V_{PI} values for a relay with 2 equally-sized gate electrodes (472 μm^2). Due to body curvature resulting from undesirable strain gradient within the structural layer, the actuation gap (T_{ACT}) is larger at the center than at the edge of the body so that V_{PI} for the center gate electrode (input combination [1, 0], where '1' ≡ V_G) is higher than that for the edge gate electrode (input combination [0, 1]). The measured I_{DS}-V_G characteristics for double-gate operation, *i.e.* with the 2 gates tied together (input combination [1, 1]), are shown in Figure 5.6. The gate switching voltages can be tuned by biasing the body electrode, as shown in Figure 5.7. V_{PI} for single-gate operation (with the other gate grounded) is more sensitive to body bias because of the additional electrostatic force between the inactive (grounded) gate and the body.

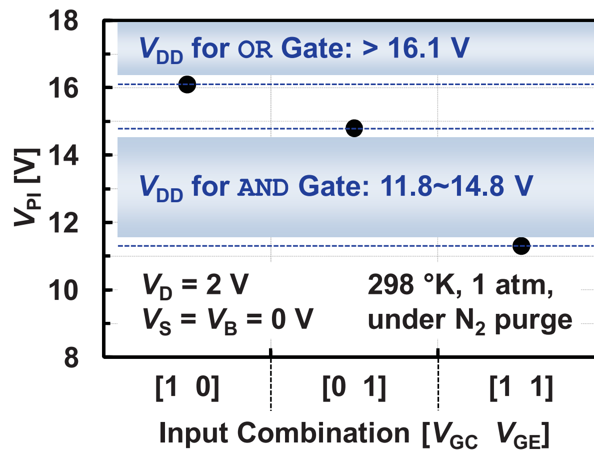


Figure 5.5: Measured pull-in voltages (V_{PI}) of the 2-input relay. '1' ≡ V_G.

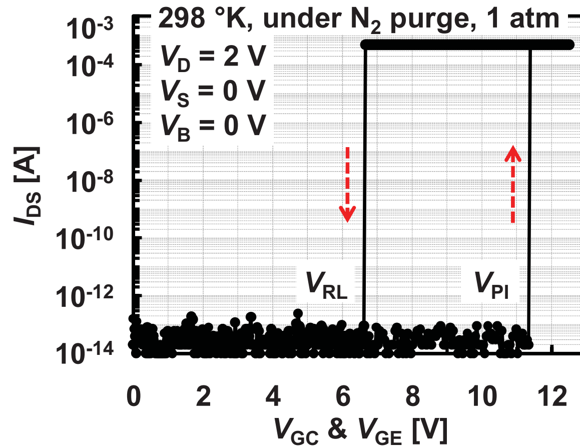


Figure 5.6: Measured I_{DS}-V_G characteristic; Both gate electrodes were swept to 12.5 V.

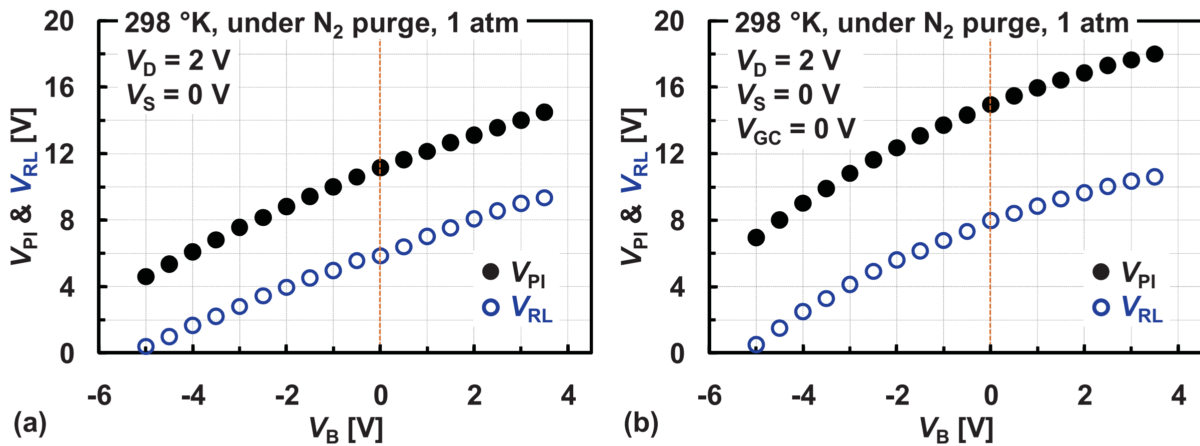


Figure 5.7: Measured V_{PI} and V_{RL} with changes in body-bias voltages. **(a)** Both gate electrodes were swept. **(b)** The gate edge electrode (V_{GE}) was swept.

Figure 5.8 demonstrates that a single 2-input relay (with drain electrodes connected to a power supply and source electrodes connected to an oscilloscope with 1 MΩ internal resistance) can perform either the **AND** function or the **NAND** function, depending on the body bias voltage. If the input voltage level is increased sufficiently (reference Figure 5.5), a 2-input relay performs the OR function. This is demonstrated in Figure 5.9, which also shows that the output voltage range is solely determined by the drain bias voltage and hence can be different for each set of source/drain electrodes. This feature can be leveraged for voltage level conversion applications.

Figure 5.10 shows measured 2-input relay turn-*on* delay (t_{DELAY}) as a function of the driving gate voltage, for single-gate and double-gate operation. Double-gate operation is faster due to larger electrostatic force between the gates and the body. In both cases, t_{DELAY} is reduced with increasing gate overdrive. A negative body bias also improves t_{DELAY} because it serves to reduce T_{ACT} and T_{CONT}, resulting in larger electrostatic force and shorter distance to switch *on*. Figure

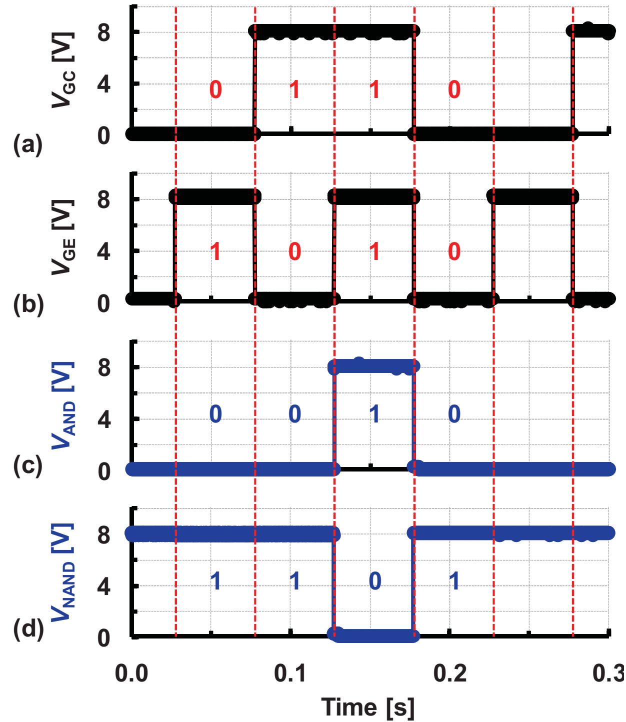


Figure 5.8: Measured chronogram for 2-input **AND** and **NAND** gates illustrating full-rail swing at the output. The drain electrodes were connected to the power supply (V_{DD} = 8 V), and the source electrodes were tied together to form the logic gate output node. **(a)** **INPUT 1** signal. **(b)** **INPUT 2** signal. **(c)** **AND** output signal. V_B = -4 V. **(d)** **NAND** output signal. V_B = 12 V.

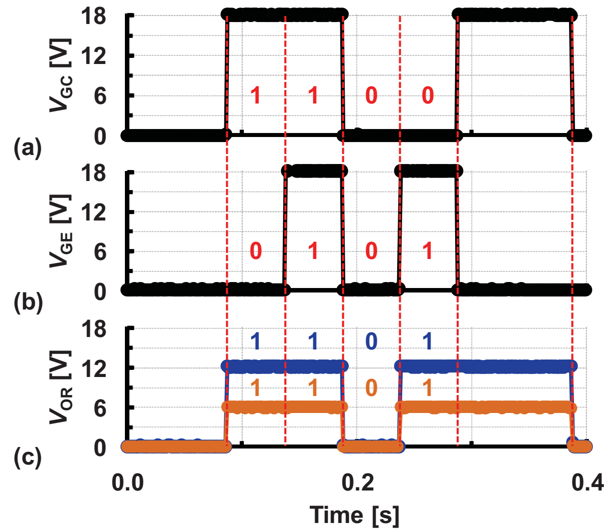


Figure 5.9: Measured chronogram for a 2-input **OR** gate showing voltage level conversion. The left and right drain electrodes were at 12 V and 6 V, respectively. Each source electrode serves as the logic gate output node. V_B = 0 V. **(a)** **INPUT 1** signal. **(b)** **INPUT 2** signal. **(c)** **OR** output signals at 12 V (left source) and 6 V (right source).

5.11 shows that *on*-state resistance (R_{ON}) is fairly stable, staying below 10 kΩ after 3×10^7 *on/off* cycles, which is adequate for digital logic applications [3].

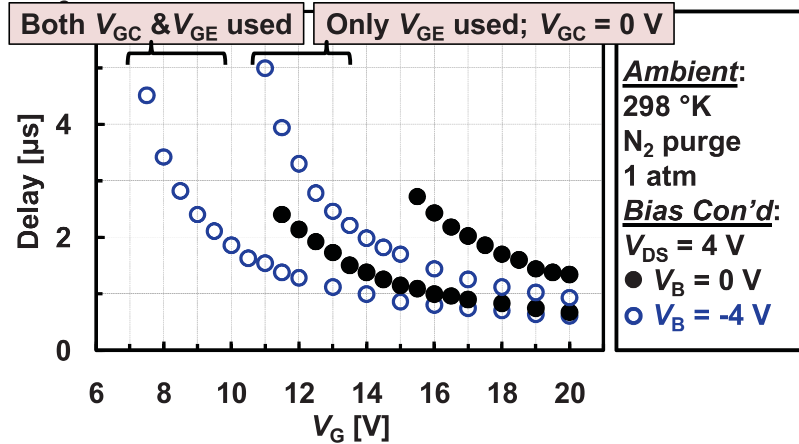


Figure 5.10: (a) Plan view scanning electron micrograph (SEM) of the 2-input relay. **(b)** Cross-sectional SEM (X-SEM) from focused ion beam (FIB) cut through A-A' in (a).

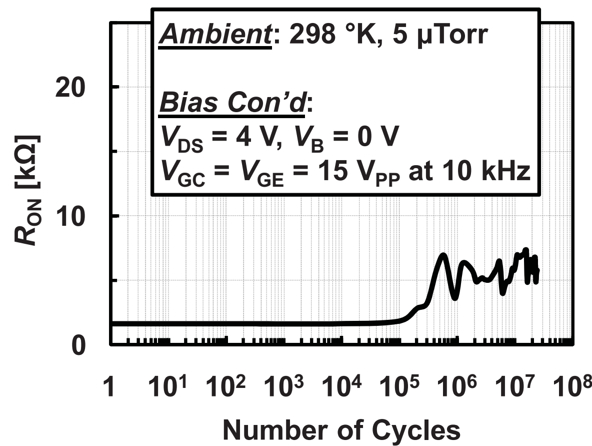


Figure 5.11: Measured 2-input, dual-source/drain relay endurance.

5.3. 3-Input Relay

By partitioning the gate layer into 3 electrodes (Figure 5.12), various 3-input logic functions are realized by adjusting the input voltage level, as shown in Figure 5.13. (Again, V_{PI} for each gate electrode is different due to body curvature resulting from an undesirable strain gradient within the structural layer.) Measured I_{DS}-V_G characteristics for the outer two gates (Y and Z) tied together (input combination [0, 1, 1]) are shown in Figure 5.14. The additional input electrode increases the versatility of the relay: basic logic functions, such as **AND**, **MAJORITY** (***X•Y+Y•Z+X•Z***), and **OR**, are achieved by adjusting the input voltage level, as demonstrated in Figure 5.15. Similarly as for the 2-input relay, t_{DELAY} can be improved by body-biasing and/or increasing gate overdrive (Figure 5.16).

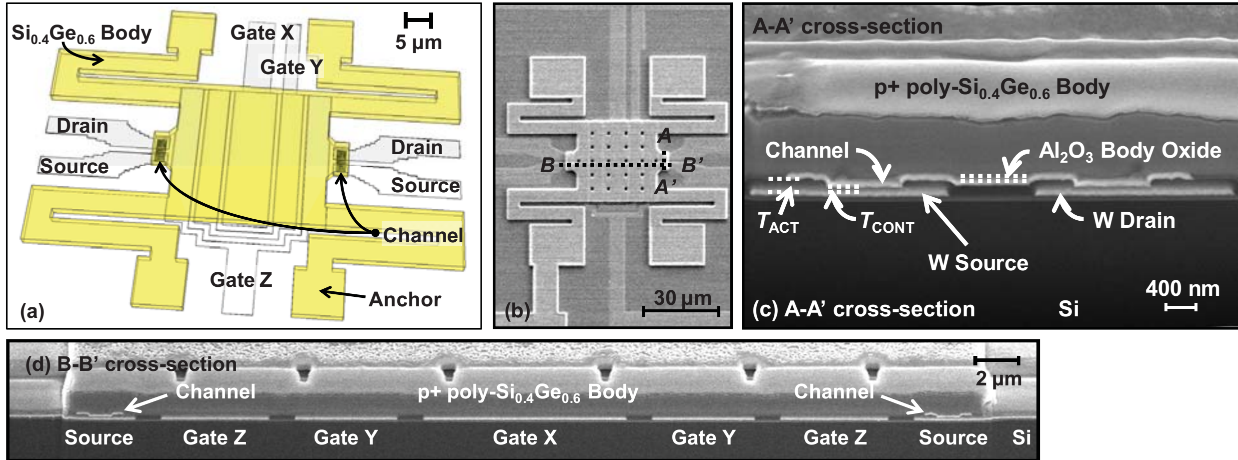


Figure 5.12: (a) Three-dimensional (3-D) schematic of a 3-input relay structure. The vertical dimensions are as shown in **Figure 5.3**. All gate electrodes are equally sized (288 μm^2 each). **(b)** Plan-view SEM of the 3-input relay. **(c)** X-SEM from FIB cut through A-A' in (b). **(d)** X-SEM through B-B' in (b).

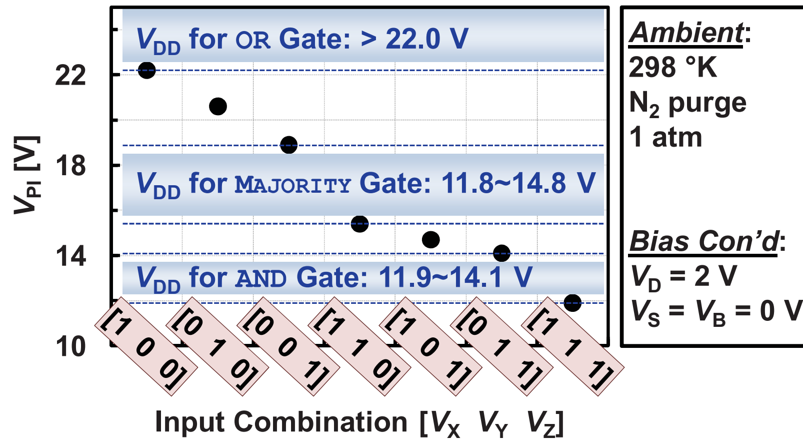


Figure 5.13: Measured V_{PI} of the 3-input relay. '1' ≡ V_G. The logic expression for a **MAJORITY** gate is ***X•Y+Y•Z+X•Z***.

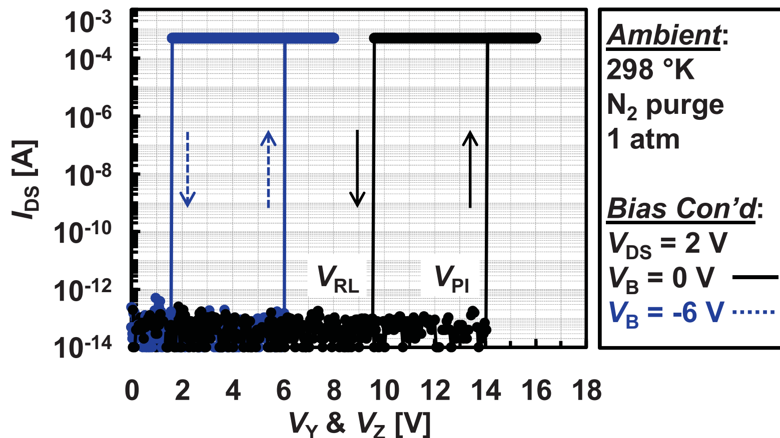


Figure 5.14: Measured I_{DS}-V_G characteristic of the 3-input relay with *vs.* without applying a body-bias voltage.

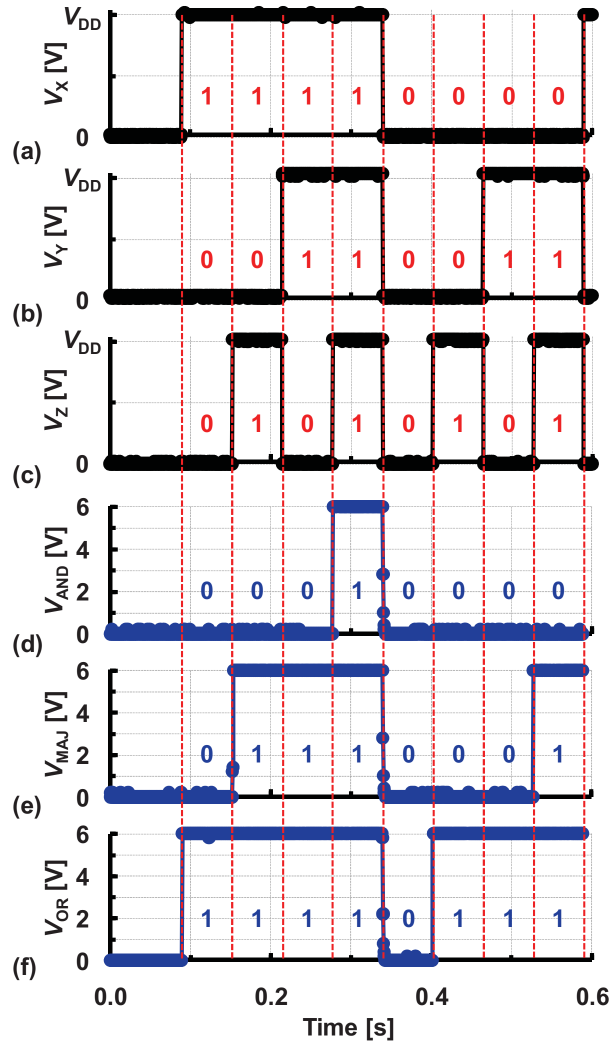


Figure 5.15: Measured chronogram for 3-input **AND**, **MAJORITY**, and **OR** gates. The drain electrodes were connected to the power supply (6 V), and the source electrodes were tied together to form the logic gate output node tied to **GND** via the 1 MΩ internal resistance of the oscilloscope. **(a)** **INPUT X** signal (to "Gate X"). **(b)** **INPUT Y** signal (to "Gate Y"). **(c)** **INPUT Z** signal (to "Gate Z"). **(d)** **AND** output signal. **(e)** **MAJORITY** gate (***X*•*Y*+*Y*•*Z*+*X*•*Z***) output signal. **(f)** **OR** output signal.

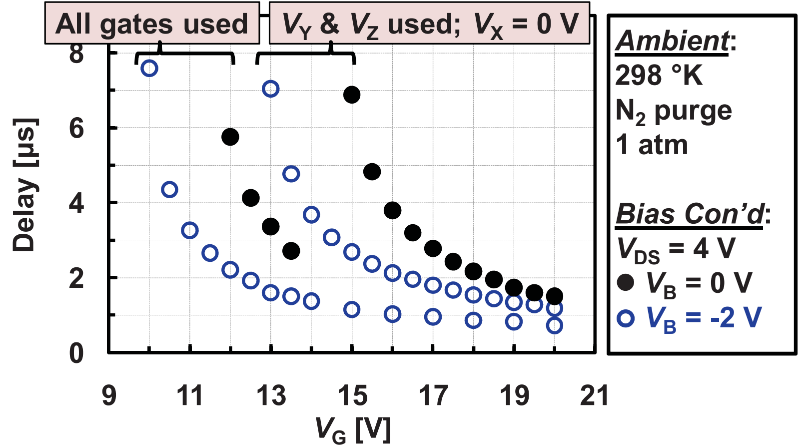


Figure 5.16: Measured turn-*on* delay of the 3-input relay with increasing gate overdrive.

	ΔV_{PI} = 0.9 V	ΔV_{PI} = 1.6 V
Electrode Area	Value	Value
$T_{GATE.X}$	192 μm²	96 μm²
$T_{GATE.Y}$	136 μm²	88 μm²
$T_{GATE.Z}$	536 μm²	680 μm²
T_{BODY}	1024 μm²	1024 μm²

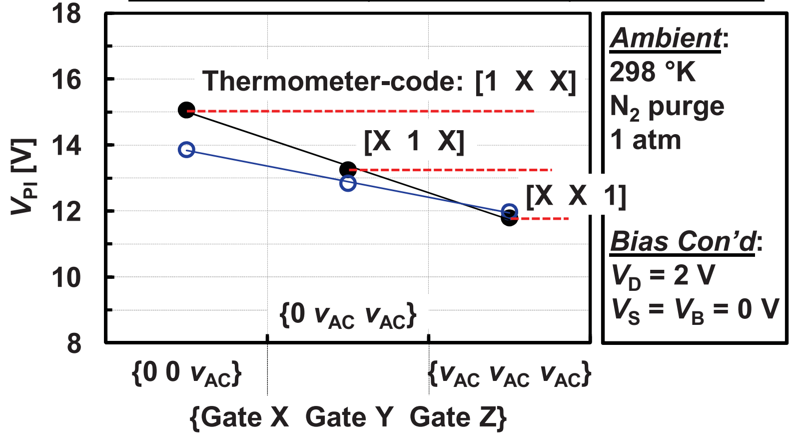


Figure 5.17: Measured, linearly-increasing V_{PI} of a 3-input relay with differently-sized gates for an analog-to-digital converter (ADC).

For use as a flash-type analog-to-digital converter (ADC), the switching voltages for single-gate (most significant bit, MSB), double-gate, and triple-gate (least significant bit, LSB) operation must decrease linearly to properly decode an analog input signal level (v_{AC}). This can be achieved by properly sizing the gate electrodes, as shown in Figure 5.17. As an example, if v_{AC} is between 11.7 V and 13.3 V, only the LSB relay (with v_{AC} driving all 3 input electrodes) turns *on*; the other two relays (with v_{AC} driving only the 2 smallest input electrodes, or with v_{AC} driving only the smallest input electrode) do not turn *on*, so that the thermometer code [001] is generated.

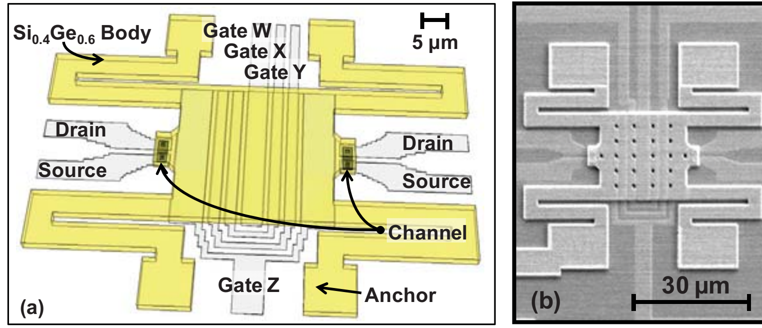


Figure 5.18: (a) 3-D view of a 4-input relay. The vertical dimensions are as shown in **Figure 5.3**. The "Gate Z" is weighted heavier (400 μm^2) than the other gate electrodes (128 μm^2 each). **(b)** SEM of the relay.

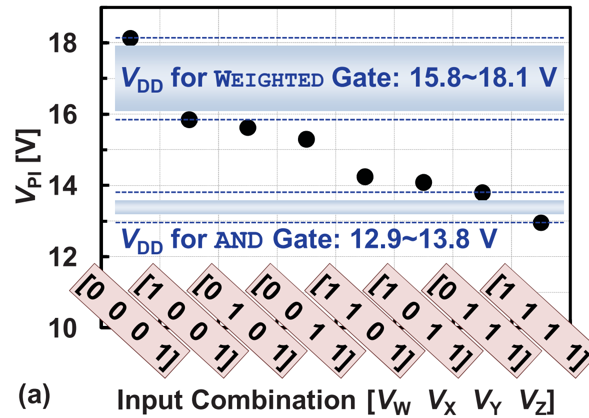


Figure 5.19: (a) Measured V_{PI} of the 4-input relay at 298 °K under N_2 purge at 1 atm. '1' ≡ V_G. V_{DS} = 2 V and V_B = 0 V. The **BOOLEAN** expression for a **WEIGHTED** gate is ***Z*•(*W*+*X*+*Y*)**.

5.4. 4-Input Relay

Figure 5.18 shows a 4-input relay structure used to demonstrate the possibility of performing complex logic functions by adjusting the areas of the input electrodes. The area of the outermost gate electrode (Gate Z) is designed to be larger than that of the other gate electrodes (Gates W, X, and Y) to realize the function ***Z*•(*W*+*X*+*Y*)**, *i.e.* the relay turns *on* only when the input signals to the larger electrode and at least one of the other gate electrodes are "high." Figure 5.19(a) shows the input voltage ranges required for the relay to perform the **AND** or **WEIGHTED** (***Z*•(*W*+*X*+*Y*)**) functions, and Figure 5.19(b) shows the measured I_{DS}-V_G characteristic for quadruple-gate operation, *i.e.* with the 4 gates tied together (input combination [1, 1, 1, 1]). Note that the switching hysteresis is relatively large, which can adversely affect the relay functionality (such that it does not turn *off* if an input signal changes from "high" to "low"). Recent improvements to the relay fabrication process and design mitigate this issue, as demonstrated by the measured I_{DS}-V_G characteristic in Figure 5.19(c) [4]. Similarly as for a 3-input relay, the gate electrodes of the 4-input can be sized to generate each bit of a 4-bit thermometer code for flash-type ADC applications (Figure 5.20).

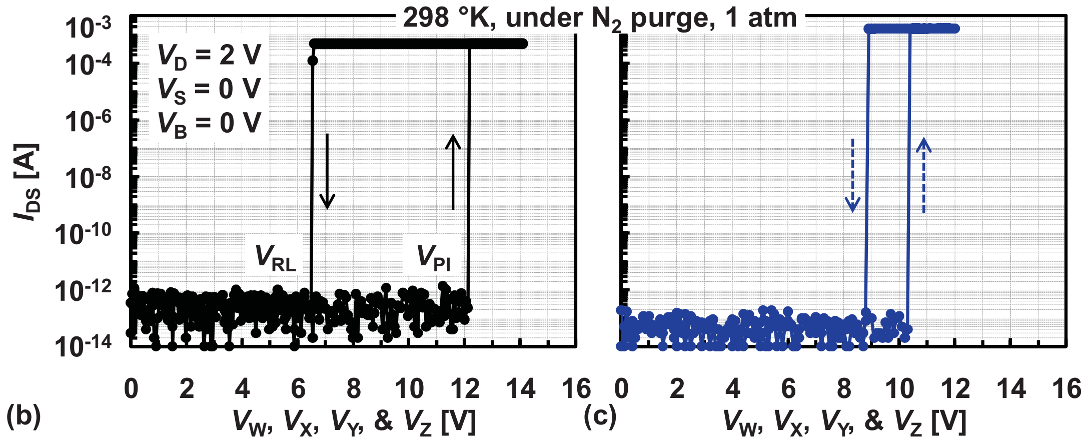
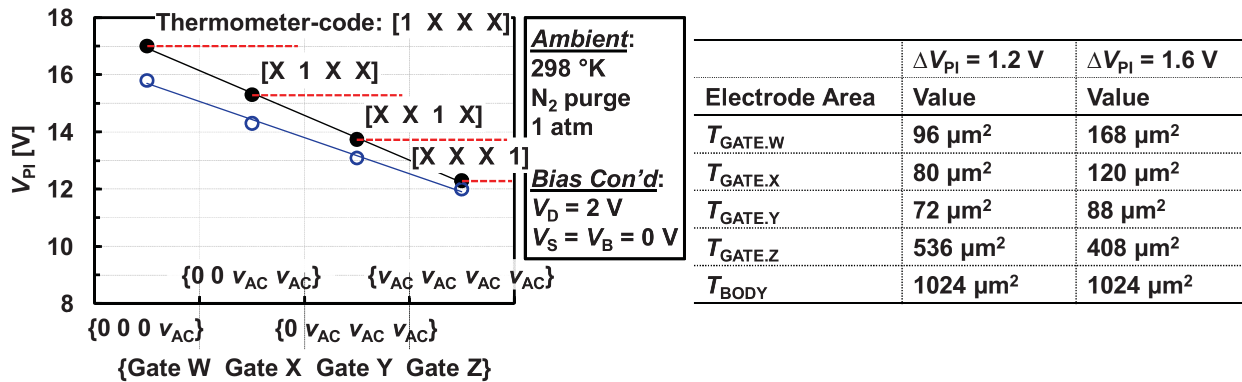


Figure 5.19: (**b**) Measured I_{DS}-V_G characteristic of the 4-input relay used herein *vs.* **(c)** that of an improved relay showing a reduced hysteresis [5].

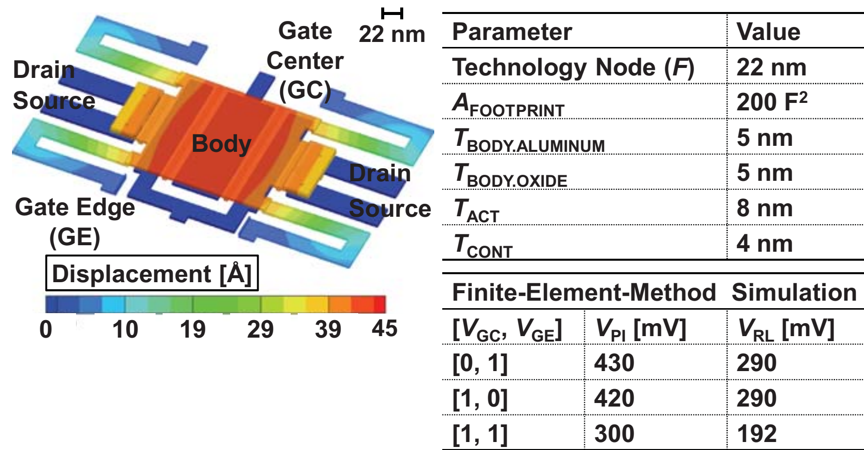


	ΔV_{PI} = 1.2 V	ΔV_{PI} = 1.6 V
Electrode Area	Value	Value
$T_{GATE.W}$	96 μm²	168 μm²
$T_{GATE.X}$	80 μm²	120 μm²
$T_{GATE.Y}$	72 μm²	88 μm²
$T_{GATE.Z}$	536 μm²	408 μm²
T_{BODY}	1024 μm²	1024 μm²

Figure 5.20: Measured, linearly-increasing V_{PI} of a 4-input relay with differently-sized gate electrodes for use in ADC.

5.5. Multi-Electrode Relay Scaling

The dimensions of multi-gate, dual-source/drain relays must be scaled down to reduce their operating voltage for improved energy efficiency as well as improved layout area efficiency. The operation of a multi-electrode relay scaled to a 22-nm-equivalent technology node was simulated using Finite-Element-Method (FEM) software [6], [7]. The results in Figure 5.21 indicate that sub-500-mV operation is achievable (*c.f.* V_{DD} projected for 8-nm CMOS technology is 0.6 V [8]).



Parameter	Value
Technology Node (F)	22 nm
$A_{FOOTPRINT}$	200 F^2
$T_{BODY.ALUMINUM}$	5 nm
$T_{BODY.OXIDE}$	5 nm
T_{ACT}	8 nm
T_{CONT}	4 nm

Finite-Element-Method Simulation		
[V_{GC}, V_{GE}]	V_{PI} [mV]	V_{RL} [mV]
[0, 1]	430	290
[1, 0]	420	290
[1, 1]	300	192

Figure 5.21: A 2-input relay scaled to a 22 nm technology node.

5.6. Summary

In this chapter, multi-gate, multi-source/drain relays are proposed and demonstrated for compact implementation of ultra-low-power digital electronics.

5.7. References

[1] K. Akarvardar, D. Elata, R. Parsa, G. C. Wan, K. Yoo, J. Provine, P. Peumans, R. T. Howe, and H.-S. P. Wong, "Design Considerations for Complementary Nano-Electro-Mechanical Logic Gates," in *IEEE International Electron Device Meeting (IEDM '07) Technical Digest*, pp. 299-302, December 2007.

[2] F. Chen, M. Spencer, R. Nathanael, C. Wang, H. Fariborzi, A. Gupta, H. Kam, V. Pott, J. Jeon, T.-J. King Liu, D. Marković, V. Stojanović, and E. Alon "Demonstration of Integrated Micro-Electro-Mechanical (MEM) Switch Circuits for VLSI Applications," in *IEEE International Solid-State Circuits Conference (ISSCC '10) Technical Digest*, pp. 150-151, February 2010.

[3] M. Spencer, F. Chen, C. Wang, F. Chen, R. Nathanael, H. Fariborzi, H. Kam, V. Pott, J. Jeon, T.-J. King Liu, D. Marković, E. Alon, and V. Stojanović, "Demonstration of Integrated Micro-Electro-Mechanical Switch Circuits for VLSI Applications," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 308-320, January 2011.

[4] H. Kam, V. Pott, R. Nathanael, J. Jeon, E. Alon, and T.-J. King Liu, "Design and Reliability of a Micro-Relay Technology for Zero-Standby-Power Digital Logic Applications," in *IEEE International Electron Device Meeting (IEDM '09) Technical Digest*, pp. 809-812, December 2009.

[5] R. Nathanael *et al.*, submitted to *IEEE 25th International Conference on Micro-Electro-Mechanical Systems (MEMS '12)*.

[6] *CoventorWare^{TM}2008 Designer Reference Manual*, Coventor, Inc., Cary, NC, 2008.

[7] *CoventorWare^{TM}2008 Analyzer Reference Manual*, Coventor, Inc., Cary, NC, 2008.

[8] International Technology Roadmap for Semiconductors (2009 Edition). *[Online]. Available: http://www.itrs.net/reports.html*

Chapter 6

Conclusions

This dissertation aimed to solve one of the important societal challenges we are facing today -- reducing greenhouse gas emissions by improving energy efficiency of future information-processing systems. To that end, a nano-electro-mechanical (NEM) relay technologies that offer zero *off*-state leakage current and thus allows ultimate V_{DD} scaling have been investigated.

In this work, design challenges for NEM relay technology were addressed. In particular, improvements in relay design and process technology, which led to the successful demonstration of purely-mechanical digital IC building blocks, were discussed. In addition, multi-electrode relays that would enable smarter design and compact implementation of zero-static-power digital electronics were discussed.

As a final remark, the author believes that NEM relay technology will pave a revolutionary pathway to realize truly "greener," "smarter" computing of future electronics systems (*e.g.* table-top and/or portable super-computer).

6.1. Contributions of This Work

The key contributions of this work are summarized, as follows:

In Chapter 2, improvements in four-terminal (4-T) relay design were demonstrated to provide for improved electrostatic integrity, *i.e.* a strong body bias effect with equally effective relay switching using either the movable electrode or the underlying actuation electrode as the gate, a dramatically reduced drain bias effect on the gate switching voltage, and a minimal actuation due to the parasitic channel-to-body potential.

In Chapter 3, a 4-T relay was demonstrated to function as a memory cell. Thus, a 4-T relay logic technology (reference Chapter 2) can offer embedded memory capability at no additional process cost.

In Chapter 4, a see-saw relay design that provides for perfectly complementary switching behavior that is symmetric about V_{DD} / 2 was proposed and demonstrated. This design can maximize the operating voltage margin, minimize the crowbar current, and improve the

reliability (by allowing active turn-*off*) of a relay-based logic technology. The complementary see-saw relay design is versatile for implementing basic logic functions and a latch function with a single structure.

In Chapter 5, multi-gate, multi-source/drain relays were proposed and demonstrated to enable smarter design and compact implementation of ultra-low-power digital electronics without incurring additional process cost.

6.2. Future Directions

Source/Drain (and Channel) Engineering: The Tungsten (W) film used for the source/drain and channel electrodes tends to oxidize, once it is exposed to air. The W native oxide (typically composed of conductive WO_2 and W_2O_5 and insulating WO_3 [1]) always resides in the contacting regions (at the surface of the source/drain and channel) of a released device. This oxide formation is undesirable, in that I_{ON} (and hence R_{ON}) of the relay can increase over the operating lifetime of the relay.

Similarly as for the switching mechanism of Resistive-Random-Access-Memory (ReRAM) [1], a relay in the *on*-state can show a high-resistance state when the drain-to-source bias (V_{DS}) is not sufficiently high -- perhaps due to the insulating WO_3 component. When the V_{DS} is increased beyond a certain level, however, it shows a low-resistance state -- perhaps because conducting filaments are formed (WO_2 and/or W_2O_5) in the oxides between the source/drain and channel electrodes.

In order to ensure reliable operation of relay-based circuits, stable I_{ON} (and hence R_{ON}) must be achieved. An alternative contact material, *e.g.* Ruthenium, conductive Ruthenium Oxide, or Titanium Nitride, should be adapted for the source/drain and channel electrodes.

Air-Gap Thickness Engineering: In this work, the as-fabricated, contact-gap-to-actuation-gap thickness ratio ($T_{CONTACT}$ to T_{ACT}) was chosen to be 0.5. This ratio needs to be smaller than one third, in order to minimize the hysteric switching behavior -- by allowing the relay to operate in non-pull-in mode. Otherwise, V_{DD} scaling would be limited by the release voltage of the relay. To achieve most energy-efficient relay operation, however, the relay should operate in pull-in mode (with a $T_{CONTACT}$-to-T_{ACT} ratio of 0.7 to 0.8) [2]. Therefore, energy-performance tradeoff between a relay chain operating in non-pull-in-mode *vs.* multi-stage CMOS inverter chain needs to be evaluated, if non-pull-in-mode operation is preferred for relay scaling.

Surface (or Contact) Engineering: A relay that operates in non-pull-in-mode would not eliminate the hysteric switching behavior, due to surface adhesion that is always present when the channel and source/drain surfaces are brought into contact. This hysteric switching behavior

is exacerbated if the relay is hermetically sealed in an inert environment, because the absence of any chemical films and absorbates at the interface usually increases surface adhesion [3]. From endurance standpoint, however, the contact surface should remain clean to ensure good R_{ON} stability over a longer lifetime [4]. Therefore, a novel material, *e.g.* Graphene, that shows low surface energy and adhesion, high current-carrying capability, and high stability in harsh environments, might be a compelling surface-coating material to minimize hysteresis as well as to achieve stable R_{ON}.

Gate Stack Engineering: The gate stack of a relay should show low strain gradient (on the order of 10^{-5} /μm or less) to achieve low operating voltages, since these require the formation of very small actuation and contact gaps with good uniformity. In addition, the gate stack material should be free of any mechanical reliability issues, such as fatigue, creep, and stress rupture, for use in digital IC products. These might be difficult to achieve with any polycrystalline film, because of columnar growth of micro-grains that cause strain gradient in film thickness and because of the micro-voids, dislocations, and defects that are inherently present along film grain boundaries. Therefore, a new material might be needed for the relay structure. For example, the hinge material of the Texas Instruments Digital Mirror Display (DMD) devices -- the amorphous $TiAl_3$ sputtered with 4 % Oxygen -- was found to yield the lowest stress relaxation [5]. With this material incorporated, a large array of DMD devices (more than 500,000 elements) that can rotate through a large angle (± 10 degrees) has been demonstrated to operate reliably with more than 10^{12} cycles without any mechanical failures (such as fatigue, stress rupture, and creep) [6], [7].

6.3. References

[1] W. C. Chien, Y. R. Chen, Y. C. Chen, A. T. H. Chuang, F. M. Lee, Y. Y. Lin, E. K. Lai, Y. H. Shih, K. Y. Hsieh, C.-Y. Lu, "A Forming-Free WO_x Resistive Memory Using a Novel Self-Aligned Field Enhancement Feature with Excellent Reliability and Scalability," in *IEEE International Electron Device Meeting (IEDM '10) Technical Digest*, pp. 19.2.1-19.2.4, December 2010.

[2] H. Kam, T.-J. King Liu, V. Stojanović, D. Marković, and E. Alon, "Design, Optimization, and Scaling of MEM Relays for Ultra-Low-Power Digital Logic," *IEEE Transactions on Electron Devices*, vol. 58, no. 1, pp. 236-250, January 2011.

[3] B. Bhushan, *Introduction to Tribology*, New York, NY: John Wiley & Sons, 2002.

[4] Q. Ma, Q. Tran, T.-K. A. Chou, J. Heck, H. Bar, R. Kant, and V. Rao, "Metal contact reliability of RF MEMS switches," *Proceedings of the SPIE*, vol. 6463, pp. 646305-1-13, January 2007.

[5] J. Tregilgas, "How We Developed an Amorphous Hinge Material," *ASM International Advanced Materials and Processes Magazine*, pp. 46-49, January 2005.

[6] M. R. Douglas, "Lifetime Estimates and Unique Failure Mechanisms of the Digital Micromirror Device (DMD)," in *Proceedings of the IEEE Annual International Reliability Physics Symposium (IRPS '98)*, pp. 9-16, March 1998.

[7] A. B. Sontheimer, "Digital Micromirror Device (DMD) Hinge Memory Lifetime Reliability Modeling," in *Proceedings of the IEEE Annual International Reliability Physics Symposium (IRPS '02)*, pp. 118-121, April 2002.