ABSTRACT

Title of Document:	ENABLING HARDWARE TECHNOLOGIES FOR AUTONOMY IN TINY ROBOTS: CONTROL, INTEGRATION, ACTUATION
	Tsung-Hsueh Lee, Doctor of Philosophy, 2016
Directed By:	Professor Pamela A. Abshire, Department of Electrical and Computer Engineering

The last two decades have seen many exciting examples of tiny robots from a few cm^3 to less than one cm^3. Although individually limited, a large group of these robots has the potential to work cooperatively and accomplish complex tasks. Two examples from nature that exhibit this type of cooperation are ant and bee colonies. They have the potential to assist in applications like search and rescue, military scouting, infrastructure and equipment monitoring, nano-manufacture, and possibly medicine.

Most of these applications require the high level of autonomy that has been demonstrated by large robotic platforms, such as the iRobot and Honda ASIMO. However, when robot size shrinks down, current approaches to achieve the necessary functions are no longer valid. This work focused on challenges associated with the electronics and fabrication. We addressed three major technical hurdles inherent to

current approaches: 1) difficulty of compact integration; 2) need for real-time and power-efficient computations; 3) unavailability of commercial tiny actuators and motion mechanisms. The aim of this work was to provide enabling hardware technologies to achieve autonomy in tiny robots.

We proposed a decentralized application-specific integrated circuit (ASIC) where each component is responsible for its own operation and autonomy to the greatest extent possible. The ASIC consists of electronics modules for the fundamental functions required to fulfill the desired autonomy: actuation, control, power supply, and sensing. The actuators and mechanisms could potentially be post-fabricated on the ASIC directly. This design makes for a modular architecture.

The following components were shown to work in physical implementations or simulations: 1) a tunable motion controller for ultralow frequency actuation; 2) a nonvolatile memory and programming circuit to achieve automatic and one-time programming; 3) a high-voltage circuit with the highest reported breakdown voltage in standard 0.5 μm CMOS; 4) thermal actuators fabricated using CMOS compatible process; 5) a low-power mixed-signal computational architecture for robotic dynamics simulator; 6) a frequency-boost technique to achieve low jitter in ring oscillators. These contributions will be generally enabling for other systems with strict size and power constraints such as wireless sensor nodes.

ENABLING HARDWARE TECHNOLOGIES FOR AUTONOMY IN TINY ROBOTS: CONTROL, INTEGRATION, ACTUATION

By

Tsung-Hsueh Lee

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2016

Advisory Committee:
Professor Pamela Abshire, Chair
Professor Nikhil Chopra, Dean's Representative
Professor Timothy Horiuchi
Professor Alireza Khaligh
Professor Robert Newcomb
Professor Elisabeth Smela

Acknowledgments

I would like to express my sincere appreciation to Dr. Abshire for her guidance and support through my whole doctoral course. This dissertation would not have been possible without her help. I would also like to thank my other committee members, Dr. Chopra, Dr. Horiuchi, Dr. Khaligh, Dr. Newcomb, and Dr. Smela, for their constructive feedbacks and comments. I am especially thankful for Dr. Smela for attending regular meetings to discuss the micro-fabrication project with me and provide valuable suggestions.

My thanks also go to my colleagues who have directly or indirectly contributed to this dissertation: Dr. Marc Dandin, Dr. Timir Datta-Chaudhuri, Ms. Deepa Sritharan, Mr. Andrew Berkovich, Mr. Michael Kuhlman, Mr. Bathiya Senevirathna, Mr. Alex Castro, Mr. Ivan Penskiy, Dr. Haoyu Wang, Dr. David Sander, Dr. Babak Nouri, and Mr. Nolan Ballew. I also thank Mr. Tom Loughran, Mr. Jonathan A. Hummel, and Mr. John Abrahams in the Maryland FabLab for their help on my micro-fabrication project. I thank Mrs. Melanie Prange for being so concerned about my situation and for her encouragement. I am grateful to all my friends in Taiwan and the United States for being supportive, listening to my complaints, and making my life more fun.

My greatest appreciation goes to my wife Yi-Ting Hou for her love over my entire doctoral course. She gave up her professional career and came with me to an unknown country to support me pursing my Ph.D. degree. I greatly thank my parents Chen-Nan Lee and Yue-Yin Lee for their unconditional love and support. They have been always tirelessly trying to provide me the best they can offer.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AFM	Atomic Force Microscope
ASIC	Application-specific Integrated Circuit
BJT	Bipolar Junction Transistor
BLPS	Body Length Per Second
CMOS	Complementary Metal-Oxide-Semiconductor
COTS	Commercial Off-the-shelf
CPG	Central Pattern Generator
CS	Current-starved
CSI	Current-starved Inverter
CSVCO	Current-starved Voltage-controlled Oscillator
CTE	Coefficient of Thermal Expansion
DB	Divide-by
DB2	Divide-by-2
DDB2	Dynamic Divide-by-2
DFF	D Flip-flop
DVFS	Dynamic Voltage Frequency Scaling
EDS	Energy Dispersive Spectrometry
FD	Frequency Divider
FGPLL	Floating-gate Phase-locked Loop
FOM	Figure-of-merit
GCSINV	Gated Current-starved Inverter
GIDL	Gate-induced Drain Leakage
GR	Guard Ring
HV	High Voltage
IR	Infrared
PCB	Printed Circuit Board
PDMS	Polydimethylsiloxane
PGMEA	Propylene Glycol Monomethyl Ether Acetate
MEMS	Microelectromechanical Systems
MITEs	Multiple Input Translinear Elements
MOSFET	Metal-oxide-semiconductor Field-effect Transistor
NMOS	N-type Metal-Oxide-Semiconductor
PCB	Printed Circuit Board
PDMS	Polydimethylsiloxane
PFD	Phase-frequency Detector
PLL	Phase-locked Loop
PMOS	P-type Metal-oxide-semiconductor
PSD	Power Spectral Density
PVA	Polyvinyl Alcohol
RF	Radio Frequency
RHC	Receding Horizon Control
RNG	Random Number Generator
RO	Ring Oscillator

RRHC	Randomized RHC
SMA	Shape Memory Alloy
SFDR	Spurious-free Dynamic Range
SoC	System on Chip
TCA	Transconductance Amplifier
TG	Transmission Gate
TGDB2	Transmission Gate Divide-by-2
THD	Total Harmonic Distortion
TSC	Triangle-to-sine Converter
μC	Microcontroller
VCO	Voltage-controlled Oscillator
VLSI	Very-large-scale Integration
WSAMP	Wide-swing Amplifier

Chapter 1: Introduction

1.1 Research Motivation and Purpose

It is hard to comprehensively define what a robot is. They appear in a large variety of forms and have become closely linked to our daily lives. Robots are now widely used in production lines for assembling automobiles and electronics devices [1]. They can also be found in applications like home cleaning (iRobot), military tasks [2], mining [3], and entertainment [4]. Recently, they have been used to replace human workers in warehouses [5], facilitate complex surgery [6], and provide healthcare [7].

As the technology progresses, electronics and mechanical components can be made smaller and more powerful. This enables people to build smaller robots. There have been many exciting examples of tiny robots (from a few cm^3 in volume to much less than one cm^3) in the last two decades [8-27]. Tiny robots have advantages of being small, stealthy, and disposable. Although they are individually simple and limited, a large group of these robots, commonly called a swarm, has the potential to work cooperatively and accomplish complex tasks that are far beyond their individual capability. Two good examples in nature are ant and bee colonies. Tiny robots have the potential to be able to operate in small and hazardous spaces that are not reachable by humans and large machines. One application might be to search for survivors and collect data in a disaster site which is exposed to fire, radiation, or fatal poisons [9]. Tiny robots have the potential to perform operations with minimal change to the

environment and avoid further damage, for example a rescue operation in an unstable infrastructure or a space full of inflammable gas. They are also suitable for applications like military scouting, infrastructure and equipment monitoring [28], micro-assembly [28], nano-handling , and possibly medicine [28, 29]. In conclusion, we believe tiny robots will redefine some existing applications and broaden many others.

Most of the applications mentioned above require autonomy of the tiny robots. Autonomy of robots was defined as the integration of sensing, actuation, power, and control on a single chassis by Churaman [19] or a micro-system by Fukuda [30]. Although robotic autonomy is task-dependent, generally speaking there are a few requirements that are more important: 1) Autonomous robots sense the environment, acquire and extract information that is useful to complete the designed tasks; 2) They possess local control (or computational power) to make decisions related to their tasks and not to destroy themselves and the mission; 3) They should have reasonable mobility to maneuver around by being able to move at a speed of one half body length per second (BLPS); 4) The first three actions can be performed for an extended period that is a few times longer than the tasks require; 5) Most importantly, the robots have to operate independently with no or limited human assistance, providing instructions or energy for example.

Some robot platforms already demonstrated a high level of autonomy, for example iRobot, Honda ASIMO [31], HRP-3 [32], and Sony AIBO. However, they are usually much larger than many cm^3. When robot size shrinks down to a few cm^3 and below, current approaches are no longer valid. Existing tiny robots do not meet

all of these requirements and only present limited autonomy. Therefore, the purpose of this research is to provide enabling hardware technologies for tiny robots with a focus on the ***electronics*** module (control and computation), ***fabrication*** of tiny mechanism, and ***integration*** of these components to achieve a high level of autonomy and bridge the technology gap between large autonomous robots and tiny ones.

1.2 Literature Survey

1.2.1 Existing Tiny Robots

In this work, we are interested in tiny robots with a volume of a few cm^3 and below. In the following we summarize in the following several representative tiny robots that were contributed by other researchers in the literature. Table 1.1 summarized important features that are related to robotic autonomy. In order to get more insight of the bottleneck to shrink the size of autonomous robots, we include some robots with a volume of many cm^3 that are larger than the interested size. The robots are ordered in descending volume and inherent functionality to some degree. The volume presented here is not the exact volume of the robots but the space the robots will occupy while they operate. Some of the volumes were approximated if not provided by the authors. These robots will be briefly introduced; the first paragraph will introduce some important features of each robot while the second paragraph will focus on the electronics design and moving mechanism. Most of these projects integrated the electronics modules on a printed circuit board (PCB) while some other approaches were also pursued. A variety of actuators and moving mechanisms were fabricated for the tiny robots.

Table 1.1 Summary of Important Features of Existing Robots (red indicates where the robots fail to meet the general requirements)

Robot	Size (cm^3)	Actuation	Abs. Speed, BLPS	Motion Mechanism	Comm.	Sensor	Independent Control (type)	Power, Spec.	Duration
Kilobot	23.33	Motor	1 cm/s, 0.33	Stick-slip	IR	Distance	Y (Swarm, Local)	Battery, 160 mAh	3-24 hr
HAMR[3]	14.1	Piezoelectric	3 cm/s, 0.625	Walking	None	None	N	Battery, 8 mAh	2 min
RoACH	11.25	SMA	3 cm/s, 1	Walking	IR	None	N (External)	Battery, 20 mAh,	9 min
Alice	7.93	Motor	4 cm/s, 2	Wheel	IR, radio	Proximity	N (External, Local)	Battery, 69 mAh	10 hr
MiCRoN	4.32	Piezoelectric	0.4 mm/s, 0.03	Stick-slip	IR	AFM	N (External, Local)	Inductive, ~5 V	∞
Jumpingbot	0.112	Chemical	Jump 8cm, 11 BL	Jumping	None	Optical	Y (Local)	Capacitor, 6 V	8 min
I-SWARM	0.064	Piezoelectric	0.207 mm/s, 0.05	Stick-slip	IR, optical	Distance	N (External, Local)	Solar, 1.4 mW	∞
WalkingS	0.038	Thermal	6 mm/s, 0.43	Walking	Wire	None	N (External)	Wire, > 1 W	∞
Siliconbot	0.017	Electrostatic	0, 0	Walking	None	None	Y (Local)	Solar, 100 μW	∞

Kilobot (23.33cm^3), Rubenstein, Harvard University [8, 33]

Kilobot (shown in Figure 1.1) is sized 3.3 cm in height and 3 cm in diameter. This robot was designed to test collective algorithm on hundreds or thousands of robots and was aimed for mass production and collective operations. Kilobot has part cost of $14 and takes only 5 minutes to assemble by hand. The authors claimed that it can be collectively programmed, powered on, and charged and allows a single user to operate a large swarm easily [8].

Kilobot features an onboard microcontroller (μC) to perform computations and control its behavior. The robot has two sealed coin shape vibration motors. The vibration is converted to a forward force based on stick-slip principle. Due to the low efficiency of this moving mechanism, the maximum speed is only 1 cm/s, 0.33 BLPS. Moreover, the robot can only move on smooth surface. The control signals for the motors are generated by the μC using pulse width modulation. It is also able to communicate with neighbors and overhead control station via an infrared (IR) transceiver. The robot can estimate the distance between itself and neighbors by measuring incoming IR light during communicating. The robot is equipped with a 3.4 V, 160 mAh rechargeable lithium-ion battery which can power the robot to remain active for 3-24 hours.

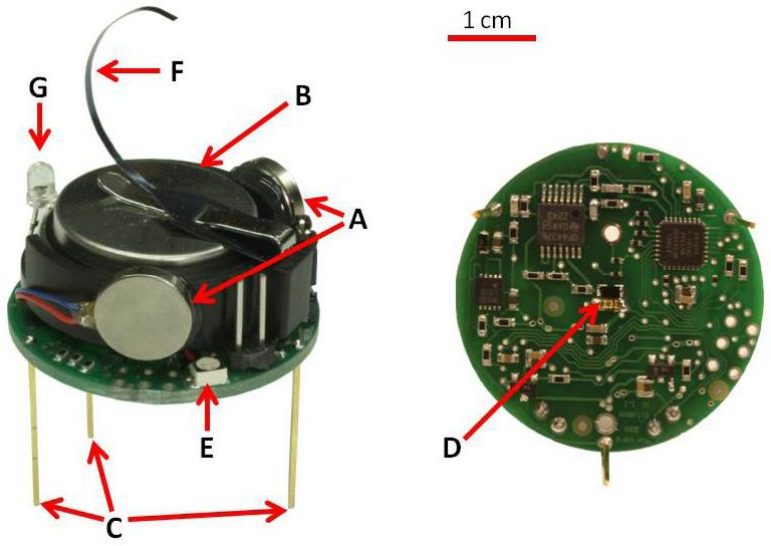


Figure 1.1 Components on Kilobot: A) Vibration motor, B) Battery, C) Supporting legs, D) Infrared transceiver, E) RGB LED, F) Charging tab, and G) Light sensor [8]. Figure taken directly from the reference [8].

HAMR^3 (14.1 cm^3), Baisch, Harvard University [9]

The third generation Harvard Ambulatory MicroRobot (HAMR^3) in Figure 1.2 is a 1.7g ambulatory hexapod robot with a size of 4.7 × 2.0 × 1.5 cm^3. This work focused on developing a light-weight walking platform to study the dynamics of small scale locomotion and to improve design and fabrication techniques.

HAMR^3 does not have any sensor or designed intelligence but has the potential to include them. A ramped square wave is used as the driving signal for the actuators and is generated by filtering a binary output from the μC. HAMR^3 has 9 piezoelectric actuators with actuation voltage of 200 V to control the lift and swing of the six legs. The 200 V bias is produced by an inductor boost converter from a 3.7 V, 8 mAh, 330 mg lithium polymer battery. However, this robot is not capable of turning. The battery can only power the robot to walk for 2 minutes at an optimal actuation frequency of 20 Hz when the average speed is 3 cm/s (0.625 BLPS).

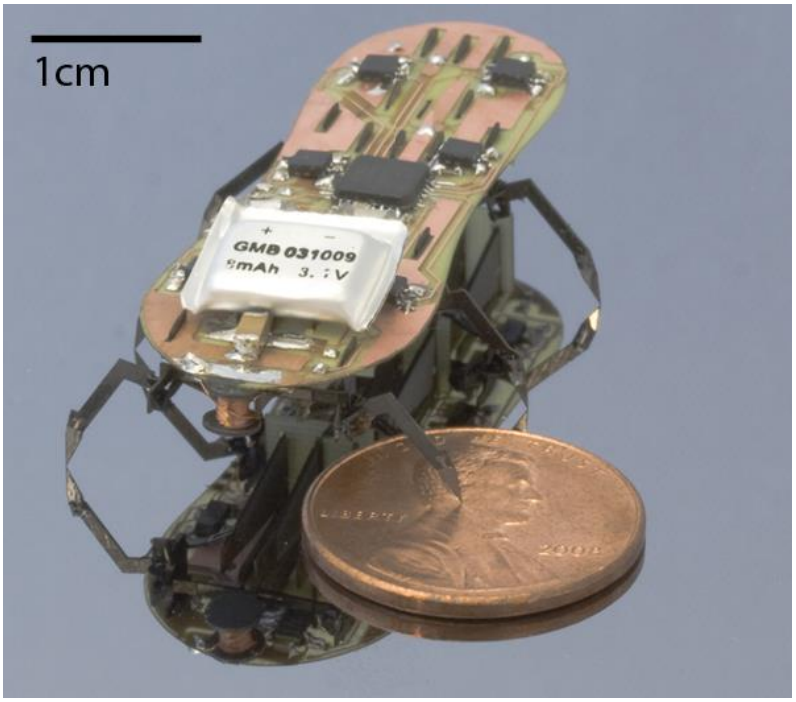


Figure 1.2. HAMR^3 on a US one cent coin [9]. Battery, electronics components, and actuators are on both sides of the PCB. Figure taken directly from the reference [9].

RoACH (11.25 cm^3), Hoover, U. C. Berkeley [10]

RoACH is another hexapod robot sized 3 × 2.5 × 1.5 cm^3 and weighs 2.4 g (shown in Figure 1.3). The skeleton of the robot was fabricated using smart composite microstructures process.

RoACH chose shape memory alloy (SMA), as its actuator. SMA is one type of thermal actuator requiring 13.6 V and 60 mA for actuation. A standard boost converter is responsible for providing the high voltage supply for the SMA. Two actuators are respectively driven by two channels of 19 KHz based pulse width modulation signal generated by a μC to achieve alternating tripod gait and some degree of turning. The maximum speed is 3 cm/s, 1 BLPS. The onboard battery is a 3.6 V, 20 mAh lithium polymer battery which accounts for 847 mg, over one third of the total weight of the robot. Continuous actuation at 3 Hz resulted in a maximal

running time over 9 minutes which indicated an average power consumption of 0.48 W. The robot can communicate with the station using infrared data association (IrDA) protocol for gait change. Sensors were not implemented on this platform.

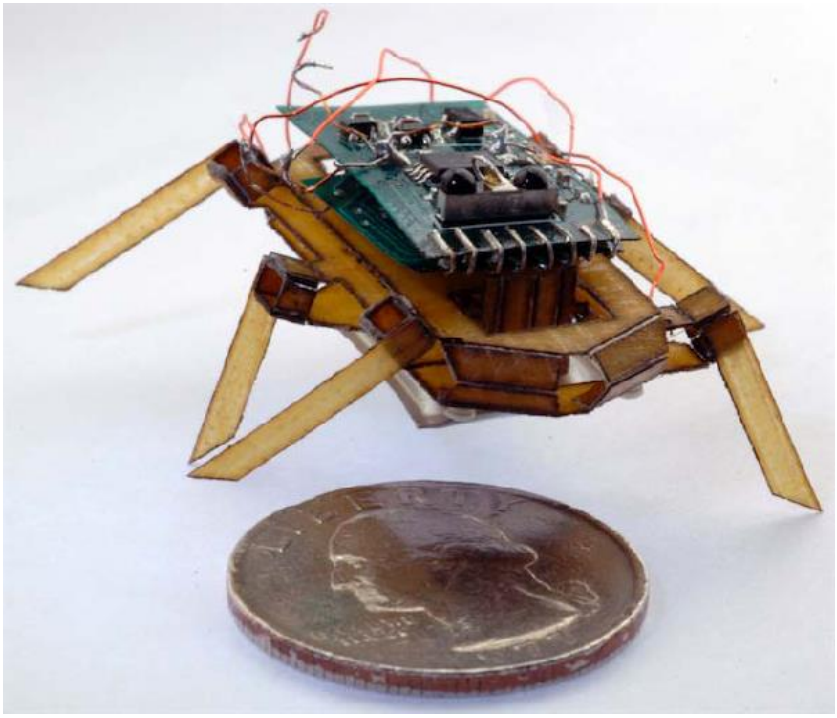


Figure 1.3. RoACH with a U.S. quarter as scale [10]. The robot has a skeleton fabricated using a micro process and electronics on the top. Figure taken directly from the reference [10].

Alice (7.93 cm^3), Caprari, EPFL [11-14]

Alice has many variants across several years: Alice 97, Alice 98, Alice 99, Alice 2002, AliceAllTerrain, LAMAlice, and more. However, the details of each version have not all been reported. The information used here is mostly from an implementation reported in 2001 [13]. The development of Alice sought a modular solution to make the robots more accessible to different applications. Many "plug and play" modules were implemented to enhance the functionality of the base module. Its size is 2.1 × 2.1 × 1.8 cm^3 with a weight of 5 g.

The motion of Alice [13] is provided by two wheels driven by 2 watch motors. Each of the watch motors consumes 1.5 mW. The control signal for the motors is from the μC. Maximum speed of Alice is 40 mm/s, 2 BLPS. Obstacle avoidance and wall following were implemented on the μC. The batteries last for 10 hours under an average overall power consumption of 4-10 mW. Alice is equipped with four proximity sensors which can be used for local communication. A radio transceiver is also available onboard for long distance communication. The robot is powered by three button batteries, each of which is 1.5 V and 23 mAh.

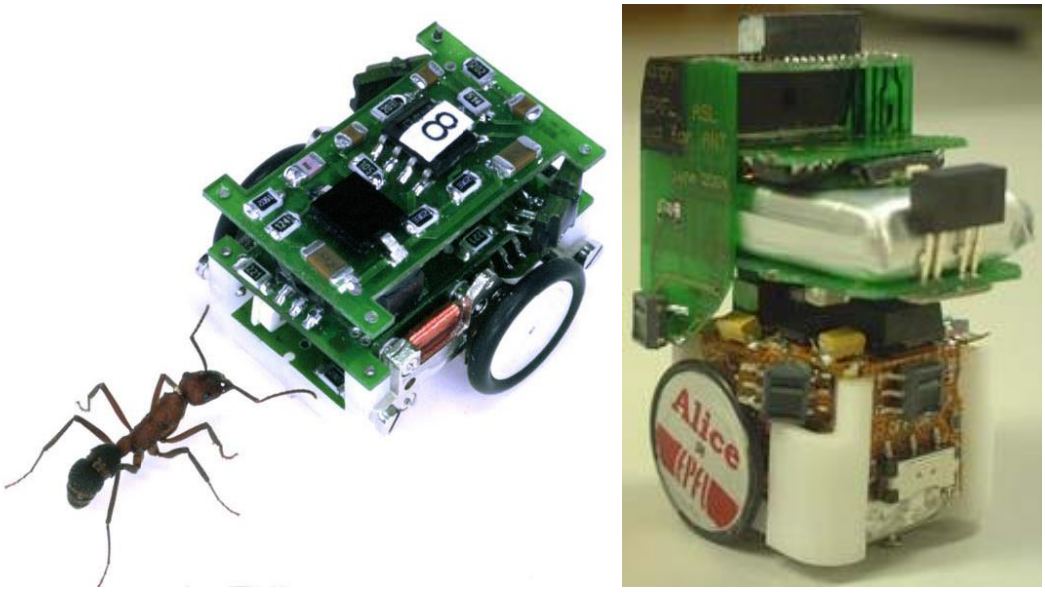


Figure 1.4. (Left) Alice in basic configuration [13]. Figure taken directly from the reference [13]. (right) Alice 2002 with ANT extension which includes one large battery, a 128 pixels linear camera, 2 proximity sensors, and 1 long distance sensor [14]. Figure taken directly from the reference [14].

MiCRoN (4.32 cm^3), European Union Project [15-17]

MiCRoN was designed to handle tasks like cellular manipulation and nano-manipulation with an integrated atomic force microscope (AFM) tip. The size of this robot is 1.2 × 1.2 × 3.0 cm^3 (Figure 1.5).

MiCRoN has some electronics modules implemented using an application specific integrated circuit (ASIC) approach. Three types of programmable waveforms are generated by the control ASIC and then amplified to high voltage (20 V) by two amplification ASICs to drive the piezo-actuators. Piezoelectric actuation provides locomotion of the robot and positioning of the AFM tip. A maximum speed of 0.4 mm/s (0.03 BLPS) was achieved. The power for this robot is provided by a power floor through inductive coupling. The robots are wirelessly powered by a voltage supply of 3.5-6 V. However, this power floor also limits the working range of the robots. The robots communicate with the station to acquire instructions and send back information via IrDA protocol.

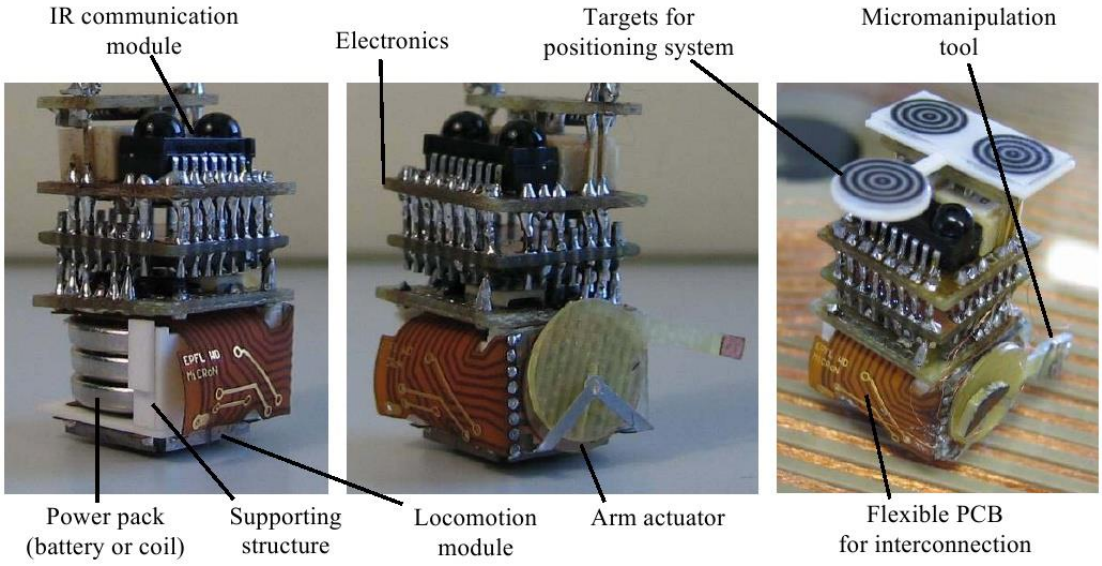


Figure 1.5. Three pictures showing the structure of MiCRoN. The power floor is shown under the robot in the rightmost picture [34]. Figure taken directly from the reference [34].

Jumping Microrobot (0.112 cm^3), Churaman, University of Maryland [18, 19]

Jumping Microrobot is a robot that is able to jump. The chassis is made of Loctite 3525. It is sized 0.4 × 0.7 × 0.4 cm^3 and weighs 300 mg.

Jumping Microrobot has sensing, control, actuation and power integrated onto one chassis. A circuit was designed to follow a simple behavior: sense a change in light intensity and response with a jump. A special chemical actuation, energetic nanoporous silicon, is used on this robot. The chemical requires 150 mA for 100 us to ignite. The power of this robot is from two capacitors (10 and 22 μF) which are pre-charged to 6 V. The robot was demonstrated to jump 8 cm height. However, the current design can only offer one time actuation.

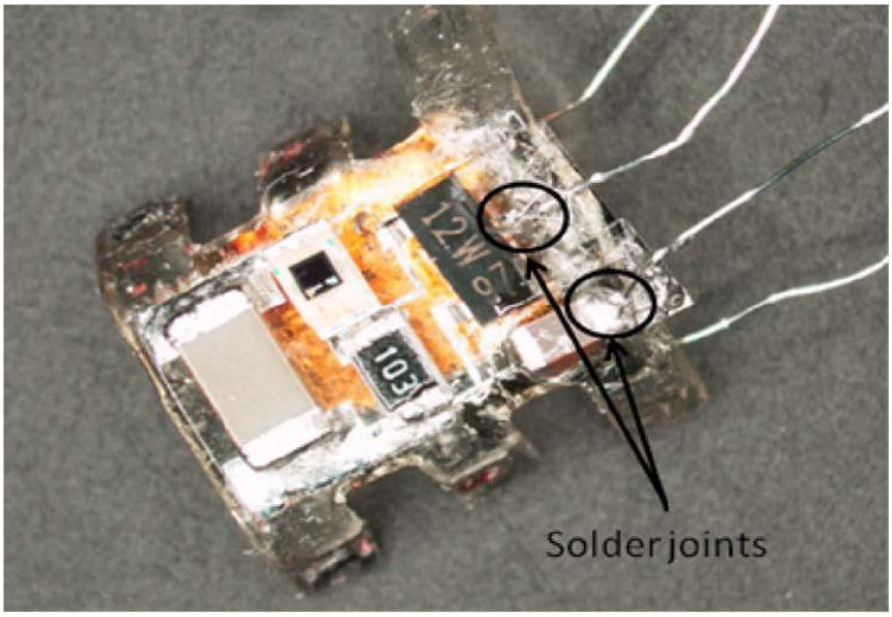


Figure 1.6. Jumping Microrobot with integrated components onto its polymer chassis using low melting temperature alloy [19]. Solder wires are used to charge the onboard capacitors. Figure taken directly from the reference [19].

I-SWARM (0.064 cm^3), European Union Project [20-23, 28]

The size and weight of I-SWARM are 0.4 × 0.4 × 0.4 cm^3 and 70 mg, respectively. The components are integrated on a flexible PCB which is then folded to the configuration shown in Figure 1.7.

The authors for I-SWARM pursued an ASIC approach to minimize the size of the robot. A system on chip (SoC) was implemented for the required control and computation of the robot. The local communication between neighboring robots is via IR while communication from the station to the robots is via on/off of a light source. Since this robot does not have any non-volatile memory, it has to be programmed for about 50 minutes every time after start-up. Fortunately, the programming can be done collectively. A vibration contact sensor on the robot can detect objects in front of the robot. Piezoelectric actuators are used to provide motion of the robot. The driving signal is a 3.6 V square waveform divided from a high frequency signal from the SoC. The robot can move forward at a speed of 0.207 mm/s (0.05 BLPS) but is not able to turn. The robots have to work in a controlled area which is illuminated by a high intensity light. Under this condition the solar cells can generate 1.5 mW to power the robot.

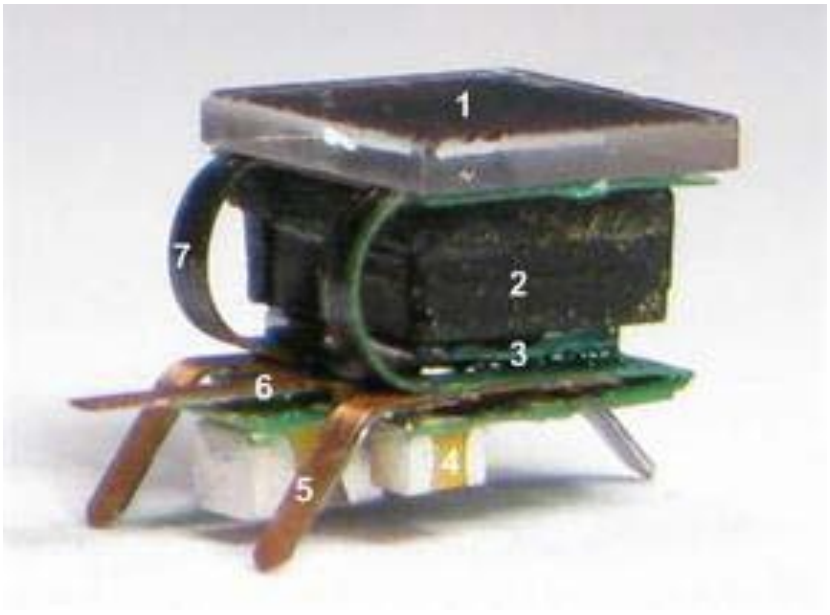


Figure 1.7. I-SWARM picture taken directly from the reference [21]. 1) Solar cells, 2) IR module, 3) SoC ASIC, 4) Capacitors, 5) Locomotion unit, 6) Vibrating contact sensor, and 7) Flexible PCB. F

Walking Silicon (0.0375 cm^3), Ebefors, Royal Institute of Technology [24-26]

The goal of this project was to fabricate a micro-robot that is capable of conveying large loads compared to the robot weight. One interesting feature of this robot is that it can be used upside down to more rapidly convey heavier loads. This robot is sized 1.4 × 0.7 × 0.05 cm^3 and weighted 80 mg. The robot can carry a maximum of 2500 mg as shown in Figure 1.8.

This robot does not have any onboard electronics. The control signal is provided externally from wires. The robot has a silicon chip as its substrate. Twelve legs with polyimide joint were fabricated on the silicon substrate. The legs are thermally actuated. The maximum speed 6mm/s (0.43 BLPS) was achieved under actuation signal of a 18 V, 100 Hz, square wave, a power consumption of 1.1 W, and a load which is four times the robot weight. The robot is tethered using bond wires to acquire power and control signal.

Figure 1.8. Walking silicon robot carries a load. Power and control signal are provided externally through the bond wires [26]. A match on the right is used to indicate the size. Figure taken directly from the reference [26].

Silicon Robot (0.017 cm^3), Hollar, U. C. Berkeley [27]

The main goal of this project was to fabricate an untethered walking micro-robot. The weight of this robot is 10 mg and its size is 0.85 × 0.4 × 0.05 cm^3.

The authors pursued a three-chip solution to fulfill robotic functions of powering, control, and actuation as shown in Figure 1.9. The control chip outputs digital control signal for the legs based on a clock signal. The control signal is then converted to 50 V by the high voltage buffer. Solar cells were fabricated on the same chip as the buffers. The 90 solar cell array generates 50 V and 100 uW to power the robot. The high voltage buffer and solar cells require a specialized process and is not compatible with the digital control chip. Therefore, multiple chips are needed. Electrostatic actuation of two legs with different length makes the robot crawl but only one of the tested robots walked for the first 250 cycles on a silicon surface.

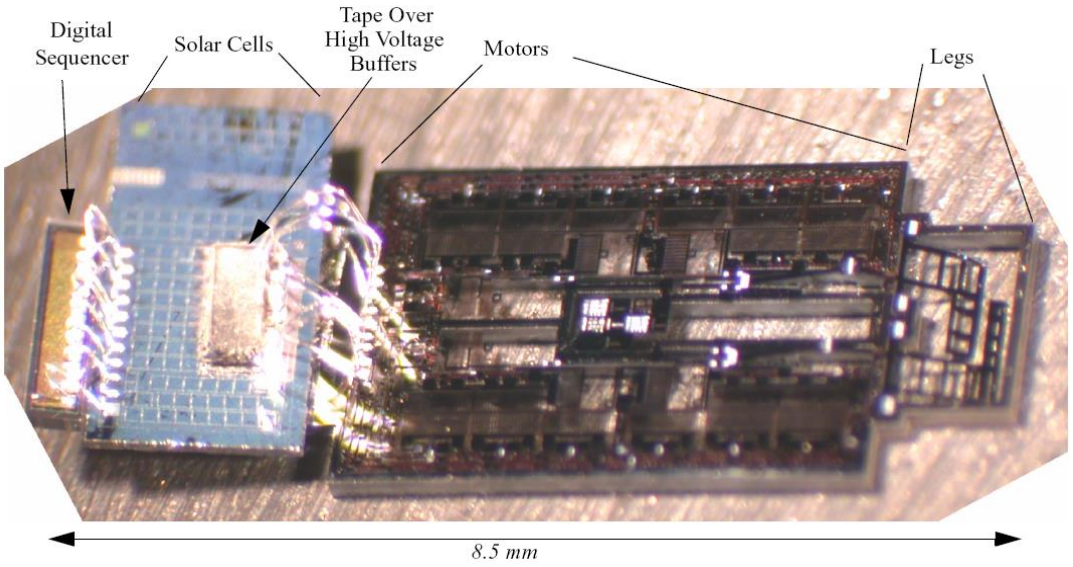


Figure 1.9. Silicon Robot with three chips with all major components indicated in the picture [27]. Figure taken directly from the reference [27].

1.2.2 State of the Art: Tiny Robots

Some important features were summarized in Table 1.1. Among these robots we have introduced in the last section, only Kilobot and Alice achieve a high level of autonomy. Although they are significantly larger than 1 cm^3, they still do not meet all the general requirements we have identified. Kilobot has inefficient and unpredictable actuation; it can only move at one third BLPS. Its precise control of the movement depends on feedback from other robots. Alice relies on external instructions to perform tasks and manual change of the batteries to remain powered. Based on this review, it is clear that there remains a significant technology gap between the large autonomous robots and the tiny robots as shown in Figure 1.10. This technology gap will be introduced in the next section.

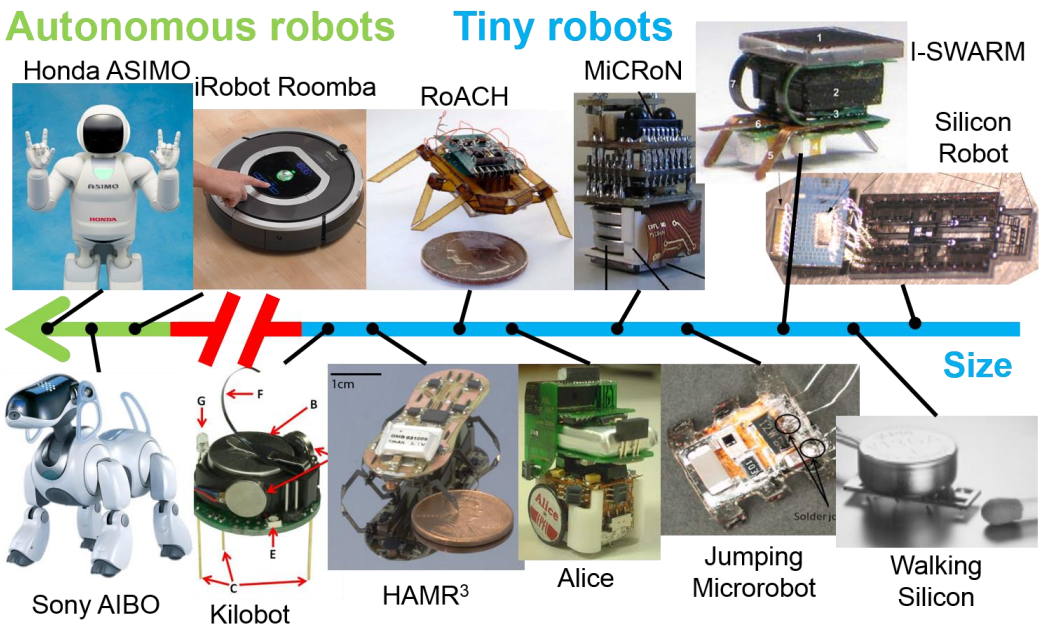


Figure 1.10. Some representative robots ordered by size showing the technology gap between autonomous robots and tiny robots. Figures taken directly from the references.

1.3 Discussion of Challenges on Electronics and Motion Mechanisms

Most of these introduced robots were assembled using commercial off-the-shelf (COTS) components for electronics modules and mechanisms. This popular approach in the robotics community offers short design time, flexible integration, fast manufacturing turnaround time, and low cost for prototyping. The electronics modules on Kilobot, HAMR^3, RoACH, Alice, Jumping Microrobot, and part of MiCRoN are all COTS. When the robot size scales down to a few cm^3 and below, size and power constraints become strict. Multiple COTS components that are required to provide complex functionality for the robot are no longer affordable in terms of size. Also, COTS components do not usually offer the best performance. They might have redundant devices to consume extra power or be overdesigned to enhance robustness and flexibility. Under such scenarios, COTS components become unfeasible to meet all our requirements with limited volume and power. COTS mechanisms that can apply to tiny robots and produce reasonable motion are not available at this scale. Therefore, they have to be specially fabricated.

Kilobot, HAMR^3, RoACH, Alice, and I-SWARM utilize a μC for all the required computations. The computations might include but are not limited to control of all the other components, processing of sensor data, handling communications, generating actuation control signal, and making decisions. In our experience processing sensor data to estimate distance on a TI MSP430 (a 16-bit μC) needed to be carefully optimized so that it could run in real time [35]. Therefore, the 8-bitμCs that are used for computation in small robots due to their low power feature are not computationally powerful enough and not energy-efficient enough to handle all the

computations while meeting tight power budgets. Thus, the μCs have to operate intensely at a high clock rate and remain active for most of the time. This might result in malfunctions like exceeding the power budget or causing delays in response.

Another challenge is the requirement of separate complementary-metal-oxide-semiconductor (CMOS) chips or PC boards for integration of HV devices. Many actuators that are suitable for tiny robots require high operating voltage to attain reasonable motion. I-SWARM compromised on high voltage electronics and can only move at an extremely slow speed (0.05 BLPS). Many sensors that can be integrated on CMOS chips benefit from having a HV bias. Since HV devices are usually not compatible with standard CMOS technologies, the necessity of high voltage electronics fabricated on a separate chip or board prevents compact integration. For example, MiCRoN and Silicon Robot have ASICs that were fabricated using a specialized high voltage process and are separated from the control ASICs operating at normal voltage.

To sum up, when the robots size shrinks down, current approaches are no longer valid. We have introduced three distinct problems that hinder the development of autonomous tiny robots. We have described three problematic approaches that are widely used in implementation of current robots: 1) usage of COTS electronics components and mechanisms; 2) centralized computations relying on a low-end μC; 3) separate chip or board for high-voltage devices. We therefore have identified three major challenges that need to be addressed for the realization of autonomous tiny robots: 1) compact integration; 2) real-time and power-efficient processing; 3) unavailability of COTS tiny motion mechanisms.

In order to overcome these challenges, we have to seek other solutions. First, alternate electronics modules and motion mechanisms other than COTS components has to be adopted to meet the strict size and power requirements. Second, dependence on and loading of μCs have to be decreased to achieve real-time and power-efficient processing. Third, high voltage devices are desired to be integrated with other low voltage electronics on the same chip.

1.4 Proposed Approach to Implement Tiny Walking Robots

The proposed approach seeks to provide enabling technologies for the electronics module and motion mechanism design for fully autonomous tiny robots. To overcome or mitigate the three challenges summarized in the last section we propose a decentralized single-ASIC platform where each component is responsible for its own operation and autonomy to the greatest extent possible. Many of the components can be custom designed and fabricated to meet specific specifications and requirements. The comparison for COTS and full-custom implementation approach for electronics design are summarized in Table 1.2. Full-custom generally provides much better performance such as smaller sizes, less overhead, higher robustness, and lower cost for mass-production but requires more design effort and higher cost more for prototyping stage.

Table 1.2 Comparison of COTS and full-custom implementation approach for electronics design. Numbers given here were estimated based on general cases.

Approach	COTS	Full-custom
Size	Larger than 10 mm^3 due to individual packages	Can be as small as a single die, < mm^3
Overhead	Additional functions not required	Designed to meet specific requirements
Integration	More flexible and easier	Difficult although fewer components
Design time	Several weeks	Several months
Turnaround time	1 month	> 3 months and rely on the foundry
Robustness	Low due to integration of more components	High
Cost	Low for prototyping (several dollars per part)	Low for mass production (< 1 dollar per part)

The platform consists of all the electronics modules with HV circuits and can be integrated with other required components so the size is minimized. Four fundamental functional components, **actuation**, **control** (or computation), **power** supply, and **sensing**, are included to fulfill the general autonomy requirements for tiny robots. Note that these functional components are not completely exclusive from one another. For example actuation and sensing might involve some computations, power supply and actuation might involve some control. Other specialized components such as communications can also be included. The system may need central processing for coordination but this would argue against the use of central processing for component specific functions. The platform serves as the body of the robot; the actuators and motion mechanisms will be post-fabricated on the ASIC directly. Walking is not as efficient as using wheels but is potentially more adaptable to different kinds of terrains. This design makes for a modular architecture. This system is an improved implementation inspired by Walking Silicon [24-26] and

Silicon Robot [27]. Silicon Robot adopted multiple-chip integration to fabricate micro-robots with onboard control. The functions for the three chips are: 1) A CMOS chip to provide control signals to the actuators; 2) A solar-cell and high-voltage chip to harvest energy and to drive the actuators. 3) A silicon chip with actuators. Walking Silicon was the first demonstration of walking at chip scale. However, the control is provided through wires. Our design seeks a single chip platform with multiple functionalities and could potentially be a minimum-size implementation for micro-robots.

An illustration of the final form is shown in Figure 1.11. It is called a "legged chip". The legged chip idea was first proposed in a National Science Foundation (NSF) funding proposal that was later awarded (award # 0931878). The proposal was submitted by four faculty members at the University of Maryland, College Park: Dr. Nuno Martins (ECE Dept., PI), Dr. Pamela Abshire (ECE Dept., co-PI), Dr. Elisabeth Smela (ME Dept., co-PI), and Dr. Sarah Bergbreiter (ME Dept., co-PI).

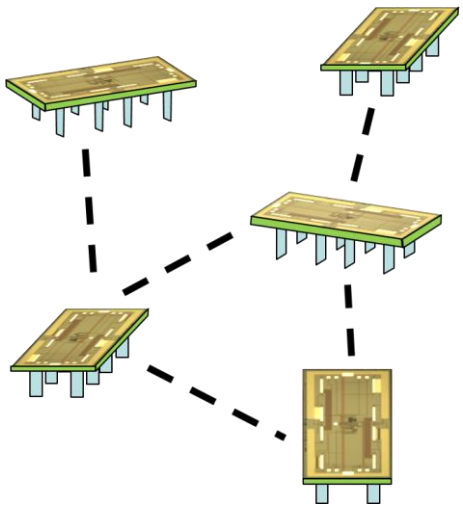


Figure 1.11. Legged chip consists of a single unpackaged ASIC (the rectangles with patterns) with mechanisms (blue color) fabricated directly on top of the CMOS. The moving mechanism in this figure is walking with legs. Dashed lines represents the communications network between the robots.

One implementation for the electronics would naturally be a full-custom single ASIC system (demonstrated in Figure 1.12). Design of a μC could take tens of engineers to spend a few years to finish. In a commercial implementation the μC could be custom designed or addressed by purchasing intellectual property (IP) from a vendor. Many open source architectures are also available such as OpenCores [36]. There are already some mature technologies to fabricate sensors on chip including temperature [37, 38] and image (light) [39, 40]. Some others can be integrated on chip without too much effort, for example pressure [41, 42] and gas [43, 44]. Technologies for energy storage like capacitors and batteries can also be integrated on chip [45, 46]. An intermediate option is an ASIC including most components that require specialized computations to achieve or accelerate the function, for example signal processing, control, and power electronics. At the same time COTS components can still be used for those functions that require extra design effort. This implementation retains most advantages from single ASIC implementation like good performance and flexible integration while greatly reducing design efforts.

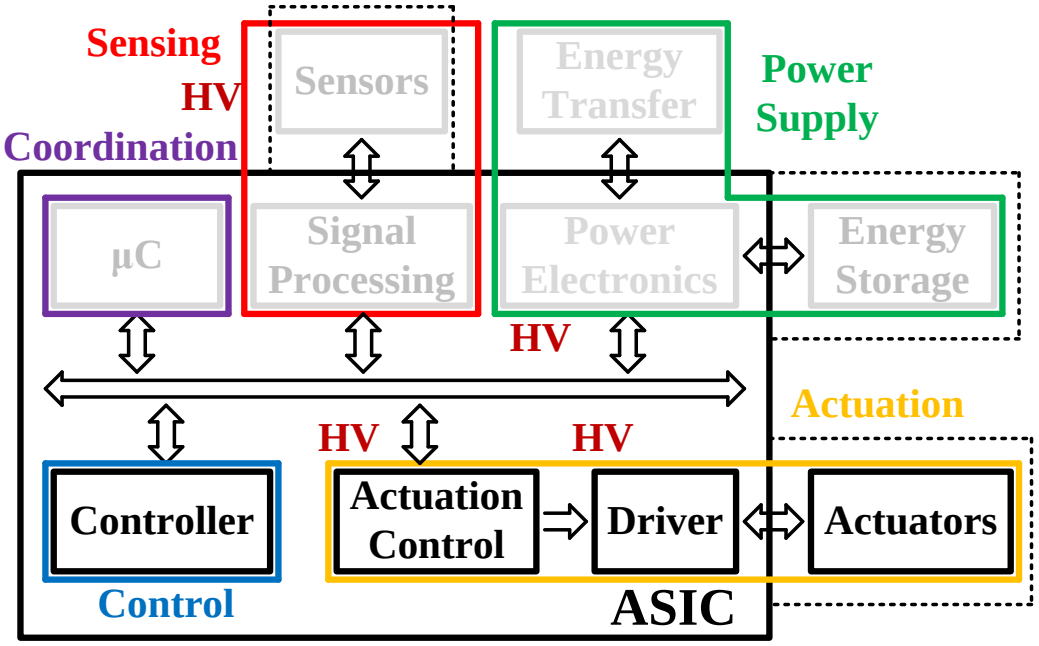


Figure 1.12. One possible implementation of the proposed platform diagram including coordination and four major functions required for an autonomous robot, control, power supply, sensing, and actuation. The black function blocks are the major focus of this thesis. HV device implemented in this work can be applied to functional blocks labeled HV. Other light grey blocks were outside the scope of this research. Extended dashed line indicates that some options of these function blocks can be implemented on the same chip.

How our proposed approaches overcome the challenges will be addressed, in order, in the following paragraphs. To achieve compact integration and high performance, a full-custom ASIC platform presents many advantages over COTS components. Full-custom circuits can be designed to match specific demands and have better size and power performance compared to COTS components which have greater overhead. Therefore, fast and power-efficient processing as well as compact integration can be achieved. Integration of multiple functions on an ASIC greatly reduces the difficulties to assemble the robot. It also reduces the noise and power leakage introduced by wire bonds or PCB routing used for the integration of multiple

components. A smaller number of connections between components also has the potential to increase robustness and reliability of the robots.

HV devices compatible with standard CMOS processes are essential for the integration of HV circuits in a single chip. These HV circuits appear in the driver block and possibly in the power electronics block. Although there are commercially available HV processes which have both HV and low voltage circuits fabricated on the same chip, these processes are usually optimized for HV devices and degrade the performance of low voltage circuits [47]. Since HV circuits only occupy a small portion of the whole system, it is preferable to build HV circuits using a standard CMOS process instead of using a specialized HV process. With this compatibility, the size of the robots can be shrunk further.

To ease the high demands on μC as described in the second challenge, each component of the decentralized system has its own function specific mixed-signal circuits to share the load. As a result, real-time processing can be accomplished by this parallel and custom-design approach. Previous designs converted the analog signal from sensors or antennas into a digital signal at the front-end and rely on digital circuits (mostly μC) to handle most of the computations. We propose to keep as much computation and processing in the analog domain as possible because, in most cases when the processing can be performed in analog domain, analog circuits beat the digital circuits by at least one order of magnitude in terms of area and power to perform the same task [48]. The μC in this system is only responsible for simple computation and coordination of other components and can stay in low power idle mode or sleep mode for most of the time.

Tiny mechanisms and actuators will be custom fabricated to deal with the unavailability of COTs components at this scale. The mechanisms will be fabricated on the ASIC platform directly; this configuration has the potential to achieve the minimum size of the robots. The actuators need to have enough force and displacement. They can be fabricated using a microelectromechanical systems (MEMS) process. The process has to be carefully designed so that the ASIC is not physically damaged or characteristically altered during post-fabrication of the actuators. The operating condition of the actuators must be compatible with the ASIC, for example current density, voltage, and temperature.

1.5 Organization

Under the proposed platform this dissertation focused on the parts that have been seldom addressed in the robotics research. Our research was divided into six parts, each part is critical to the success of the decentralized ASIC platform and is summarized as follows. First, we describe ultralow-frequency actuation control signal generation that is useful for **actuation** and **control**. Second, we describe programmability and analog storage of mixed-signal circuits that are useful for **actuation** and **control**. Third, we describe CMOS compatible HV devices for driver and power electronics that are useful for **actuation** and **power**. Fourth, we describe tiny actuators that are useful for **actuation**. Fifth, we describe low-power computation that is useful for **control** and potentially for **sensing**. Finally, we describe a clock generator for dynamic power management that is useful for **control** and coordination. The first four parts were integrated and presented as an actuation system.

This work is presented in six chapters. Each chapter describes one part of the platform mentioned in the last paragraph. The work underlying this dissertation has been reported in a few publications. In all papers except the 2013 SPIE paper, I am the first author, and I designed and performed all the experiments; Dr. Pamela Abshire is the sole co-author and served as the editor of the final manuscript in addition to providing guidance for the experimental and theoretical investigations. A summary of the chapters is as follows:

Chapter 2 describes the design of the actuation signal generator and its optimization. The contents were published in Proceedings of 2012 IEEE Midwest Symposium on Circuits and Systems [49]. The first version of actuation signal generator was initiated as a final project in a mixed signal VLSI design class at the University of Maryland (instructed by Dr. Abshire) by undergraduate students Robert Bailey, Benjamin Chang, Angel Diaz, and Daniel Sher. Mr. John Turner continued the circuit design and submitted the first version of design for fabrication. However, the chip was not able to achieve the low frequency we desired. Part of the pad frame layout drawn by Mr. Turner was reused for all the later unpackaged chip design. None of the circuits in the first version were reused.

Chapter 3 discusses non-volatile storage and programming for mixed-signal circuits. We used the actuation signal generator as an example. This circuit design and simulation results appeared in Proceedings of 2012 IEEE Midwest Symposium on Circuits and Systems [50]. The measurement results and the experiments showing the controller driving the actuators in Chapter 5 is being prepared for submission to IEEE Transactions on Circuits and Systems II.

Chapter 4 provides the design and optimization of HV N-Type Metal-Oxide-Semiconductor (NMOS) in a standard CMOS process as well as its characterization. This chapter features an article published in 2013 IEEE Sensors Journal [51] and another article published in Proceedings of 2012 IEEE Sensors Conference [52]. Dr. Marc Dandin assisted on writing a testing equipment control program in Matlab for a Keithley source measurement unit (SMU).

Chapter 5 discusses the design, optimization, and fabrication of the tiny mechanisms. Ms. Deepa Sritharan helped on some of the fabrication process and did all of polydimethylsiloxane (PDMS) curing for packing. Dr. Bavani Balakrisnan designed an initial fabrication procedure and some masks for photolithography. Her choice of SU-8 as one layer of the thermal actuators remained unchanged in the final process. Her Matlab program, which supports multi-layer actuator simulation, was used for optimization of the actuators. Dr. Elisabeth Smela, Dr. Pamela Abshire, Ms. Sritharan, and I had regular meetings to discuss the fabrication so part of the design was a group effort. I did most of the optimization, fabrication, and testing. The experiments showing the controller driving the actuators in this chapter and the measurement results in Chapter 3 is being prepared for submission to IEEE Transactions on Circuits and Systems II.

Chapter 6 presents a low-power computational approach for designed intelligence. We used odometry as an implementation example. This odometry circuit was published in Proceedings of 2013 SPIE8725 [48]. Mr. Michael Kuhlman is the first author of the paper, defined the problem, designed the preliminary mixed-signal circuits, and did the error analysis. He initiated this circuit as his final project in a

graduate circuit class instructed by Dr. Horiuchi at the University of Maryland. I modified the circuit architecture, improved the performance (especially reduced error to an acceptable level), implemented a digital counterpart for comparison, and did data analysis to understand the experiments. Dr. Abshire did the final editing and provided suggestions for the experiment. An improved sine shaper circuit was also presented which was later published in Proceedings of 2014 New Circuit and Systems Conference [53].

Chapter 7 discusses a low-jitter oscillator similar to the actuation signal generator. This circuit provides a clock signal and exhibits a tradeoff between clock quality and power; it is suitable for dynamic system management for a power constrained system like ours. This work was accepted for publication in IEEE Transactions on Very Large Scale Integration Systems in February, 2016 [54].

Chapter 8 concludes this dissertation and discusses future directions.

1.6 Contributions

Chapter 2:

- Utilized a frequency division technique to achieve ultralow frequency oscillation, without impractically large chip area.
- Developed a design approach for the ultralow frequency motion controller to guide design choices to optimize area, power, and phase noise. The choices are guided by a simulation-based capacitance model that elucidates the constraints among the underlying design parameters. Different performance metrics cab be selectively weighted in the optimization to meet specifications.

Chapter 3:

- Designed and implemented a programmable floating-gate phase-locked loop. The circuit is able to store the frequency information non-volatilely and can be programmed automatically. This design is the first demonstration incorporating a floating gate structure in a phase-locked loop.
- Integrated the floating-gate phase-locked loop controller on a legged robot to demonstrate leg control.

Chapter 4:

- Demonstrated the highest reported breakdown voltage (> 40 V, 10 V higher than the previous 30 V [51]) for NMOS transistors in a standard 0.5 μm CMOS process without modifying any standard masks
- Identified the breakdown mechanisms for the devices tested: gate-induced drain leakage breakdown and avalanche breakdown.
- Compared drain-centered and source-centered circular structures for HV NMOS transistors directly on the same chip, showing that the optimal structure is drain-centered.

Chapter 5:

- Demonstrated the utility of a cantilever simulation tool to optimize the blocking force and bending angle of thermal actuators. The actuators were capable of lifting three times the weight of the robot. Found compatible materials for the sacrificial and electrode layers during fabrication of the actuators.
- Demonstrated leg control on the thermal actuators with control signals provided by the actuation controller chip.

Chapter 6:

- Developed a mixed-signal computational architecture for odometry that achieved 60x and 330x energy efficiencies compared to digital ASIC and μC implementations, respectively [48].
- Developed a design methodology for sine shapers to achieve arbitrary input voltage range and its mapping to angles

Chapter 7:

- Proposed a frequency-boost jitter reduction technique for ring oscillators and achieved the best phase noise figure of merit at one operating frequency and competitive performance at other frequencies compared with other reported designs [54]. This figure of merit is a commonly used metric to compare across oscillators operating at different settings and fabricated using different technologies.

Chapter 2: Ultralow Frequency Actuation Control for Legged Chip*

2.1 Control Signal for Robot Motion

Researchers have used bio-inspired control like central pattern generator (CPG) [55-57] to generate complicated gait; CPG has multiple coupled processes that produce rhythmic outputs even without actuator or sensory feedback. However, CPG is unnecessary for tiny robots that cannot perform complicated gait. For tiny robots, motion control often relies on a periodic signal to drive actuators, for example Kilobot, HAMR^3, RoACH, MiCRoN, I-SWARM, Walking Robot, and Silicon Robot. The signal is usually a square wave with tunable frequency and duty cycle to control motion of the robot. Frequencies of the periodical signal ranges from less than 1 Hz to several KHz depending on the mechanical response of the actuators. Legged robot (HAMR^3, RoACH, Walking Silicon, and Silicon Robot) used two phased periodic signal to achieve alternating gait by dividing the legs into two groups. Figure 2.1 illustrates the use of two phased square wave signals to control leg actuation in a legged chip, a robot that uses a CMOS chip as its substrate and micro-electro-mechanical systems (MEMS) actuators as its legs. This ciliary moving mechanism was shown in multiple works [58-60].

* Most of the material in this chapter was originally published as " T.-H. Lee and P. A. Abshire, "Design methodology for a low-frequency current-starved voltage-controlled oscillator with a frequency divider," in Proc. *IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2012, pp. 646-649." © 2012 IEEE.

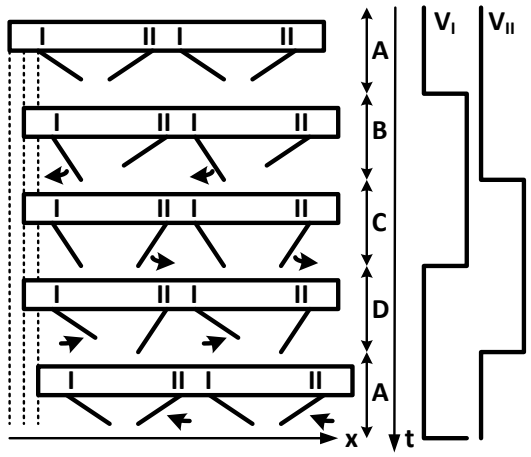


Figure 2.1. Gait illustration and associated 4-state signals (A, B, C, and D) [58-60]. Actuators and control signals are divided into two groups (I and II).

Physical characteristics of micro-scale actuators such as resistance and capacitance are not expected to be well controlled, so it will be hard to set a predetermined actuation frequency for efficient motion control. In addition, it is highly desirable for the control signal to be adaptable to different kinds of actuators and systems, and to be able to compensate for fabrication tolerances and assembly variations between individual robots. Therefore, the output frequency has to be tunable over a wide range of frequencies (0.1 Hz to 10 Hz).

The first version of this two phased control generator was initiated as a final project in a mixed signal VLSI design class at the University of Maryland (instructed by Dr. Abshire) by undergraduate students Robert Bailey, Benjamin Chang, Angel Diaz, and Daniel Sher. Mr. John Turner continued the circuit design and submitted the first version of design for fabrication. However, the chip was not able to achieve the low frequency we desired. Part of the pad frame layout drawn by Mr. Turner has

been reused for all the unpackaged chip design here. None of the circuits in the first version were reused.

2.2 Design of Control Signal Generator

2.2.1 Architecture for Low-Frequency Oscillator

We chose voltage-controlled oscillators (VCOs) as the core circuit to implement the square wave signal generator with programmable frequency. Whereas most design efforts are directed toward high frequency oscillators, low frequency oscillators (<1Hz to several Hz) present unique design challenges and have many interesting applications including not only motion control for tiny robots but also biomedical, biological [61, 62], geophysical, audio, control [63].

CMOS oscillators can easily go up to hundreds of MHz or GHz [64, 65]. Therefore, means are required to bridge this frequency gap between the CMOS and the actuation frequency. The most straightforward way to implement a low frequency oscillator is by scaling up the size of components in the oscillator. However, increasing the component size only decreases the frequency linearly and leads to impractical design parameters. This can be explained using an example of a relaxation oscillator which has a current source to either charge or discharge a capacitor. The current direction is controlled by a controller with two thresholds. When the capacitor is being charged and its voltage reaches the upper threshold, the controller switches the current direction and vice versa. This yields a triangle waveform bouncing between two thresholds. One setting to achieve 1 Hz oscillation is to have a capacitor of 1 nF, a current source of 2 nA, and a triangle swing of 1 volt.

A 1 nF capacitor implemented with a poly-poly capacitor would occupy more than 1 mm^2 chip area which is impractical. A current source of 2 nA is at thermal noise current level and we lost linear control at this scale of current. Therefore, linear scale does not work for our application. Here we introduced a frequency divider (FD) in order to achieve low frequency operation. By cascading multiple divide-by-two (DB2) circuits the frequency decreases exponentially for a linear increase in area. However, frequency dividers increase the total power consumption and introduce an additional parameter (frequency division) and tradeoffs into the design process. The architecture of the control signal generator is shown in Figure 2.2. The function of the D flip-flop (DFF) at the output of FD will be discussed in more details in section 2.2.3.

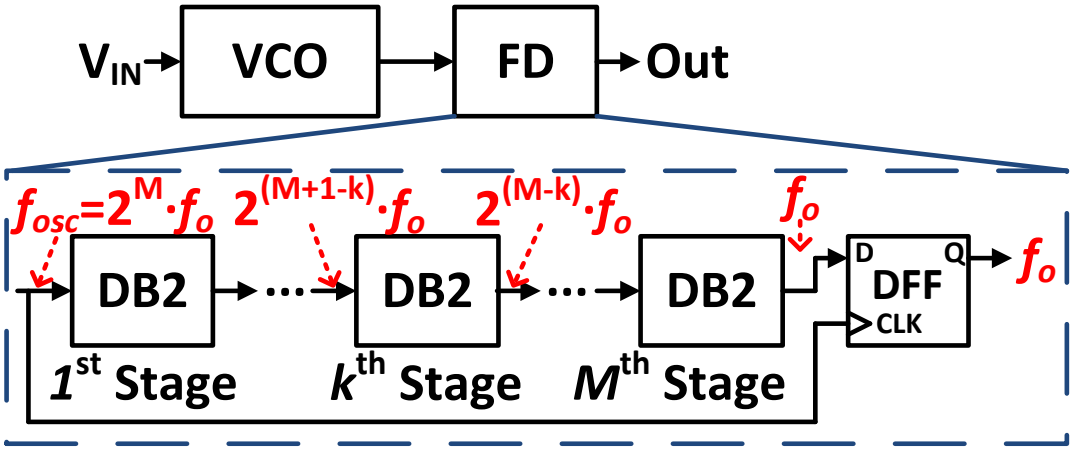


Figure 2.2. Block diagram of the low frequency oscillator which consists of a VCO and a FD. Oscillation frequency of VCO is f_{osc} and desired oscillator output frequency is f_o.

2.2.2 Current Starved VCO Design

In comparison with other VCOs such as LC-tank oscillators and source-coupled oscillators, current-starved (CS) VCOs (CSVCOs) exhibit the widest tuning

range which is favorable for our application and at the same time achieve a practical balance among area, power, and phase noise as reported by Hsieh et al. [66] and Miyazaki et al. [67]. Therefore, this work focuses on the optimization of a CSVCO based square wave signal generator with programmable frequency.

This design implemented an *N*-stage CSVCO shown in Figure 2.3. Two inverters not shown in the figure were added to the output to sharpen the edges of output signal. The output inverters are important especially when the frequency is slow. The VCO gain was linearized with a large resistor R (100 KΩ). Since the desired frequency is low, a current divider stage, M_3 and M_4, was added to reduce the current I_{CSI} mirrored to the CS inverters (CSIs) by making the W/L ratio of M_3 smaller than M_2. The feedback capacitor C_f was included to enhance the effective capacitance of the inverter. Although the capacitance per area of available capacitors are usually several times smaller than that of gate oxide which the transistors can provide, some other effects take place and make C_f useful. Oscillation frequency in this structure mainly depends on I_{CSI} (controlled by V_{IN}), number of CSI stages, N, and size of CSI, W_{Ip}/L_{Ip} and W_{In}/L_{In}. Moreover, we will let $W_{In}=W_{Ip}=L_{Ip}=L_{In}=W$ and this will be explained in section 2.3.

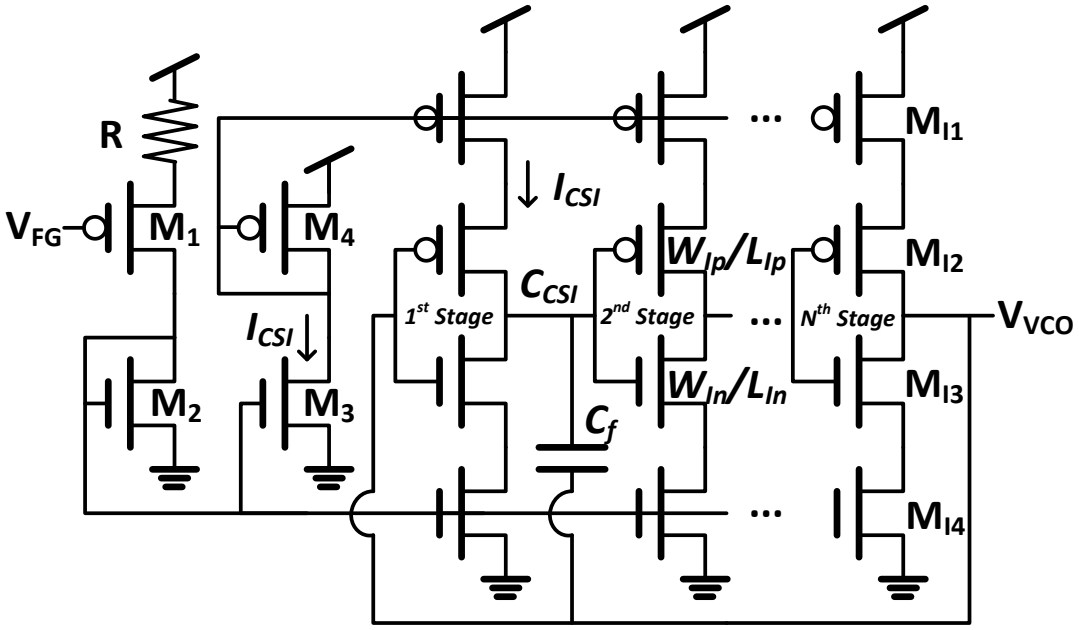


Figure 2.3. CSVCO with a linearizing resistor R, one current divider stage M_3 and M_4, and one feedback capacitor with capacitance C_f.

2.2.3 Frequency Divider Design

The FD circuits were implemented by cascading *M* stages of DB2 circuits which yields an internal oscillation frequency of $f_{osc}=2^M \cdot f_o$. DB2 was implemented by connecting *Qэ* output of a DFF to its *D* input and making CLK input and *Q* output as the input and output of the DB2 respectively. This configuration can be seen in Figure 2.4. The output port, Q, changes state when a positive edge is coming in the clock input port. Therefore, the output stays constant for a whole input clock cycle and, then, changes state at the next cycle. The frequency division by 2 is achieved hence.

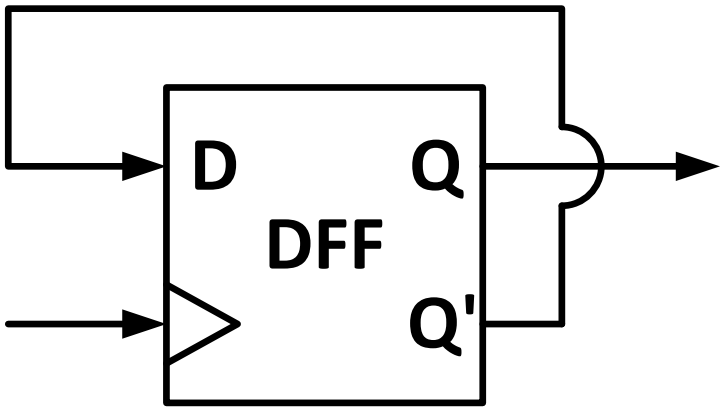


Figure 2.4. Configuration of the DFF based DB2 circuit.

Two kinds of implementations of DFF were introduced: transmission-gate and dynamic as show in Figure 2.5 (a) and (b). The dynamic DFF has a few advantages over the transmission gate DFF. Dynamic DFF has 11 transistors, 7 less than transmission gate DFF, and, thus, occupies less area and has less leakage current. The total capacitance of internal switching nodes for dynamic DFF is less than transmission gate DFF, and, thus, consumes less dynamic power. Although the dynamic DFF shows both lower area and lower power consumption, it only works when input signal changes often enough because its function relies on temporary charge storage in capacitance and could be destroyed by leakage. In our simulation, dynamic DB2 (DDB2) works for input frequency higher than 1 KHz. Moreover, we set the output frequency, f_o, to be 1 Hz, a simulation determined optimal actuation frequency for the MEMS legs that will be used as in Figure 2.1. Under this condition, DDB2 was only used for the first (M-10) stages when M is larger than ten or otherwise all stages would be transmission gate DB2 (TGDB2).

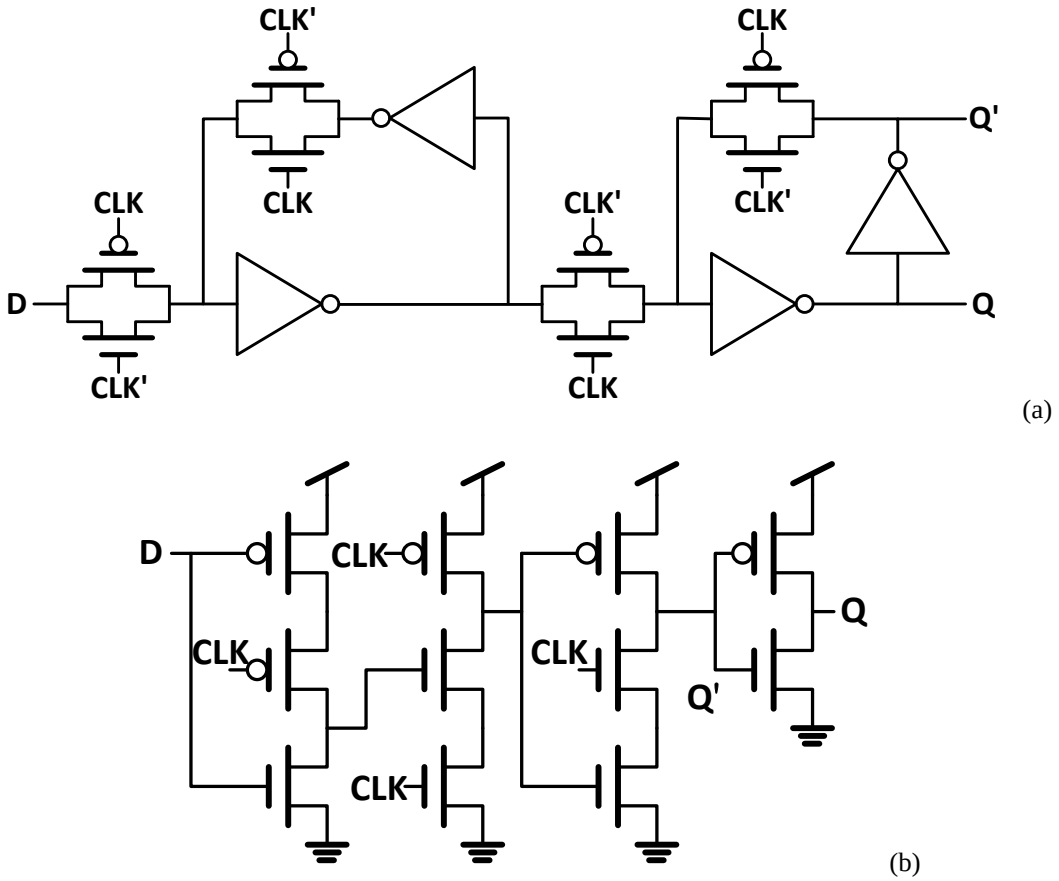


Figure 2.5. (a) Transmission gate based DFF. The inverter for clock inversion is not shown. (b) True single phase dynamic DFF.

The function of the DFF at the output (see Figure 2.2) is to eliminate the timing error introduced by the DB2 circuits. The timing jitter of the CSVCO for each cycle can be considered as a random variable Δt_1. Without this DFF the timing jitter at the oscillator output is the sum of multiple random variables $\Delta t_1 + M \cdot \Delta t_2$ assuming random variables are stationary, where Δt_2 is the timing jitter of each DB2. However, with the DFF the output would only change at the rising edge of the signal generated by the CSVCO. The timing error consequently only depends on the CSVCO and becomes Δt_1.

2.3 Oscillator Optimization

Simulation of oscillators is usually computationally intensive and slow [68-70]. The simulation takes several cycles for the oscillator to achieve steady state and it will require another tens of cycles to get statistical result like frequency and phase noise. It also requires fine timing step (one hundredth of the oscillation period) to capture the details of the huge transient. Therefore, it is desirable to accomplish as much of the design as possible in an analytical or model-based framework. Most studies on optimization of oscillators, for example Leung [71] and Abidi [72], have focused on reducing timing jitter or phase noise. In generation of control signals for tiny robots, timing jitter or phase noise of the signal is not as critical as area and power constrains because the frequency stability is not strictly important. However, it would still be a benefit to consider the phase noise to prevent unstable actuation from happening. We propose a design methodology which considers area and power as well as phase noise. A model for effective capacitance of a CSVCO was developed to serve as a bridge between the design parameters and these metrics when VCO current and frequency are predetermined.

In our case, the desired output frequency is 1 Hz. I_{CSI} was chosen to be 290 nA so that the bias current only consumes 1 μW with a 3.3 V supply. Moreover, for best tuning range, V_{IN} should be in the middle of the linear VCO gain region (~1.2 V). Then, the current bias of CSVCO (M_1 to M_4 and the resistor) was designed accordingly.

From section 2.2, there are four remaining design parameters *N*, *M*, *W*, and C_f. If performance metrics, area, power, and phase noise, can each be directly related to

the four design parameters and the combination (product for example) of the metrics forms a cost function, it becomes a four-dimensional optimization problem. Since *N* and *M* are integers and have a practical range, it is actually a two-dimensional optimization problem given one *N* and *M* combination; we find optimal *W* and C_f given *N* and *M* and then find the best set among all possible *N* and *M* combinations. However, *W* and C_f cannot be chosen arbitrarily given *N*, *M*, and an output frequency f_o. They have to follow a relationship so that the output frequency is as desired. We propose a simple and accurate model for effective capacitance of a CSVCO which can be used to avoid exhaustive simulation. The model relates C_f to *W* and eliminates one dimension for the problem while providing a good approximation of the oscillation frequency. Given *N* and *M*, the optimization becomes a one-dimensional search over *W* using approximations of area, power, and phase noise that only depend on these design parameters.

2.3.1 Effective Capacitance Model and Frequency Estimation

The oscillation frequency f_{osc} of a CSVCO without the feedback capacitor can be approximated as suggested by Baker [73]

$$f_{osc} = I_{CSI} \Big/ V_{dd} \cdot C_{eff} \tag{2.1}$$

where C_{eff} is the sum of the effective capacitance at the output of each CSI and is given by

$$C_{eff} = N \cdot 5 \Big/ 2 \cdot C'_{ox} (W_{Ip} L_{Ip} + W_{In} L_{In}) \tag{2.2}$$

where *N* is the number of CSI stages and C_{ox} is gate oxide capacitance per unit area. From this analysis we infer that the function of M_{I2} and M_{I3} is to provide capacitance,

and thus, these transistors were designed to be square and the same size to minimize area ($W_{In}=W_{Ip}=L_{Ip}=L_{In}=W$). Because I_{CSI} and V_{dd} are fixed, we focused on estimating C_{eff} in order to better approximate f_{osc}. We modeled C_{eff} as the weighted sum of capacitance contributed by the transistors and the feedback capacitor:

$$C_{eff}(N,W,C_f) = K_1(W)\cdot C_{CSI}(N,W) + K_2(W)\cdot C_f \tag{2.3}$$

where $K_1(W)$ and $K_2(W)$ are regression parameters. $C_{CSI}(N,W)$ is the effective capacitance contributed by the transistors:

$$C_{CSI}(N,W) = [N\cdot 5\cdot C'_{ox}\cdot W^2 + C_L] \tag{2.4}$$

where C_L is the load capacitance of the CSVCO, input capacitance of an inverter in this case.

To determine K_1 and K_2, we first used transient simulations with different combinations of N, W, and C_f to find f_{osc}. Then, $\tilde{C}_{eff}$ is calculated as $I_{CSI}/V_{dd}\bullet f_{osc}$. For some simulations C_f was set to zero and K_1 was extracted from Equation 2.3 by

$$\tilde{K}_1(N,W) = \tilde{C}_{eff} / C_{CSI}(N,W). \tag{2.5}$$

Minimum regression error of Equation 2.5 was obtained for

$$K_1(W) = \frac{0.258\cdot W^2 + 0.357\cdot W + 6.09}{0.49\cdot W^2 - 0.0897\cdot W + 10.2}. \tag{2.6}$$

The worst case error is 0.3% and the N dependence was taken away because K_1 stays relatively constant while varying N. The regression result is shown in Figure 2.6. Next, $\underline{K_2}$ was extracted from Equation 2.3 again by replacing K_1 with Equation 2.6 and get simulated $\tilde{C}_{eff}$ with nonzero C_f as

$$\tilde{K}_2(N,W,C_f) = (\tilde{C}_{eff} - K_1(W)\cdot C_{CSI}(N,W)) / C_f. \tag{2.7}$$

For large W, K_2 is dominated by the residual error of K_1. Thus, $\tilde{K}_2$ values were held fixed once Equation 2.7 times C_f is less than 0.5 % of $K_1(W) \bullet C_{CSI}$. Minimum regression error of K_2 was given by

$$K_2(W) = \frac{0.00273 \cdot W^3 + 108.63 \cdot W^2 + 3.61 \cdot W + 228.63}{28.33 \cdot W^2 + 185.88 \cdot W - 366.26}. \quad (2.8)$$

The N and C_f dependences were also removed for simplicity while enough accuracy can still be maintained. The resulting $K_2(W)$ is shown in Figure 2.7, and the percent error in $C_{eff}(N,W,C_f)$ compared to simulated C_{eff} is shown in Figure 2.8. Errors are less than 10 % for most cases. Large errors arise for small W (< 15 μm) and particularly for large C_f (red dashed line with right-pointing triangles, solid orange lines with left-pointing triangles, and black dashed line with solid circles), a subset of configuration space where optimal results rarely occur for our specifications.

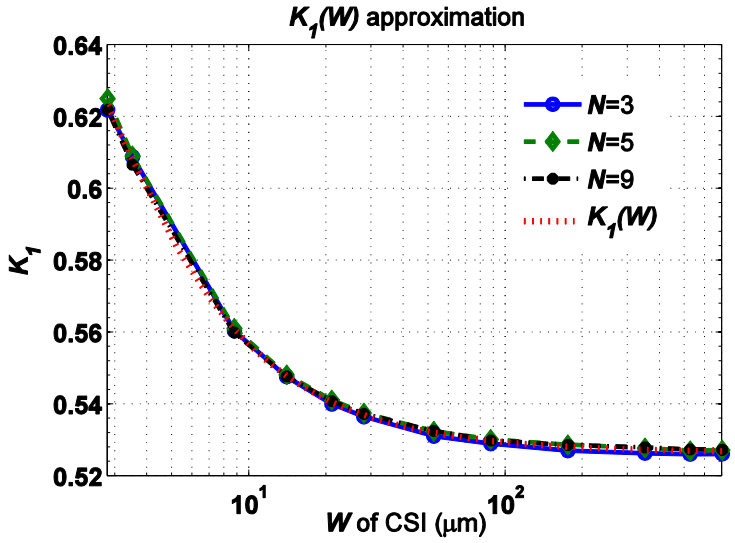


Figure 2.6. K_1 regression extracted from simulations using Equation 2.5 with zero C_f and N of 3, 5, and 9 versus different W. N dependence is negligible. Dots are data points and all lines except the red one are connecting lines. Red dashed line is a plot of Equation 2.6.

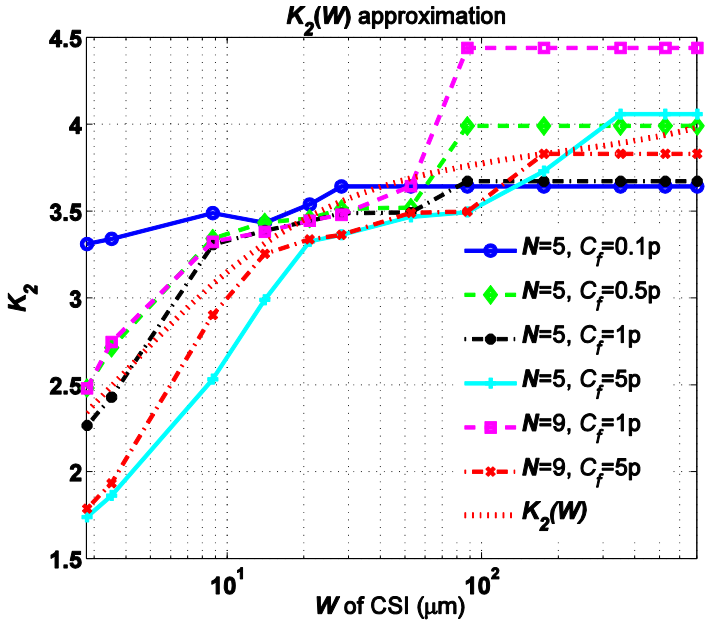


Figure 2.7. K_2 regression extracted from simulations using Equation 2.7 with nonzero C_f and N of 5 and 9 versus different W. N and C_f dependence are ignored while enough accuracy is still maintained. Dots are data points and all lines except the red one are connecting lines. Red dashed line is a plot of Equation 2.8.

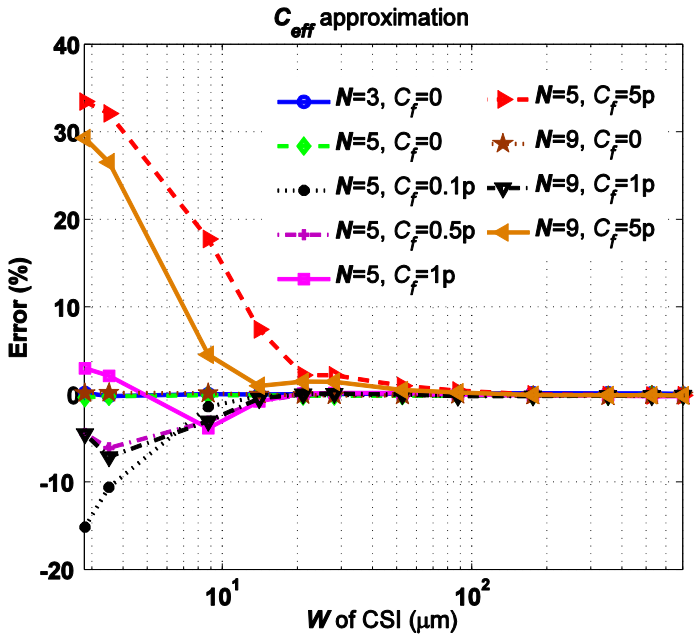


Figure 2.8. Error of our C_{eff} model versus different W compared to simulated effective capacitance. Dots are data points and all lines connecting lines.

Note that $K_1(W)$ is less than one; this occurs because Equation 2.2 overestimates the effective capacitance assuming that the transistors remain in saturation during the entire operation. However, the transistors can operate in triode region or even turn off and, thus, provide much less capacitance. Actually n-type metal-oxide-semiconductor (NMOS) almost turns off while p-type metal-oxide-semiconductor (PMOS) in the same CSI is on, and vice versa. All these effects result in that K_1 is close to half. $K_2(W)$ is somewhat larger than one as a result of: 1) Miller effect across capacitor C_f; and 2) sharing of C_f by the first and last CSIs. When W is small, K_2 increases with W because the rail-to-rail voltage range is not fully utilized by the CSVCO and voltage swing of the CSINV increases as W.

Finding $K_1(W)$ and $K_2(W)$ allows us to predict the oscillation frequency based on Equation 2.1 and Equation 2.3. More importantly this model describes analytically the relationship between W and C_f (Equation 2.3) given f_o, N, and M. It converts the optimization into one dimension and makes the optimization become possible.

2.3.2 Area Approximation

The areas of CSVCO and FD circuits were approximated using extracted data from layout design by taking advantage of the regularity of the oscillator structure. The total area is

$$A_{tot} = N \cdot A_{CSI} + A_{OV} + C_f / C'_{pp} + A_{FD} \tag{2.9}$$

where A_{CSI} is area of a single CSI, A_{OV} is the area of the resistor, transistors M_3 and M_4, and two output inverters of the CSVCO, C'_{pp} is the capacitance per unit area of the poly-poly capacitor, and A_{FD} is area of the frequency divider. The first three terms represent the area of the CSVCO. The area of a single CSI stage is A_{CSI} =

$(2W+W_{OV})\bullet(W+L_{OV})$ where W_{OV} and L_{OV} are the width and length overhead for routing and substrate contacts as well as M_{I1} and M_{I4} as shown illustrated in Figure 2.9. This area approximation for CSVCO is accurate because the circuit topology is regular. The FD area is

$$A_{FD}=\begin{cases} M\cdot A_{TGDB2} & \text{if } M\le 10 \\ 10\cdot A_{TGDB2}+(M-10)\cdot A_{DDB2} & \text{else} \end{cases} \tag{2.10}$$

where A_{TGDB2} and A_{DDB2} are the areas of a TGDB2 and a DDB2 respectively. The condition of M was determined as the discussion in section 2.2.3. Area of M_1 and M_2 was not considered because they occupy a small fraction of the total area and are not needed in all low-frequency CSVCOs.

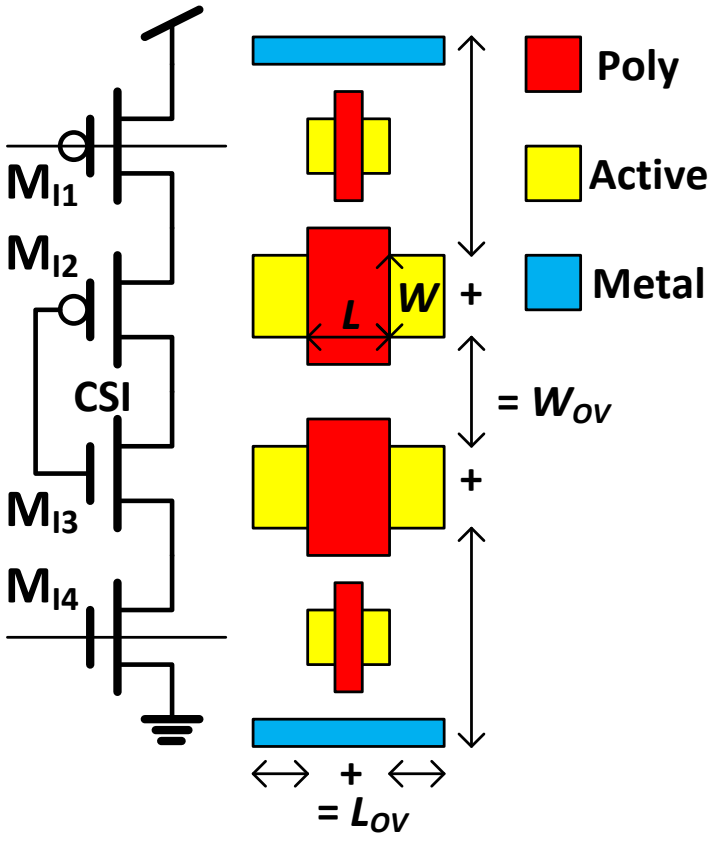


Figure 2.9. Area approximation of a CSINV. The total width (vertical direction) and length (horizontal direction) of the layout are approximately $2W+W_{OV}$ and $L+L_{OV}$ respectively.

2.3.3 Power Approximation

The average dynamic power dissipated by all CSIs is $V_{dd} \bullet I_{CSI}$ [73]. The average power of the CSVCO, P_{CSVCO}, is assumed to be $2 \bullet V_{dd} \bullet I_{CSI}$ (including current of M_3 and M_4 but not M_1 and M_2) plus the power of inverters at the CSVCO output, P_L. The power of a DB2 circuit was fit to the simulated power with a 3.3 V supply by linear regression. Power of TGDB2 and DDB2, P_{TGDB2} and P_{DDB2}, became $L_{11} \bullet f_{in} + L_{12}$ and $L_{21} \bullet f_{in} + L_{22}$ respectively, where f_{in} is the input frequency of the DB2 circuits and L_{11}, L_{12}, L_{21}, and L_{22}, are 8.74e-13, 7.39e-11, 4.70e-13, and 3.26e-10 respectively. The result is shown in Figure 2.10. For P_{TGDB2}, the regression result is slightly off at low frequency input but it is at 10^{-10} level and can be ignored. The power of the frequency divider P_{FD} is

$$P_{FD} = \begin{cases} (L_{11} \cdot 2 f_o + L_{12}) + \cdots + (L_{11} \cdot 2^M f_o + L_{12}) \\ \quad = L_{11} \cdot (2^{M+1} - 2) \cdot f_o + M \cdot L_{12} & \text{if } M \le 10 \\ L_{21} \cdot (2^{M+1} - 2^{11}) \cdot f_o + (M - 10) \cdot L_{22} + \\ \quad L_{11} \cdot (2^{11} - 2) \cdot f_o + 10 \cdot L_{12} & \text{if } M > 10 \end{cases} . \qquad (2.11)$$

P_L was calculated using the same regression method and it becomes 2.64e-13$\bullet f_{in}$+2.39e-12.

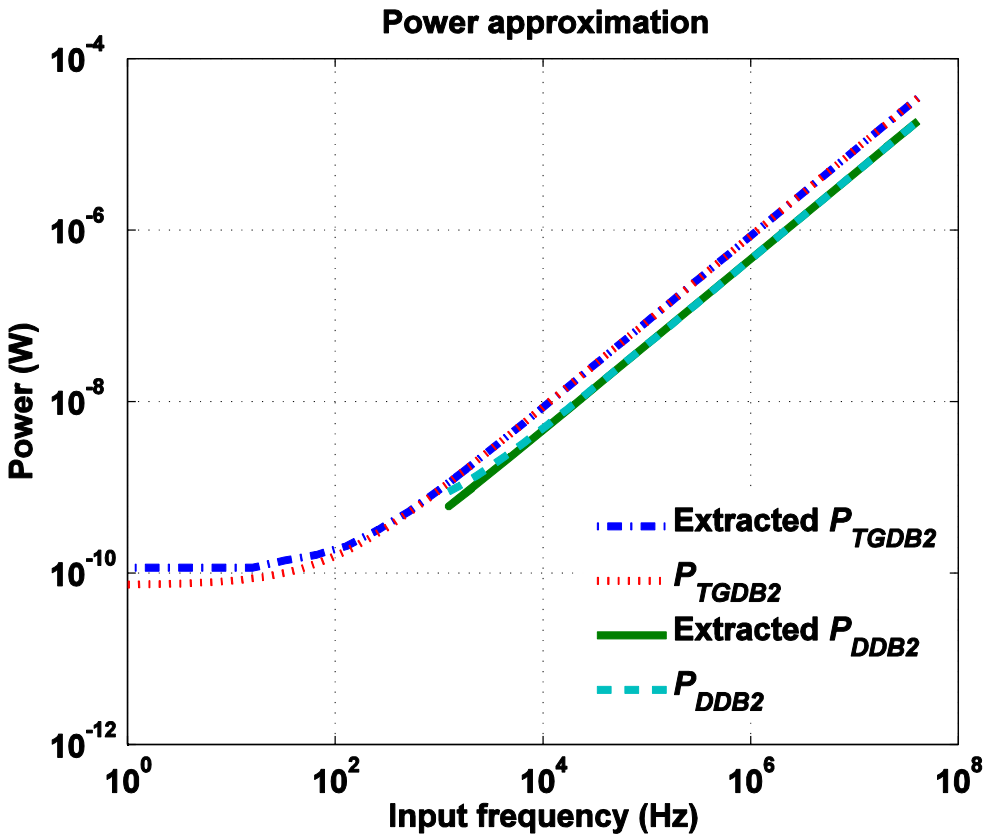


Figure 2.10. Power approximation for TGDB2 and DDB2. The simulation of DDB2 started from an input frequency of 1 KHz.

Finally the power of the whole circuit, P_{tot}, is given by

$$P_{tot} = W_{COR} \cdot P_{CSVCO} + P_L + P_{FD} \tag{2.12}$$

where W_{COR} is a correction factor to account for power other than dynamic power, for example power consumed by leakage and short circuit current. Under the assumption that dynamic power dissipates 80% of the total, W_{COR} is 1.25.

2.3.4 Phase Noise Model

In this work we assumed that the primary noise source is flicker noise (1/*f* noise) and timing noise in the CSVCO dominates. We adopt Abidi's model of phase noise model for an inverter-based ring oscillator [72]. The power spectral density is

$$N_{CSVCO}(f) = \frac{C'_{ox}}{8 \cdot N \cdot I_{CSI}} \left(\frac{\mu_N K_{fN}}{L_{In}^2} + \frac{\mu_P K_{fP}}{L_{Ip}^2} \right) \left(\frac{f_{osc}^2}{f^3} \right) = \left[\frac{C'_{ox} \cdot (\mu_N K_{fN} + \mu_P K_{fP}) \cdot f_o^2}{8 \cdot I_{CSI} \cdot f^3} \right] \cdot \left(\frac{2^{2M}}{N \cdot W^2} \right) \quad (2.13)$$

where f is the frequency offset, μ_N and μ_P are electron and hole mobility, and K_{fN} and K_{fP} are empirical coefficients for flicker noise in NMOS and PMOS. The first term (the square brackets) is constant if we are looking at a specific f, and the second term includes all dependence on design parameters. Because phase noise is typically expressed in units of dB, its optimization metric, N_{tot}, is given by

$$N_{tot} = 10 \log \left(\frac{N_{Max}}{N} \cdot 2^{2(M - M_{Min})} \cdot \left(\frac{W_{Max}}{W} \right)^2 \right) \quad (2.14)$$

where N_{Max} and W_{Max} are the largest possible number for N and W respectively, and M_{Min} is the smallest possible number for M. These factors are used to make N_{tot} positive at any time.

2.3.5 Design Flow

The optimal design parameters can be found by minimizing the product of area, power, and noise, *APNP*,

$$APNP(W_A, W_P, W_N) = A_{tot}^{W_A} \cdot P_{tot}^{W_P} \cdot N_{tot}^{W_N} \quad (2.15)$$

where W_A, W_P, and W_N are weighting factors for area, power, and phase noise respectively. Figure 2.11 illustrates the design flow. Parallelogram, rectangle, and diamond blocks indicate setting parameters, operations (simulation and calculation), and making decisions, respectively.

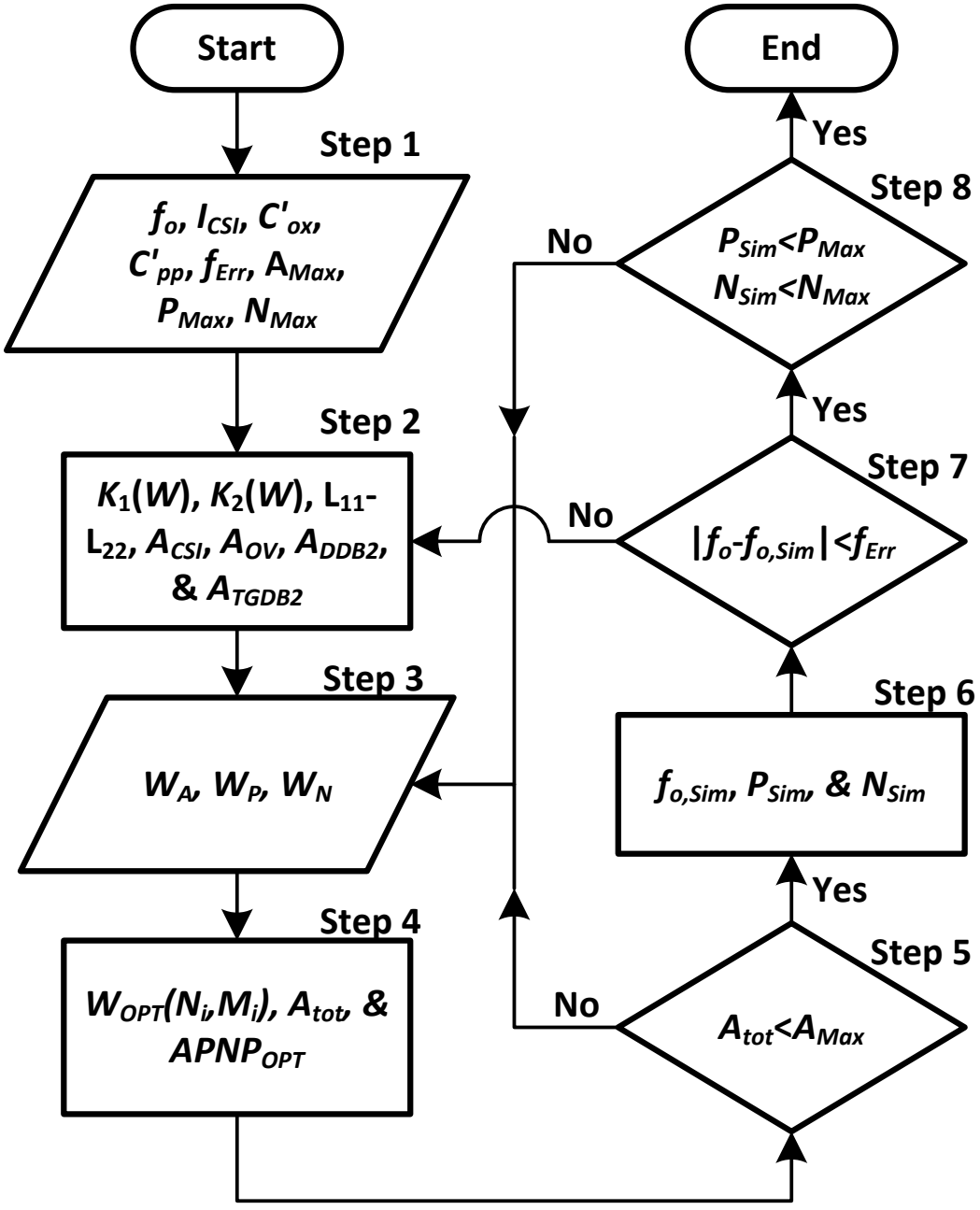


Figure 2.11. Design flow for the CSVCO.

The design constraints and process parameters are specified in Step 1. Step 2 obtains regression parameters based on simulation results and approximates the area of individual components from layout or designers' experience. The weighting factors are adjusted in Step 3. In Step 4 we select a possible set of N and M (N_i and M_i) where

optimal W for each combination ($W_{OPT}(N_i, M_i)$) is found using a Matlab function, *fminbnd*, and the upper and lower bounds are set to cover all physically possible W. Corresponding optimal C_f was found using Equation 2.3. From all these solutions optimal $APNP$ ($APNP_{OPT}$) and A_{tot} are found and the optimal design parameters are recorded. Step 5 compares the A_{tot} with the area constraint, A_{Max}. If A_{tot} exceeds the constraint, the process returns to Step 3 and puts more emphasis on area (increase W_A or decrease W_P or W_N or both). Then, the candidate design is simulated to obtain the output frequency ($f_{o,Sim}$), power (P_{Sim}), and phase noise (N_{Sim}). Since there is error in the model for effective capacitance, Step 7 checks whether the $f_{o,Sim}$ falls in an acceptable range. If it does not pass, the process returns to Step 2, adds the simulation result to regression data, and refines the regression model parameters. Step 8 checks whether power or phase noise from the simulation exceeds the constraint. If either one does not pass, the process returns to Step 3.

2.3.6 Optimization Results

Several design examples for a 0.5 μm CMOS technology with a 3.3 V power supply are presented in Table 2.1. We selected f_o and I_{CSI} to be 1 Hz and 290 nA, respectively as described above. The effect of changing one weighting factor can be seen by considering three distinct cases where the other two weighting factors are kept constant. For example changing the importance of area, power, and phase noise can be observed form data in columns [2,4,5], [2,6,7], and [1,2,3], respectively, to improve the corresponding metric accordingly. The $APNP$(5,14) versus W for W_A=W_P=W_N=1 is shown in Figure 2.12 (the second data column in Table 2.1); the $APNP$(5,13) versus W for W_A = 1, W_P = 1, W_N = 2 is shown in Figure 2.13 (the third

data column in Table 2.1). Two types of *APNP* functions were encountered. Figure 2.12 illustrates a convex function over valid *W* space and the optimal point is at the minimum of *APNP*. Figure 2.13 *APNP* illustrates a monotonic function where the optimal point is at the boundary of the range of valid *W*. The valid *W* ranges are different for these two examples because the *M* parameters are different. The *APNP*s over a possible set of *N* and *M* for two cases are shown in Figure 2.14 and Figure 2.15. Each data point in these two figures is an optimal point of $APNP(N_i, M_i)$. The optimal design parameters *N, M, W,* C_f and performance metrics A_{tot}, P_{tot}, N_{tot} for Figure 2.14 can be also found in the second data column in Table 2.1.

Table 2.1 Design Scenario

Column #		1	2	3	4	5	6	7
Weighting Factors	W_A	1	1	1	0	0.5	1	1
	W_P	1	1	1	1	1	10	55
	W_N	0	1	2	1	1	1	1
Design Parameters	*N*	5	5	5	5	5	5	5
	M	15	14	13	0	13	13	12
	W (μm)	2.80	8.87	17.71	1631	17.71	11.24	15.01
	C_f (pF)	0.96	0.83	0.012	0.87	0.012	1.95	4.05
Performance	A_{tot} (μm^2)	1.5e4	1.6e4	1.9e4	2.7e7	1.9e4	1.8e4	2.1e4
	P_{tot} (μW)	2.41	2.393	2.383	2.370	2.383	2.383	2.377
	N_{tot}	152.9	136.8	124.8	7.2	124.8	128.7	120.2

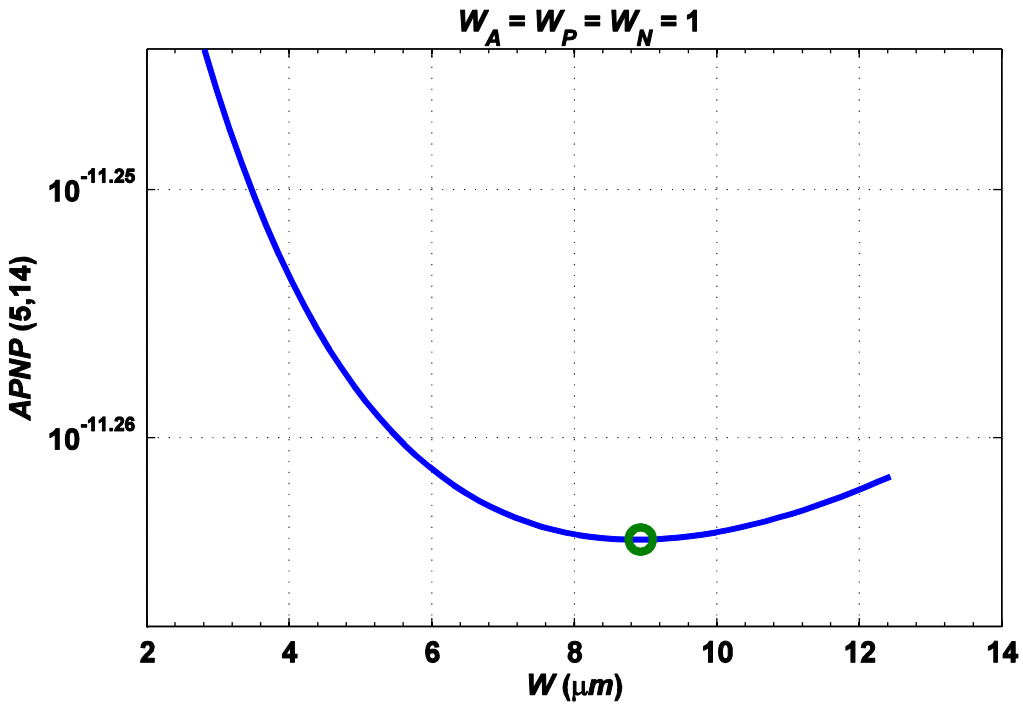


Figure 2.12. *APNP* (5,14) for W_A=W_P=W_N=1. Green circle indicates the optimal point.

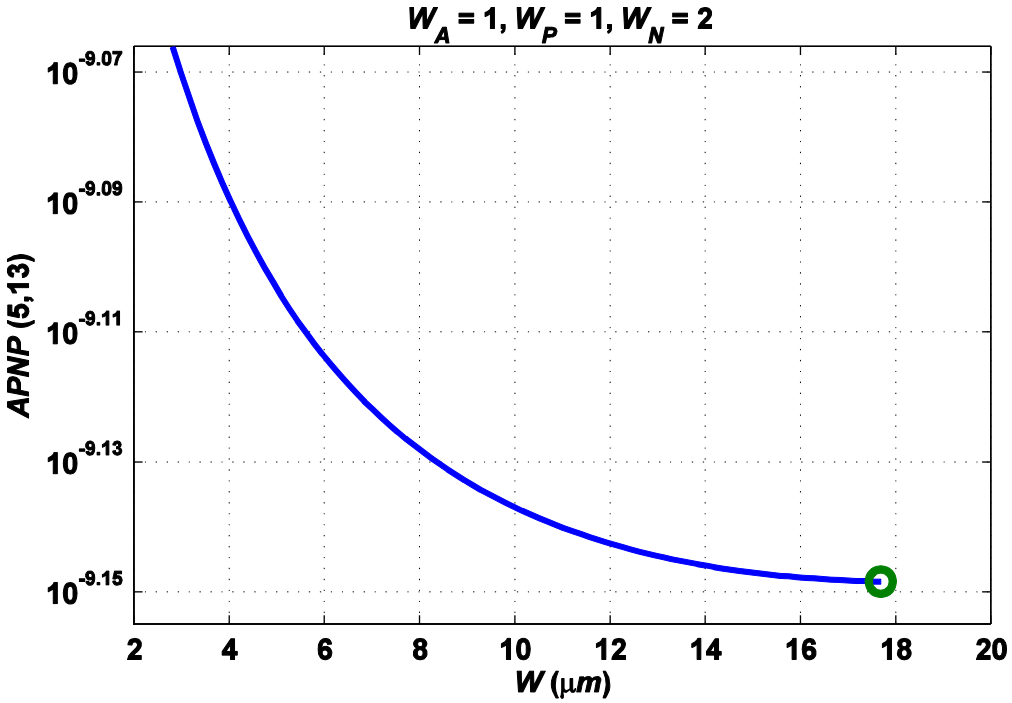


Figure 2.13. *APNP* (5,13) for W_A = 1, W_P = 1, W_N = 2. Green circle indicates the optimal point.

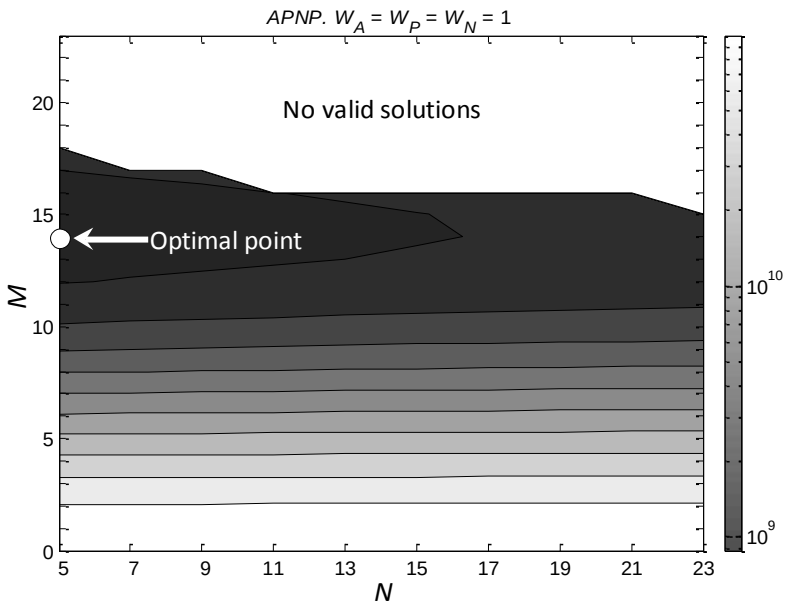


Figure 2.14. *APNP* over (N_i,M_i) for W_A=W_P=W_N=1. White spaces indicate that there is no valid solution. The output frequency cannot be achieved by the parameters in the top white space. The *W* requirement for the bottom white space is smaller than physical limitations for the CMOS process we use.

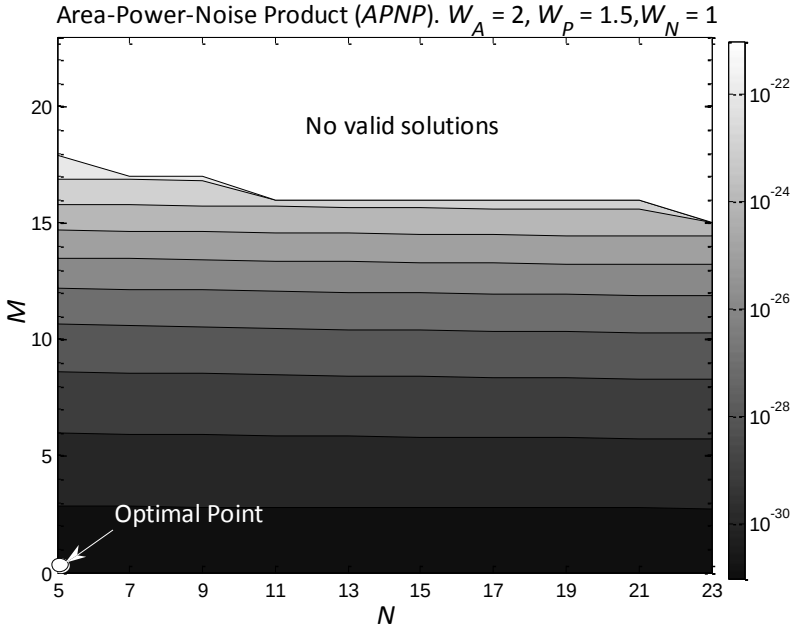


Figure 2.15. *APNP* over (N_i,M_i) for W_A = 2, W_P = 1.5, W_N = 1. White spaces indicate that there is no valid solution because the output frequency cannot be achieved by the parameters in the top white space.

2.4 *Duty Cycle Control and Phase Shift*

After the square wave is generated, two phased control signal can be generated using simple digital combinational logic circuit. Signals with different duty cycle and phase shift can be generated with combination of square waves from DB2 stages. A few examples are demonstrated in Figure 2.16. This approach can be used to generate other types of signal. Triangle and sawtooth waves can also be generated without too much effort.

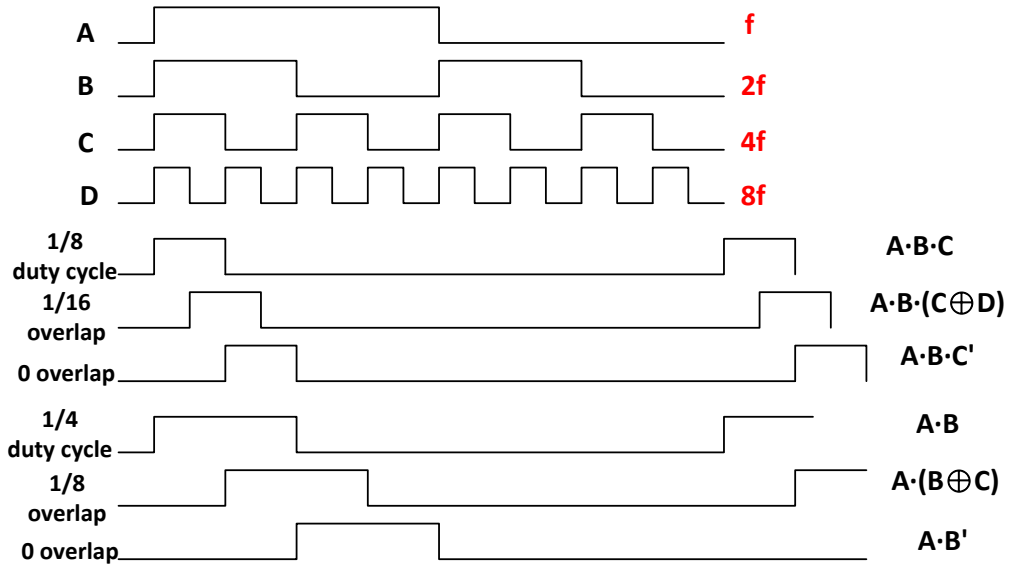


Figure 2.16. Generation of control signals of 1/8 duty cycle with zero or 1/16 overlap and 1/4 duty cycle with zero or 1/8 overlap. The generation utilizes digital combination of four square waves with frequencies of f, 2f, 4f, and 8f.

2.5 *Experiment and Results*

We implemented and tested a CSVCO (Figure 2.3) whose design parameters are summarized in Table 2.2. The CSVCO is followed by a frequency divider with 14 DB2. The measured output frequency versus input voltage under 3.3 V supply is shown in Figure 2.17. An oscillation frequency of 100 KHz corresponds to 6.1 Hz at

the output of the frequency divider. In overall the output frequency of the frequency divider is linear ranging from 0.1 Hz to 12 Hz which well covers the actuation frequency we desired for the MEMS actuators. The linearity makes the tuning easier and improves the programming accuracy as we will discuss in Chapter 3.

Table 2.2 Summary of the CSVCO design parameters. The parameters for transistors are W/L with unit of λ (0.35 μm)

N	M	C_f	R	M1	M2	M3, M4, MI1, MI4	MI2, MI3
5	14	2.04 pF	100 KΩ	6/2	60/10	10/10	8/8

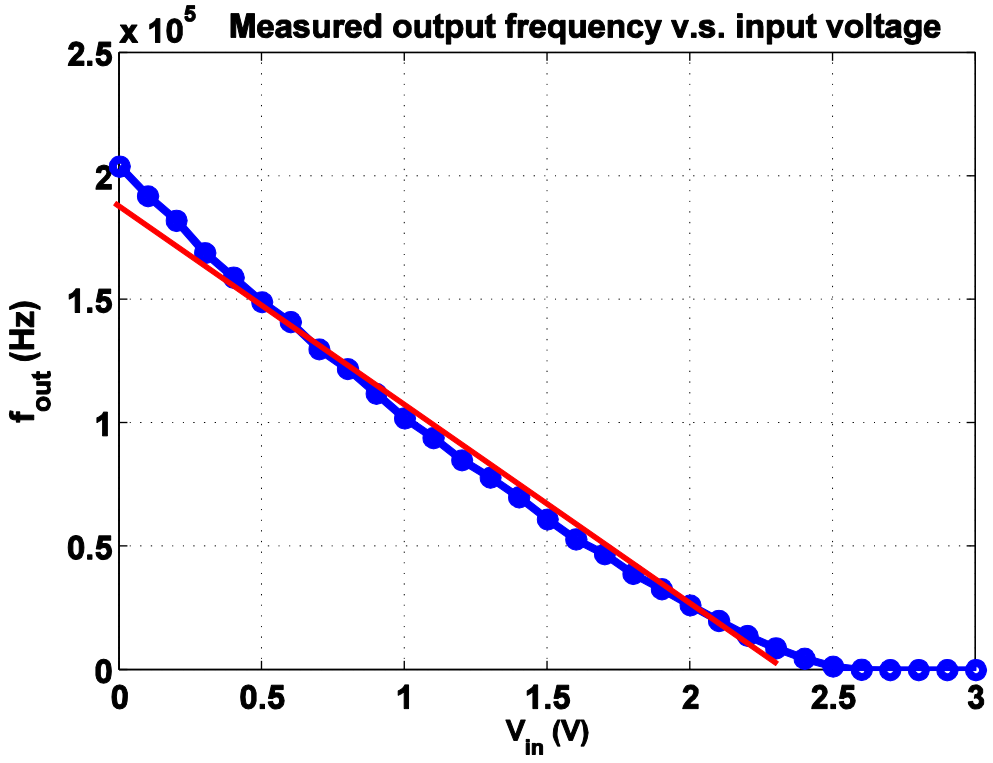


Figure 2.17. Measured output frequency of the CSVCO versus input voltage. Red line is an approximated straight line.

2.6 Discussion

2.6.1 Model-based Optimization

Oscillator simulation is usually time consuming. Moreover, we can only find a feasible set of design parameters because relying on simulation to find optimum design parameters is sometimes impossible. With the aid of analytical or model-based optimization, we can find optimum design parameters in a more timing efficient manner. This gives us a good starting point for simulation. Changes of the parameters are needed during the simulation stage because using compact model to describe a complicated system usually introduces inaccuracies. In our case the optimal parameters rarely fall in the highly inaccurate area so only minor changes of parameters are required during simulation. The optimization flow we proposed allows the designers to adjust the emphasis of different performance metrics and to meet specifications when the requirements vary. We have presented a few examples to demonstrate the effects of changing the weighting factors.

Most importantly the initial simulations have to be done only once for one process technology. Then the regression parameters can be reused for the same technology assuming the run to run variations are not significant. Small variations can be overcome during tuning the design parameters after finding the optimal parameters using the optimization flow. A good coverage of the design parameters have to be selected and simulated to get initial regression parameters. After that, simulations done during design will add to the model and further refine the regression parameters. When the number of samples increase, the model becomes more accurate.

2.6.2 Power Supply Rejection

We propose one possible method to reduce power supply rejection of the oscillator. Power supply voltage on tiny robots tends to fluctuate because the wireless power source is not stable and the on-chip power electronics do not have as good performance as larger scale ones. From Equation 2.1 we found that the output frequency, f_{osc}, is proportional to the inverse of V_{dd} if I_{CSI} and C_{eff} are constant. Our goal is to make the derivative of f_{osc} to V_{dd} zero so that the output frequency is independent of power supply voltage. This is done by

$$\frac{d(\frac{I_{CSI}}{V_{dd} \cdot C_{eff}})}{dV_{dd}} = \frac{\frac{dI_{CSI}}{dV_{dd}} \cdot V_{dd} \cdot C_{eff} - I_{CSI} \cdot C_{eff}}{V_{dd}^2 \cdot C_{eff}^2} = 0 . \tag{2.16}$$

Here, we assume C_{eff} does not depend on V_{dd}. Then, by letting the numerator of Equation 2.16 become zero, we have

$$d(\frac{I_{CSI}}{V_{dd}}) = \frac{I_{CSI}}{V_{dd}} . \tag{2.17}$$

CSI current can be solved by

$$I_{CSI} = \beta_1 \cdot V_{dd} \tag{2.18}$$

where β_1 is a constant. Since we also need I_{CSI} to be proportional to V_{IN} for the VCO to work, the current can be designed as

$$I_{CSI} = \beta_2 \cdot V_{dd} \cdot V_{IN} . \tag{2.19}$$

where β_2 is a constant. In this case we need a voltage multiplier to perform Equation 2.19 and replace the current conversion stage in Figure 2.3 (M_1, M_2, and the resistor).

The problem of implementing Equation 2.19 is that most multiplier circuits have limited linear region and would not function with one input as V_{dd}. One circuit

topology using multiple input translinear elements (MITEs) works for extended input range as shown in Figure 2.18 [74, 75]. Using this topology can also help increase the linear range for V_{in}. When all the transistors operate in saturation region, the output current can be expressed as

$$I_{out} = I_{out1} - I_{out2} = \frac{2K}{(m+2)^2}(V_1 - V_2)(V_3 - V_4) \tag{2.20}$$

where K is the transconductance parameter. If V_1 and V_3 are connected to V_{dd} and VIN respectively while V_2 and V_4 are both connected to ground, the output current becomes what we need as Equation 2.19.

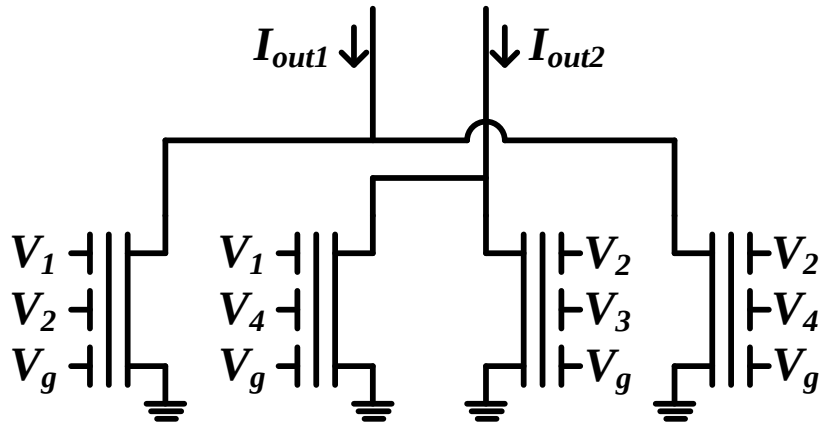


Figure 2.18. Wide input linear range MITE based voltage multiplier [74].

Although the current does not depend on V_g, it has to be carefully designed in order to keep the transistors in saturation. Since

$$V_{fg4} = \frac{V_2 + V_4 + mV_g}{m+2} \tag{2.21}$$

and V2, V4 are both 0 V, this condition has to be satisfied:

$$V_g \geq \frac{m+2}{m} V_{th}. \tag{2.22}$$

If m is set to be one, V_g has to be larger than three times the threshold voltage.

Chapter 3: Memory and Programming – Floating-Gate Phase-Locked Loop*

3.1 Introduction

In order to have autonomous operation, it is necessary to have the ability to locally store command sequences and parameters. Otherwise, initialization when powering up for a huge swarm of robots would be time-consuming or even impossible; all robots have to be programmed through some means before they can perform the tasks. At the same time, programming the memory device to store the right information is sometimes an issue. In this chapter we use the circuits discussed in Chapter 2 as a target of information storage and programming.

The control signal generator discussed in Chapter 2 depends on adjusting the input voltage to produce the right output frequency which yields high actuation efficiency. In order to achieve autonomous operation, ideally we would like all the circuits to function without external inputs (with the exception of power supply). This requires that the system has the ability to determine the right input voltage and store it locally. One possible way to store the voltage is by physically changing devices like a potentiometer. However, this method requires extra devices which cannot be integrated on chip. Another way is to have a digital-to-analog converter (DAC) and hard wire its input pins to the power supply. This method is not appropriate because

* The circuit design and simulation result in this chapter were originally published as " T.-H. Lee and P. A. Abshire, "An ultra-low frequency ring oscillator with programmable tracking using a phase-locked loop," in Proc. *IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2012, pp. 17-20." © 2012 IEEE. The measurement results and the experiments showing the controller driving the actuators in Chapter 5 are in preparation for submission to IEEE Transactions on Circuits and Systems II.

extra pins are required and hard wires are difficult to alter afterwards. Volatile storage is not a good option because initialization at startup is impractical and time-consuming, particularly for a large swarm of robots. Due to the lack of nonvolatile in I-SWARM robot, the optical programming process would take more than 45 minutes [23]. One promising approach is floating gate memories which can be directly integrated on chip. Floating gate is similar to the flash memory and is widely used for Electrically Erasable Programmable Read-Only Memory (EEPROM) [76-78].

3.2 Floating Gate Phase-Locked Loop (FGPLL)

Programming the frequency involves observing the output, analyzing the data, and applying the right control signal. This can be accomplished manually but here automatic programming was implemented in order to simplify operation and use a minimal number of external connections. During programming, bias voltages and a reference signal are applied to the system and the oscillator adjusts to the desired frequency automatically based on these inputs. After programming, the oscillator continues to generate the periodic signal at the programmed frequency requiring only power connections.

The system is essentially a CSVCO circuit placed in the feedback loop of a phase-locked loop (PLL) with a FD at the output. Components include a phase-frequency detector (PFD), an injection and tunneling programming circuit (INJ/TUN), a VCO, and two divide-by (DB) circuits (DBN_1 and DBN_2) as shown in Figure 3.1. There are three main goals of this design: 1) to generate a square wave output with programmable frequency f_o; 2) to function in normal operation mode with no external

biases or controls; 3) to program the output frequency f_o once and retain the stored value even after the power supply is removed.

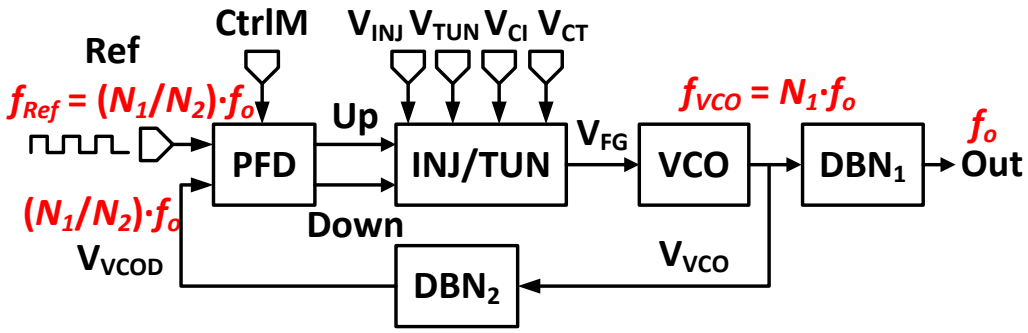


Figure 3.1. System diagram with six input pins. Desired frequency is shown at important nodes.

In normal operation mode, no external biases are required to generate the square wave output. In automatic programming mode, a reference periodic signal Ref and four biases must be supplied. The digital control signal CtrlM is by default high to ensure that no extra connection is needed in normal operation and manual programming modes. When CtrlM is low, the system is in automatic programming mode; the PLL drives the oscillation frequency to match the frequency of Ref. In manual programming mode, V_{INJ} or V_{TUN} is applied manually to adjust the frequency based on the observed output frequency.

The primary innovation in comparison to conventional PLLs is in replacing the charge pump by an INJ/TUN programming circuit. It is the first implementation combing a floating gate and PLL. Tunneling and injection mechanisms replace the charging and discharging functions for the nonvolatile storage mechanism. The oscillation signal is down-converted to low frequencies by a frequency divider.

3.2.1 Circuit Description

Phase-Frequency Detector (PFD)

The PFD was implemented with two additional OR gates at the output as in Figure 3.2. One input of both OR gates is CtrlM. When it is low, which represents automatic programming mode, the two OR gates pass the outputs as would a normal PFD. When CtrlM is high, Up and Down are always high. If sufficient voltage is applied to V_{INJ} or V_{TUN} (see INJ/TUN programming circuit), injection or tunneling begins and programming is performed manually. Otherwise, when neither V_{INJ} nor V_{TUN} is biased, the system remains in normal operation mode.

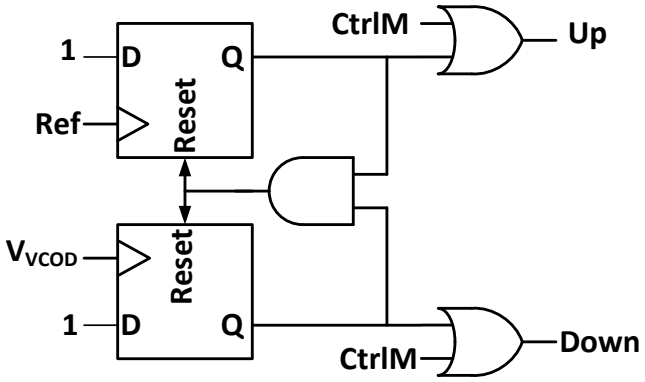


Figure 3.2. PFD with CtrlM control. Two OR gates at the output allow CtrlM to control the PFD operation.

Voltage-Controlled Oscillator (VCO)

A five-stage (N=5) CSVCO was implemented in this design (see Figure 2.3). Detailed descriptions about this circuit can be found in section 2.2.2. For the best locking range and stability of the PLL, the stored voltage V_{FG} should produce the correct output frequency in the middle of the linear VCO gain region which is 1.20 V (same for best tuning range) and corresponds with I_{CSI} ~290 nA.

Divided-By (DB) Circuit

Design of the DB circuits is explained in detail in section 2.2.3. DBN_1 is the FD circuit while DBN_2 does not have a DFF at the output. The numbers of DB2 stages of the two DB circuits are

$$N_{1,DB2} = \log_2 N_1 \tag{3.1}$$

$$N_{2,DB2} = \log_2 N_2 \tag{3.2}$$

respectively, where N_1 and N_2 are divisors of the DB circuits and are both a power of two. DBN_2 is not strictly necessary in this design; however, one stage of DB2 is used to ensure that the duty-cycle of the feedback signal is close to fifty percent.

Injection and Tunneling Programming Circuit (INJ/TUN)

Impact-ionized hot electron injection and Fowler-Nordheim tunneling are used to control the floating gate potential. As shown in Figure 2.3, high V_{FG} reflects low f_{VCO}. Thus, injection (controlled by Up) is used to decrease V_{FG} and increase the VCO frequency while tunneling (controlled by Down) has the opposite effect. The programming circuit is shown in Figure 3.3 (a). A unity gain buffer is used to buffer V_{FG} since we are not able to monitor V_{FG} directly.

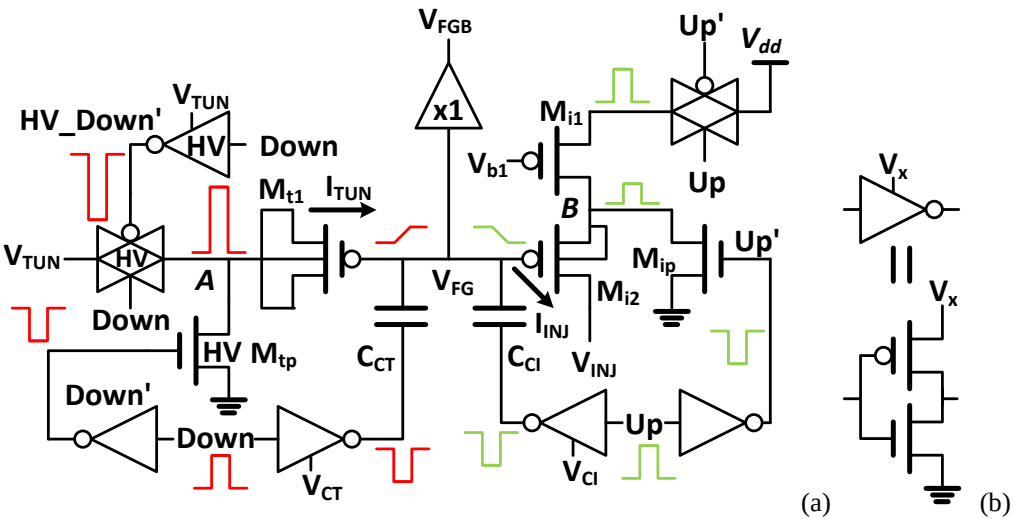


Figure 3.3. INJ/TUN programming circuit with examples of signal changes. (a) Original design with HV devices. Left half controls tunneling while right half controls injection. (b) Inverter with supply voltage different from V_{dd}.

Fowler-Nordheim tunneling is used to increase the floating gate potential. A high potential difference between the oxide junction of M_{t1} twists the energy band and makes the effective barrier lower. As a result the electrons have higher chance to to tunnel from the floating gate node, through the gate oxide, to the MOSFET body. The tunneling current, I_{TUN}, used in the simulation was given by [79, 80]

$$I_{TUN} = I_0 W_{Mt1} L_{Mt1} \exp\left(-\frac{V_f}{V_A - V_{FG}}\right) \tag{3.3}$$

where I_0 is a pre-exponential current, W_{Mt1} and L_{Mt1} are width and length of M_{t1} respectively, V_f is a device parameter depending on the oxide thickness, and V_A is the voltage at node *A*.

Impact-ionized hot electron injection is used to decrease the floating gate potential. Holes in the channel of M_{i2} is accelerated by the large field caused by high source to drain potential difference. Some holes carrying high energy collide with the semiconductor lattice and generate electron-hole pairs. The ionized electrons get high

energy and are promoted to the conduction band. They, then, are expelled the field at the drain and are attracted by the relatively high potential at the floating gate node. Electrons carrying more than 3.1 eV of kinetic energy can potentially overcome the barrier between the silicon and the oxide and inject into the floating gate node. The injection current, I_{INJ}, was modeled as [79]

$$I_{INJ} = \alpha_1 I_s \exp\left(-\frac{\alpha_2}{\left(V_{FG} - V_{INJ} + \alpha_3\right)^2} + \alpha_4 \left(V_B - V_{INJ}\right) \right) \tag{3.4}$$

where I_s is the source current of transistor M_{i2}, V_B is the voltage at node *B*, and α_1, α_2, α_3 are fitting parameters. α_4 equals to one and is used for unit consistency.

To activate tunneling we apply a high voltage across M_{t1} by passing V_{TUN} to node *A* when Down is high, and blocking V_{TUN} when Down returns low. As depicted in Figure 3.3, a pulse on Down resets the signals Downə and HV_Downə. Then, a transmission gate (TG) passes V_{TUN} to *A*. The pull down transistor M_{tp} quickly removes positive charges at *A* to prevent residual tunneling when Down returns low. To block V_{TUN}, HV_Downə has to be at least V_{TUN} minus one threshold voltage to guarantee that the PMOS in the TG does not turn on and leak V_{TUN} to *A*. Since V_{TUN} is higher than the breakdown voltage of normal metal–oxide–semiconductor field-effect transistors (MOSFETs), HV devices must be introduced. To induce injection, M_{i2} must be conducting current and have sufficient source-drain and gate-drain voltage differences. As a result, V_{INJ} is set to a negative voltage.

During programming, large voltage differences at *A* and *B* cause offset at the floating gate through capacitive coupling. In order to compensate the offset, C_{CT} and C_{CI} were added to the design [81]. Each time *A* or *B* is pulled high, compensating

capacitors are switched in the opposite direction. The capacitances of C_{CT} and C_{CI} were designed so that the offset would be minimal when V_{FG} is at 1.20 V and V_{CT} and V_{CI} are at the default voltage V_{dd}.

As described above, HV devices are necessary for tunneling but some technologies do not provide this option directly. An HV NMOS which can be fabricated in standard CMOS processes will be discussed in the Chapter 4. HV PMOS cannot be implemented without control of additional mask layers or doping density [47, 82]. To implement this design in standard CMOS technology, an alternative design is presented in Figure 3.4. Three diode-connected PMOS in series are used as a large resistor. Since M_{tp} is more than ten times stronger than the diode-connected PMOS, node *A* can be pulled to a low voltage to stop tunneling. The large voltage difference between V_{TUN} and *A* is distributed across three identical PMOS. When Down is high, *A* is charged by the series PMOS devices. The primary drawback is that charging node *A* is slower than the original design.

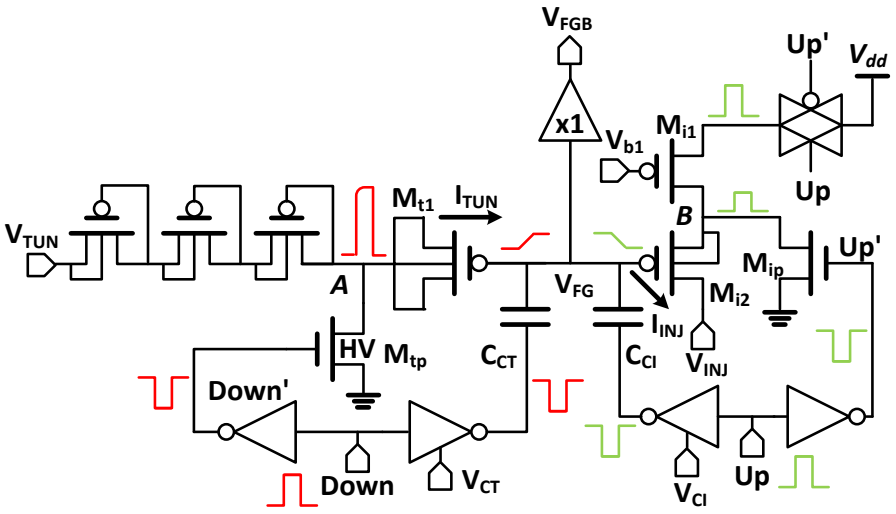


Figure 3.4. Alternative tunneling circuit with no HV PMOS. Large voltage difference across V_{TUN} and node *A* is distributed across three PMOS on the left to avoid breakdown.

3.3 Experiment and Results

3.3.1 Simulation Results

The PLL was simulated using BSIM3.3 models in a commercially available 0.5 μm 2P3M CMOS technology using 3.3 V supply. In our application, the target output frequency f_o is 1 Hz. Since $N_{1,DB2}$ and $N_{2,DB2}$ were chosen to be 14 and 1 respectively, the period of the Ref signal was 112 μs. Ideally, the output period of the VCO would be 56 μs and f_o would be 1.09 Hz. In order to set the output period to 56 μs while V_{FG} is 1.20 V, W and C_f were designed to be 2.8 μm and 2.04 pF respectively.

Transient simulation results starting from different initial V_{FG} are demonstrated in Figure 3.5. Figure 3.5 (a) starts from a voltage close to the desired voltage 1.2 V. The mean voltage is, as expected, 1.20 V and the voltage swing in V_{FG} is 0.2 %. The average, standard deviation, and maximum error of the output period of VCO are 56.00 μs, 0.13 μs, 0.49 μs respectively. The duty cycle is 50.6 %. The corresponding output frequency is 1.09 Hz. Assuming the FD is ideal (divide the frequency perfectly without introducing noise) the data then corresponds to a standard deviation of 18.14 ppm and maximum frequency error of 68.36 ppm at the system output. Figure 3.5 (b) and Figure 3.5 (c) show V_{FG} and the output period settling to their locked states after starting at initial voltages of 1.8 V and 0.8 V respectively.

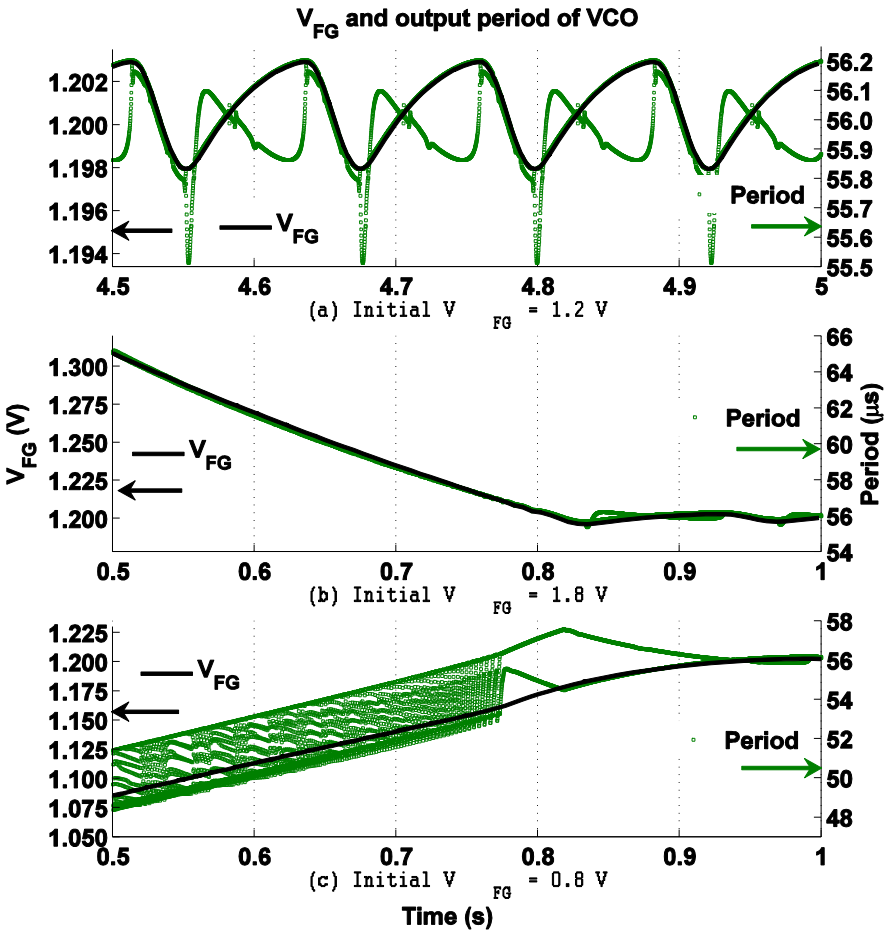


Figure 3.5. Simulation results showing V_{FG} (black solid line) and output period (green square). V_{FG} has been post-processed to remove spikes introduced by voltage coupling from programming. Data points with absolute change rate greater than 100 mV/s were set to equal to the previous data point.

3.3.2 Measurement Results

A complete motion control signal generator was fabricated in commercially available 0.5 μm 2P3M CMOS technology. Two versions were sent out for fabrication: one packaged in a standard DIP40 package and one without a package. The unpackaged chips (see Figure 5.1) were intended to potentially integrate with micromachined actuators. The only differences are the pad frame and the drivers

designed to drive the actuators in the unpackaged chips. The system diagram is shown in Figure 3.6. All results presented here were measured under 3.3 V supply voltage. A measurement result for output frequency versus reference frequency is given in Figure 3.7. For some cases, one mechanism is much stronger than the other one so the output frequency is reluctant to tune to one direction. The line with circle markers (V_{TUN} of 12 V and V_{INJ} of -2 V) has weak tunneling so the output frequency cannot be programmed to lower frequencies. On the other hand, the line with star markers (V_{TUN} of 15 V and V_{INJ} of -1 V) has weak injection so the output frequency cannot be programmed to higher frequencies. As discussed in Chapter 2 the output frequency of the oscillator plus the frequency divider well covers the actuation frequency we desired for the MEMS actuators. Figure 3.7 shows that, under proper biasing, this frequency range (0.01 Hz to 12 Hz) can be fully utilized and programmed into the controller. This figure also indicates that the response of the circuit can be changed by changing the tunneling and injection voltage especially the linear control region (also reflects the locking frequency range).

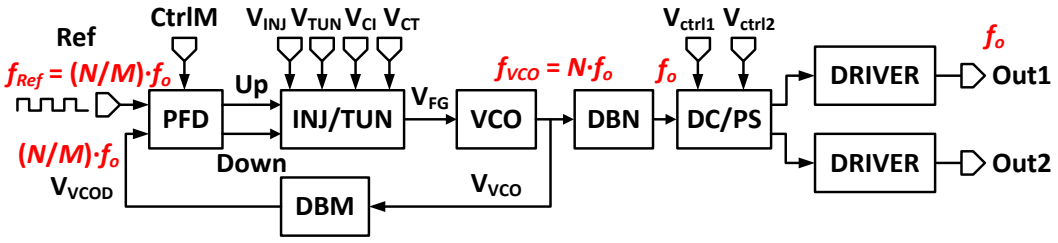


Figure 3.6. System diagram of the fabricated chip. DRIVER is to supply large current to the actuators.

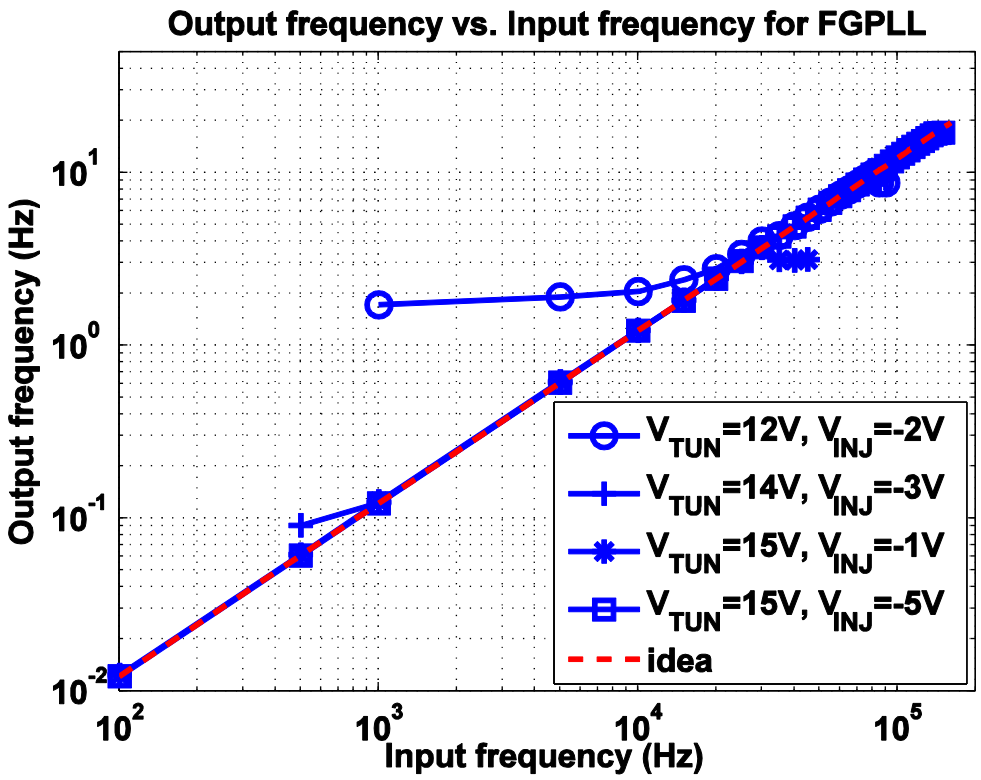


Figure 3.7. Measured output frequency (y-axis) versus input frequency (x-axis) at different tunneling and injection voltage. In the experiment, V_{CT} and V_{CI} were set to zero volt.

The power consumptions for the chip during programming and during normal operation were also measured. However, these measurements only estimate the power because there are other testing structures on the same chip and they are all connected to the same power and ground. Although the biases and the controls to the other testing structures were floating during testing, they might still consume current, at least through leakage. The power reported here was an overestimation, as it assumed that the other testing structures did not consume power. The power measurements during automatic programming (CtrlM = 0) with V_{TUN} = 14 V and V_{INJ} = -4 V were 315 μW, 356 μW, and 427 μW when programming f_{out} to 0.98 Hz, 4.9 Hz, and 7.4 Hz, respectively (input frequency equals to 8 KHz, 40 KHz, and 60 KHz, respectively).

The power measurements during normal operation (CtrlM = 1) with V_{TUN} and V_{INJ} floating were 284 μW and 396 μW when f_{out} programmed to 1.01 Hz and 9.8 Hz, respectively. It showed that most power was consumed by the oscillator and the frequency divider while the programming circuit only consumed tens of μW.

We defined locking as the error of the output frequency is less than 5 % compared to the ideal result. Then, the locking frequency range for each combination of tunneling and injection voltage can be found as shown in Figure 3.8. Here, the numbers are the locking range of the oscillator but not the output frequency for the purpose of more easily determining the range; these frequencies were calculated from the output frequency to obtain the frequency of the oscillator by multiplying it by 2^{14} (14 DB2 stages). Therefore, 100 KHz locking frequency range for the oscillator translates to 6.1 Hz locking range at the output. This result points out that higher tunneling and injection voltage (more negative) should be used in order to achieve higher locking range. We have identified five biasing conditions that yield the best locking range. Biasing for balanced injection and tunneling is also important; V_{TUN} of 16 V and V_{INJ} of 4 V do not result in the best range.

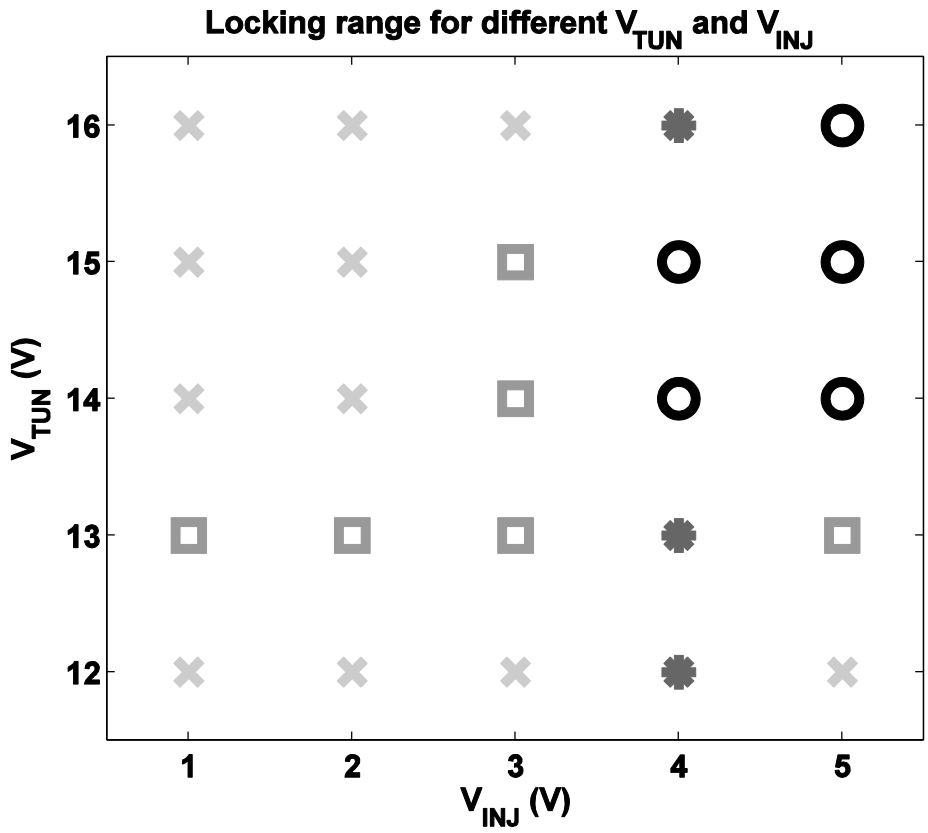


Figure 3.8. Locking frequency range measured at different tunneling and injection voltages. In the figure circle, star, square, and cross stand for a locking range larger than 100 KHz, between 80 KHz and 100 KHz, between 60 KHz and 80 KHz, and less than 60 KHz, respectively.

After extensive testing of these chips we observed that, in some chips that have been tested more intensively, the ability for the floating gate to retain the programmed state degraded. We hypothesized that this was due to the accumulated damage caused by prolonged injection and tunneling currents to the silicon oxide of Mt1 and Mi2 (Figure 3.3), respectively, which reduced the effective barriers. Thus the charge on the floating gate has a higher chance for leakage and becomes more sensitive to the environment. We also observed that the potential on the floating gate node could decrease due to spikes introduced when the chips were suddenly

connected to the power supply. These spikes might accidentally activate injection and decrease the floating gate potential.

3.3.3 Legged Robot Platform

We integrated our FGPLL chip on a legged robot platform as shown in Figure 3.9. The platform was modified from Tamiya Walking Elephant (item# 70094) to further validate its applicability for leg control. The wood work was done in the IREAP mechanical shop assisted by Mr. Nolan Ballew. Mini COTS motors (Tamiya Mini Motor Low-Speed Gearbox item# 70189) used in the platform feature compact size and provide enough torque to drive the whole robot. H-bridge was introduced to enable two way control of the legs because the FGPLL was designed for actuators that only have one way actuation. Thermal actuators recover based on passive cooling to bring down the temperature. The leg control is shown in Figure 3.10. The control is the same as the signal shown in Figure 2.1 with two square signals having overlap. This verifies that the chips can be programmed to a desired frequency and generate the desired control.

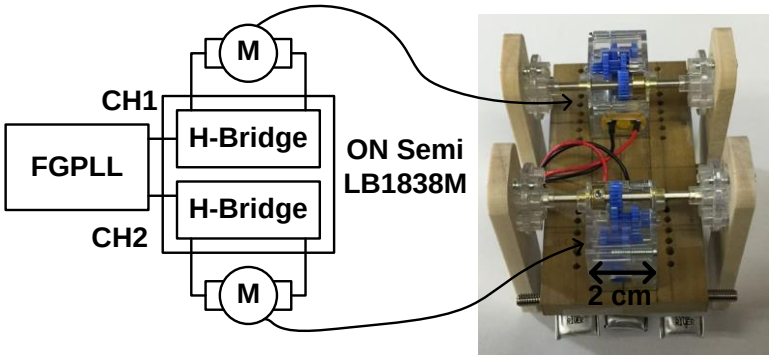


Figure 3.9. Diagram of electronics design of the control PCB for walking robot. Our custom designed chip, FGPLL, controls two H-bridge to drive two motors which drive the legs. The two H-bridges are in a single commercially available chip provided by ON Semiconductor and model number is LB1838M.

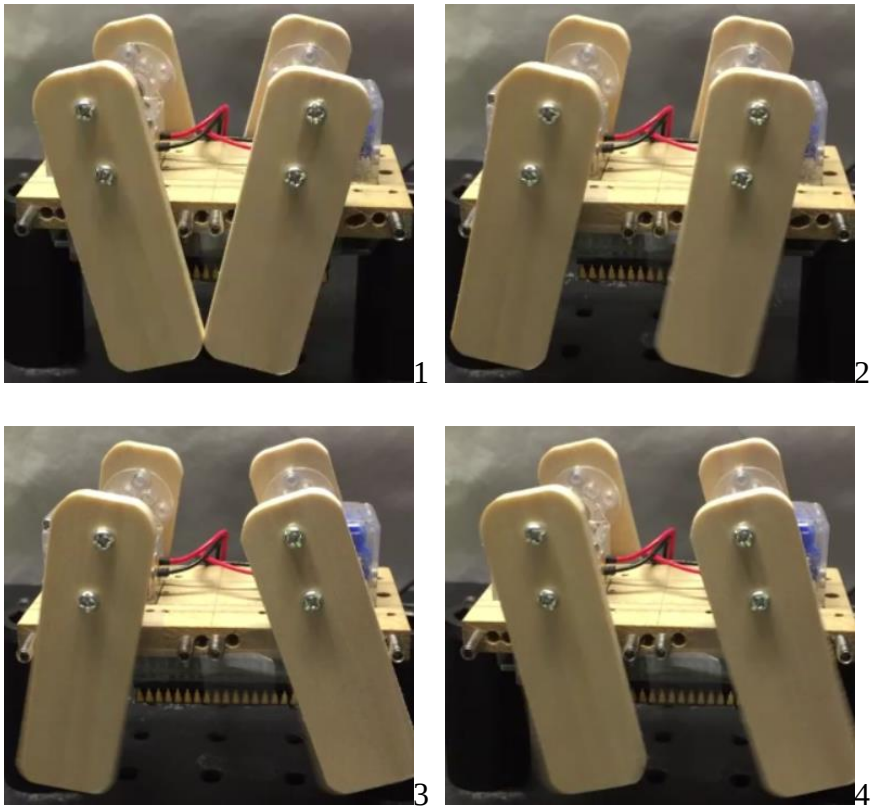


Figure 3.10. Four photos demonstrate leg control of the robot. The pair of legs on the left is controlled by Control 1 and the other pair is controlled by Control 2. Photo 1 has two controls at low level (actuators not being actuated or motor not running). Photo 2 has Control 1 high so the pair of legs on the left swing to the left. Photo 3 has both controls high so the left pair of legs remains at the previous position while the right pair of leg swing to the right. Photo 4 has Control 1 low and Control 2 high. The left pair of legs returns back to the original position while the other pair remains at the previous position. After that, legs return to the position as in photo 1.

Chapter 4: High-Voltage N-Type Metal-Oxide-Semiconductor*

4.1 High-Voltage Usage in Tiny Robots

In our FGPLL design the HV NMOS is used to assist programming of the floating gate structure. In general, to activate tunneling in the 0.5 μm CMOS process we used requires a voltage higher than the breakdown voltage of regular transistors (~12 V). Our goal was to extend the breakdown voltage to at least 20 V to activate tunneling while retaining a safety margin. The usage of the high voltage devices is not limited to only this application. It can be generally used for any circuits which require biases higher than the nominal voltage of the technology. For example, tiny scale electrostatic actuators and thermal actuators use high driving voltages (~20 V) [83-85]. HV devices also benefit sensor performance [86-89] and operating range [90, 91].

In the previous chapter we showed that pull-down and pull-up logic can be implemented with one type of HV MOSFET. HV diodes can also be implemented with the one of the techniques, an extended control gate as reported by Dandin et al. [92]. With this set of devices a complete logic can be implemented to fulfill different requirements in the robotic applications.

* Most of the material in this chapter was originally published as " T.-H. Lee and P. A. Abshire, "40 volt NMOS in a 0.5 mm standard CMOS process," in Proc. *IEEE Sensors Conference*, 2012, pp. 1-4." © 2012 IEEE and " T.-H. Lee and P. A. Abshire, "Design and characterization of high-voltage NMOS structures in a 0.5 mm standard CMOS process," *IEEE Sensors Journal*, vol. 13, no. 8, pp. 2906-2913, 2013." © 2013 IEEE.

4.2 Introduction to HV Devices

Over the past 60 years, feature scaling in CMOS technologies has aggressively reduced transistor sizes, with many benefits including reduced cost, reduced power, and increased speed [93]. These advances are inherently accompanied by technical challenges such as reduced operating voltages and reduced breakdown voltages of the devices, with operating voltages of ~1 V for recent generations of CMOS technology [93]. This voltage limitation introduces incompatibility with many sensing and actuation applications which require high voltages, for example medical instruments, aircraft and automobile electronics, and electric motors. Furthermore, when even modest voltages cannot be achieved in standard CMOS and require the use of specialized technologies, the opportunities for dense system integration are severely restricted and overall system costs increase. Therefore, HV devices in a standard CMOS process are generally enabling for many kinds of integrated sensor and actuator systems [83-92, 94, 95]. Moreover, HV devices are essential to implement floating gate programming as discussed in the previous chapter.

Sensors such as avalanche photodiodes, single-photon avalanche diodes (SPADs) [92], photoconductors [94], and PIN diodes [95] often require high bias voltages. Many actuators also require high voltages, and in some cases the utilization of high voltage devices makes it possible to implement enhanced functionality such as self-test [83, 84]. Electrostatic actuators and thermal actuators use control voltages of ~20 V [83-85]. In many cases, higher operating voltages allow better sensor performance [86-89] or extended operating range for the sensing devices [90, 91]. Additionally, in many systems the availability of high voltage devices allows more

flexibility in the design process and provides the opportunity to achieve more sophisticated control, alleviating some design constraints and enhancing performance.

In this chapter we introduce the design and optimization of four different HV NMOS structures including 1) rectangular structures, 2) drain-centered circular structures, 3) source-centered circular structures with an internal body contact, and 4) source-centered circular structures without internal body contact. To the best of our knowledge, this is the first direct comparison between the observed characteristics of drain-centered and source-centered structures on the same chip.

4.3 Background

Although there are many CMOS technologies with process specialization for high voltage devices, there are many occasions in which it is desirable to implement high voltage devices in standard CMOS technologies that are not optimized for such functionality. Low voltage devices comprise the vast majority of most chip designs, and process technology specializations for high voltage devices would degrade the performance of those low voltage devices. In addition, most designers do not have access to the foundry to control the fabrication steps, for example by adding or changing doping layers and modifying masks. Therefore, in this work we only consider high voltage MOSFETs fabricated in commercially available single-well CMOS technologies. Implementation of these devices often includes introducing a lightly doped drain area to separate the channel and the drain diffusion areas, and extending the poly layer over the intervening field oxide [47, 96]. This technique is similar to the implementation of a laterally double-diffused MOS (LDMOS) [97]. There are relatively few approaches that increase the MOSFET breakdown voltage

which are fully compatible with standard CMOS processes. Santos et al. provide a good summary of these techniques [98, 99], including the introduction of lightly-doped drain drift regions with and without source drift regions, field-ring technique [100], and gate-shifted lightly-doped drain [101].

In a process without p-type wells, it is straightforward to implement a HV NMOS by utilizing the N-well implant as the buffer region. This usually requires deliberate violation of design rules in the physical layout. Implementation of a HV PMOS in such N-well processes, however, requires additional implants and masks to achieve similar characteristics to its n-type counterpart [82]. Fortunately, many control logic or interfacing circuits can be implemented with only one type of high-voltage transistor under acceptable performance degradation [50, 102]. Therefore, this work focuses on the design and optimization of a high-voltage NMOS. The following sections offer comparisons between different high-voltage NMOS structures, discuss the design considerations, and provide characterization results.

4.4 Device Design and Optimization

To understand how different physical structures affect the performance of high-voltage NMOS devices, we implemented four different structures across three fabrication runs using the same CMOS process. While breakdown voltage is an important characteristic for a high-voltage NMOS, current-driving capability and specific ON-resistance are sometimes equally important. In order to explore the design tradeoffs, optimize these characteristics, and achieve a practical balance between these parameters, we implemented these structures using a wide range of dimensions.

The rectangular structure was reported by Ballan and Declercq is shown in Figure 4.1 [47]. L_{g1}, L_{g2}, L_{gd}, and L_{dn} stand for channel length, field poly length, distance from poly edge to n^+ drain, and distance between drain and N-well edge, respectively. The shared drain structure increases area efficiency. Three techniques were used in this work to effectively suppress avalanche and surface breakdown mechanisms. First, the N-well region serves as a lightly doped buffer region to reduce the electric field and to increase the avalanche breakdown voltage at the drain. The relationship between electric field and doping concentration is expressed by

$$\nabla \cdot E = \frac{\rho}{\varepsilon} = \frac{q}{\varepsilon}\left(p - n + D\right) \tag{4.1}$$

where E is the electric field, ε is the permittivity, q is the charge of an electron, and ρ is the charge density. The local charge density can be expressed as the sum of the free hole concentration p, the free electron concentration n, and the ionized impurity concentration D. In comparison with the electron concentration of an inverted channel, the free electron concentration of N-well is much lower. Therefore, the electric field increases more slowly in this region. Second and third, an extended poly and a region of field oxide are introduced between the drain terminal and the channel. The field oxide region blocks the highly conductive silicide layer which is normally deposited on the active region in order to lower the resistance of the drain and the source. The silicide might create a direct path from the drain to the gate and cause undesired effects, and the field oxide reduces the risk of such effects as suggested by Ouyang [103]. Additionally, relatively low voltage at the field poly gate compared to the high voltage N-well area will cause the depletion region of the junction diode between the N-well and the p-substrate to expand further on the n-side due to charge

compensation. This results in less crowded electric field lines at the surface under the field gate and increases surface breakdown voltage [47]. Although field suppression works using a standard gate, edge breakdown can occur when the gate is close to the surface. Therefore it is more effective for the gate to be graded from gate oxide up to field oxide in order to prevent avalanche breakdown occurring at the surface under the gate edge [104]. An illustration of the corresponding breakdowns and equipotential lines after introducing these three techniques are shown in Figure 4.2.

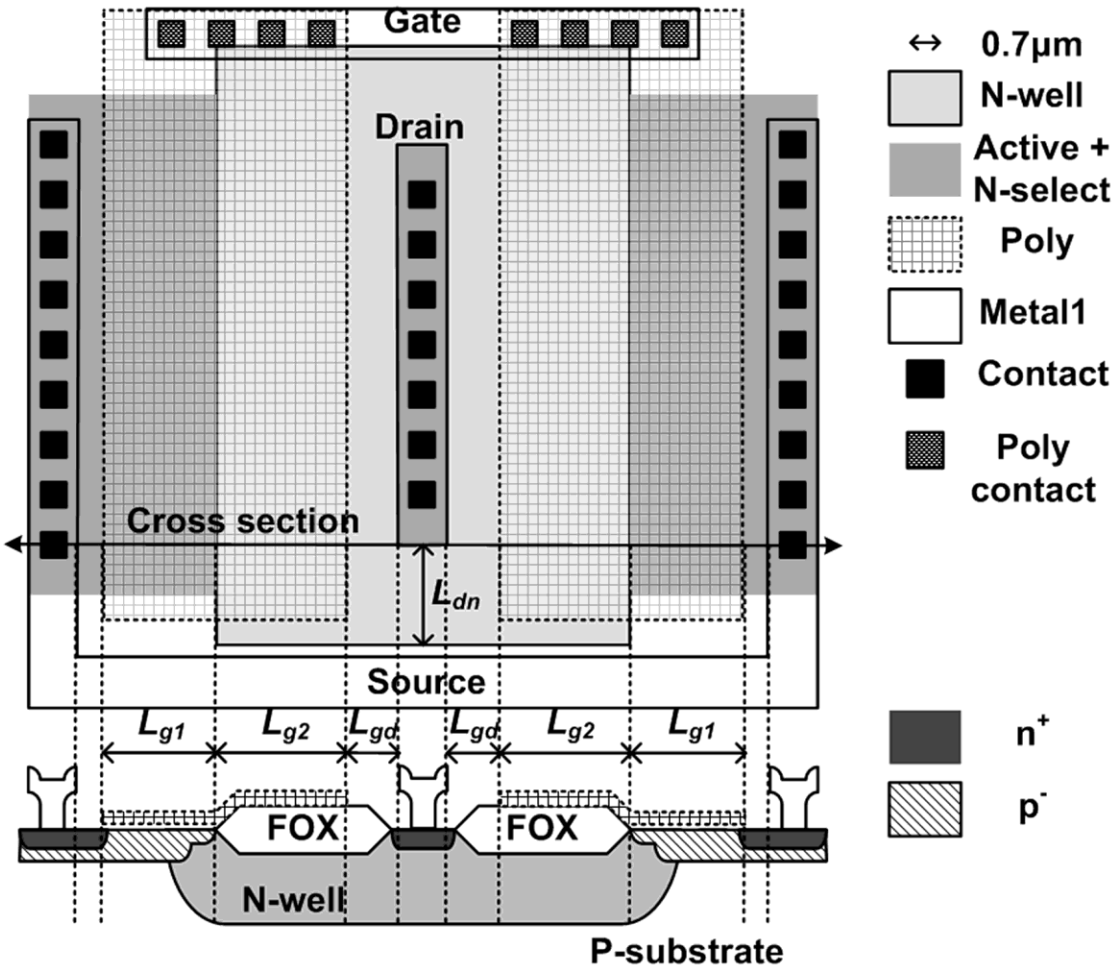


Figure 4.1. Layout and cross sectional views of the rectangular (R) HV NMOS device structure. The p^- region arises from channel-stop and threshold voltage adjust implants [47].

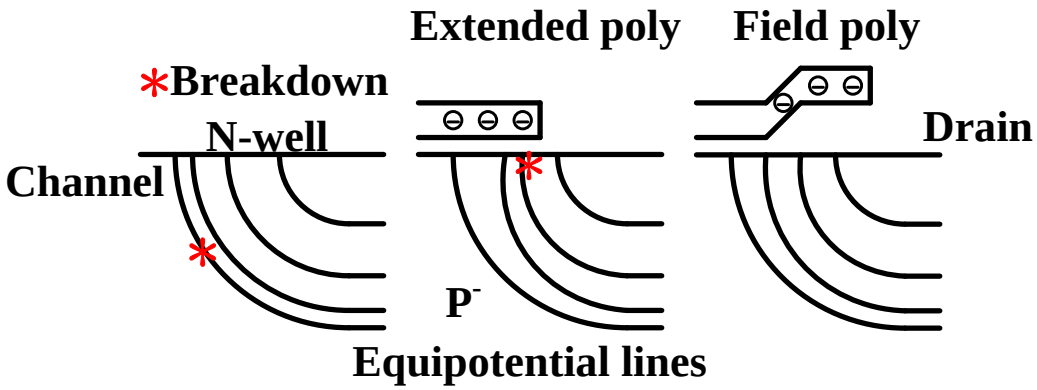


Figure 4.2. Three illustrations show the possible breakdown locations using equipotential lines as suggested by F. Conti and M. Conti [104]. The denser the lines, the stronger the file. (left) After introducing the lightly-doped drain, the densest equipotential lines occur in the interface between the drain and the channel. (middle) The equipotential lines can be altered by introducing an extended poly (can also be a metal layer) with relatively negative potential to the drain. The densest lines occur near the end of the extended poly. (right) The abrupt end of the extended poly is smoothened by having a field poly. This structure is almost free of breakdown.

The rectangular layout illustrated in Figure 4.1 has edges which might introduce edge breakdown and reduce the device performance [103, 105]. Therefore, we also implemented circular structures to eliminate edge effect as illustrated in Figure 4.3. L_d, L_b, L_{bs}, L_s, and L_{sg} are the diameter of the central drain, the diameter of the internal body contacts, the distance from body contacts to source, the source size, and the distance from source to gate. Structure C1 is drain-centered while C2 has internal body contacts surrounded by the source in the center. C3 has a source-centered structure with the internal body contacts removed. For these four structures, majority (inner) and minority (outer) carrier guard rings (GRs) (see Figure 4.4) that surround the core area of the devices were used to reduce the parasitic effects and isolate the high-voltage devices from regular devices. Only two of the structures, R and C2, are shown in Figure 4.4 because the GR dimensions L_{dn} and L_{ngr} affect their

breakdown performance, while GRs have little effect on the breakdown performance of C1.

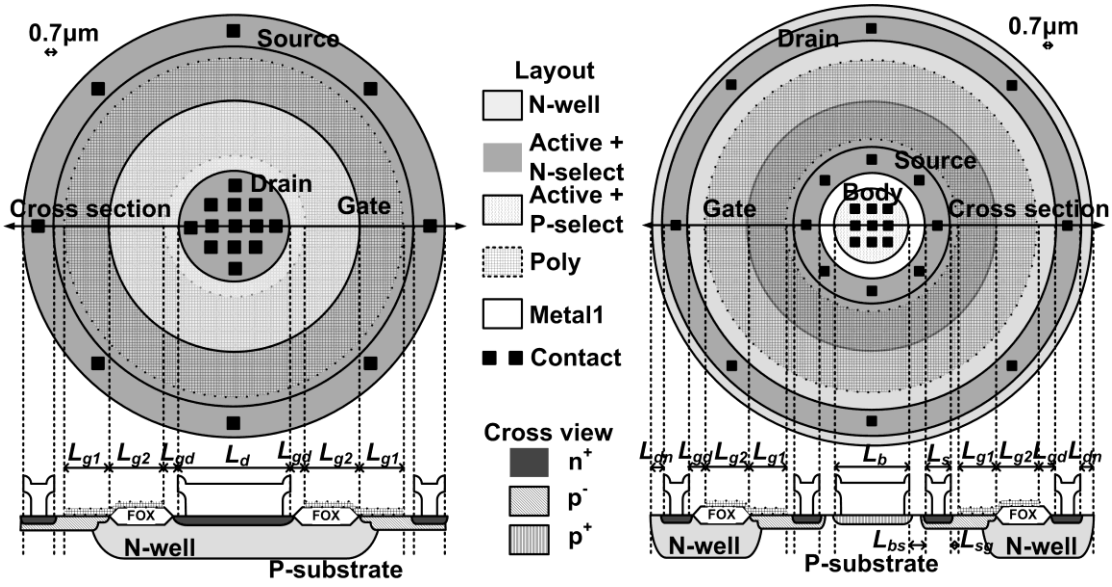


Figure 4.3. Layout and cross sectional views of two circular high-voltage NMOS structures. A drain-centered circular (C1) structure is on the left; a source-centered circular structure with internal body contact (C2) is on the right. The source-centered circular structure without internal body contact (C3) is not illustrated.

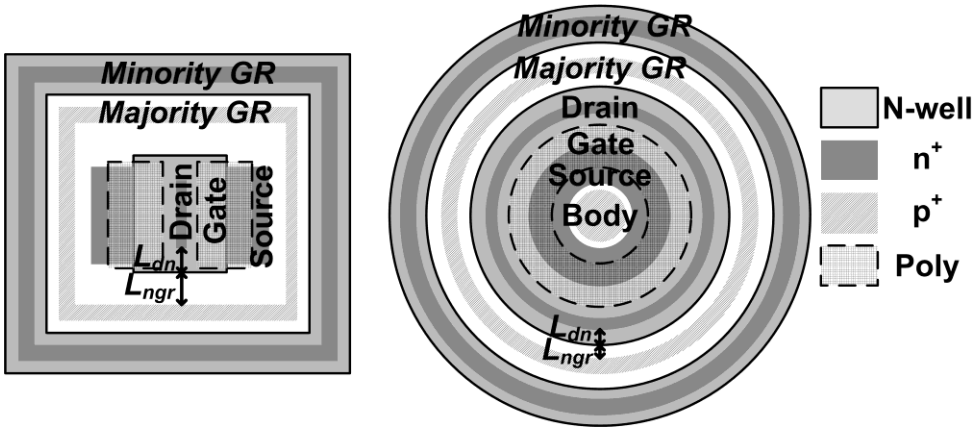


Figure 4.4. R and C2 structures with GRs. L_{ngr} is distance from N-well edge to the inner GR. C3 is again similar to C2 without the inter body contact.

In implementation 1 (I1) there are sixteen combinations of L_{g2} and L_{gd} to explore the dimensional influences on device performance; for the circular device the combinations are split between C1 and C2 configurations as summarized in Table 4.1. In implementation 2 (I2) twenty-nine C1 devices with different combinations of L_{g2} and L_{gd} were fabricated as summarized in Table 4.1. Implementation 3 (I3) has only one C2 and one C3 device with L_{g2} of 5.075 μm and L_{gd} of 1.75 μm. Other parameters were chosen as follows: L_{g1} was set to be 3.15 μm to meet the minimum channel length requirement. L_{dn} and L_{ngr} for the R structures were both chosen to be 2.8 μm which is given by the punch-through condition and design rules [47]. However, in C2 and C3, they were set to 1.05 μm and 1.75 μm as the minimum distance from the layout rules. L_d, L_b, L_{bs}, L_s, and L_{sg} were chosen to be 7.7 μm, 6.3 μm, 1.4 μm, 2.1 μm, and 0.7 μm, respectively. The channel width for the rectangular device was designed to be 28 μm while that for the circular devices is defined by other parameters as

$$W_{C1} = \pi \cdot (L_d + 2L_{gd} + 2L_{g2} + L_{g1}) \tag{4.2}$$

$$W_{C2} = \pi \cdot (L_b + 2L_{bs} + 2L_s + 2L_{sg} + L_{g1}) \tag{4.3}$$

$$W_{C3} = \pi \cdot (L_s + 2L_{sg} + L_{g1}) \tag{4.4}$$

where W_{C1}, W_{C2},and W_{C3} are the effective widths for C1, C2, and C3 devices respectively. The W/L ratio for circular structures cannot be chosen arbitrarily. Fortunately, this issue can be partially resolved using a race-track structure (i.e., extended in one dimension). Table 4.2 lists the area of each structure with and without GR.

Table 4.1 Summary of Devices Implemented for Optimization. Structure in I1; Structure in I2. - not implemented. (all units in μm)

$L_{g2} \setminus L_{gd}$	0.70	1.05	1.40	1.75	2.10
2.275	-;-	-; C1	-; C1	-; C1	-; C1
2.975	-; C1	-; C1	-; C1	-; C1	-; C1
3.675	-; C1	R,C1;C1	R,C2; C1	R,C2; C1	R,C1; C1
4.375	-; C1	R,C2;C1	R,C1; C1	R,C1; C1	R,C2; C1
5.075	-; C1	R,C1;C1	R,C1; C1	R,C2; C1	R,C2; C1
5.775	-; C1	R,C2;-	R,C2; C1	R,C1;-	R,C1; C1
6.475	-;-	-; C1	-;-	-; C1	-;-

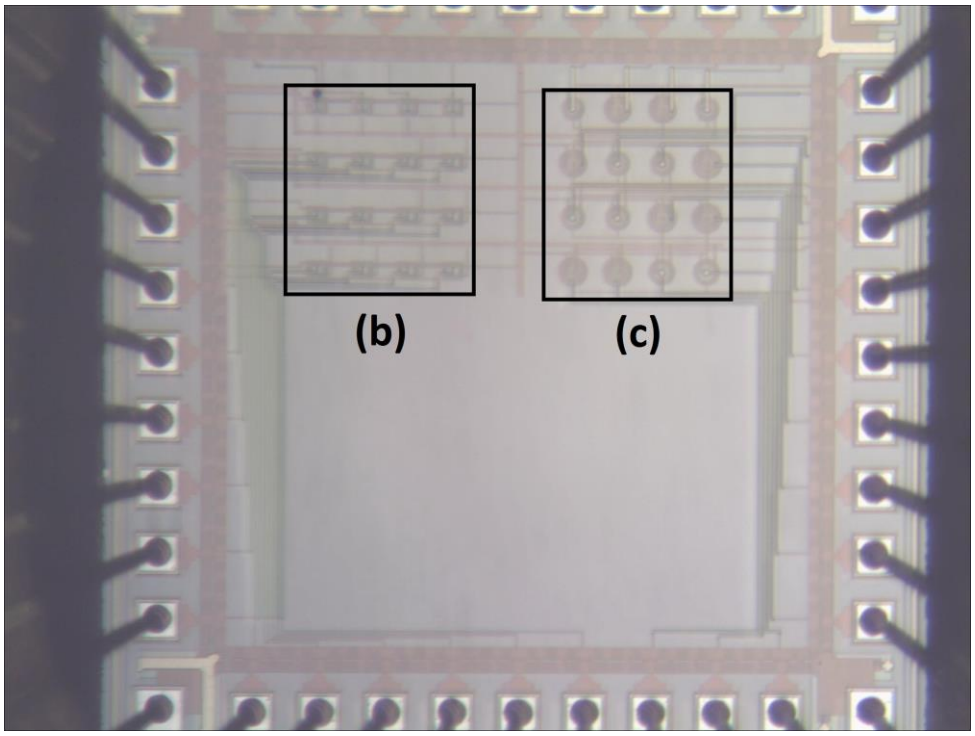
Table 4.2 Device Size with Different Structures (All units in μm)

Device	Area
R	$(W/2+4.9) \times (2L_{g1}+2L_{g2}+2L_{gd}+7)$
R (GR)	$(W/2+20.3) \times (2L_{g1}+2L_{g2}+2L_{gd}+26.6)$
C1	$\pi(W/2\pi+L_{g1}/2+3.5)^2$
C1 (GR)	$\pi(W/2\pi+L_{g1}/2+11.55)^2$
C2	$\pi(W/2\pi+L_{g1}/2+ L_{g2}+L_{gd}+3.15)^2$
C2 (GR)	$\pi(W/2\pi+L_{g1}/2+ L_{g2}+L_{gd}+12.25)^2$
C3	$\pi(W/2\pi+L_{g1}/2+ L_{g2}+L_{gd}+3.15)^2$
C3 (GR)	$\pi(W/2\pi+L_{g1}/2+ L_{g2}+L_{gd}+12.25)^2$

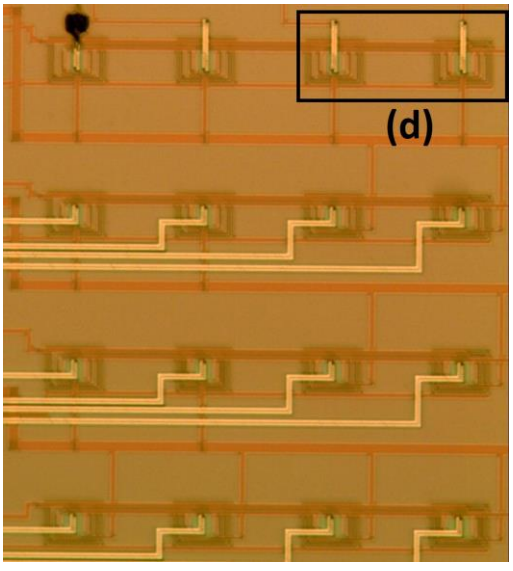
4.5 Measurement Results

The high-voltage NMOS was fabricated in ON Semiconductor C5 0.5μm N-well CMOS technology with three metal layers and nominal operating voltage of 5V. Photomicrographs of devices from I1 are shown in Figure 4.5. Two Keithley 2400 source-measure units were used for characterization. One was used to bias the transistor gate; the other one biased the drain voltage and measured drain current at the same time. A Matlab program was used to interface with the source-measure units and to collect data; this program was based on a program written by Dr. Marc Dandin. Two sets of experimental conditions are specified in Table 4.3, where *i* and *j* are indexes of drain voltage, V_d, and gate voltage, V_g, respectively. We use the notation

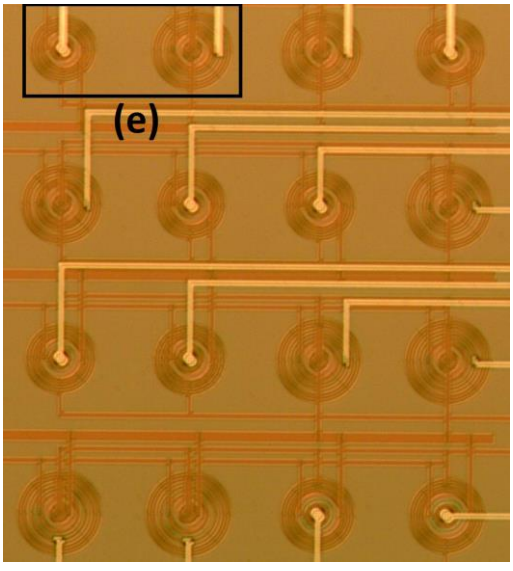
x:*y*:*z* to represent start voltage : incremental voltage : end voltage. In all cases body and source voltages were set to be zero.



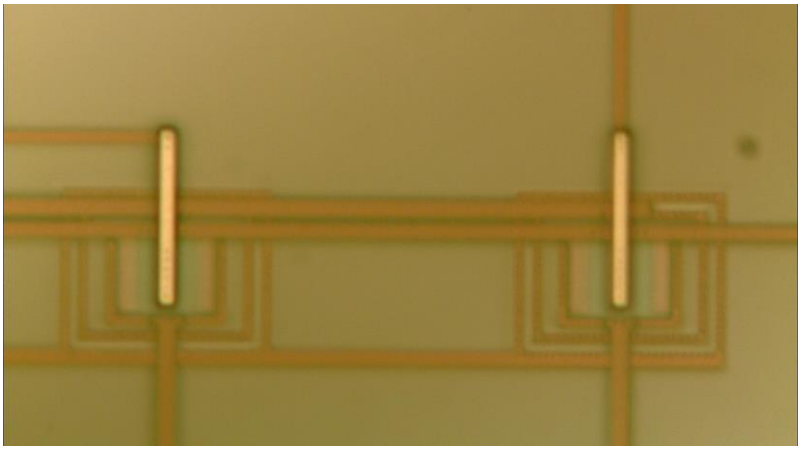
(a)



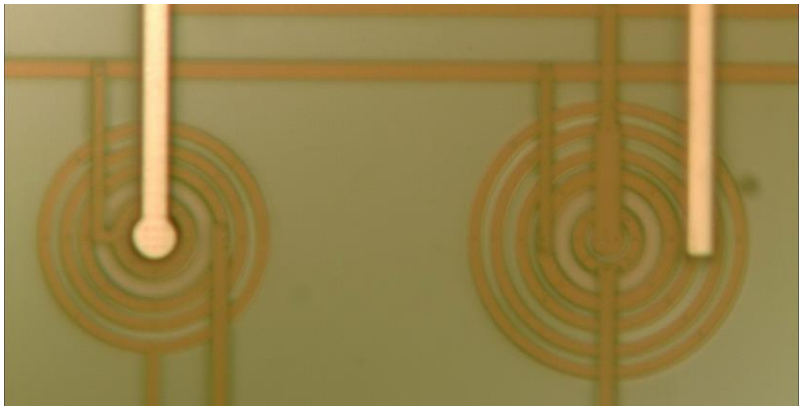
(b)



(c)



(d)



(e)

Figure 4.5. Photomicrographs of the fabricated devices. (a) Overview of I1 chip, comprising 32 R, C1, and C2 devices. (b) Rectangular devices. The metal wire connected to the top left device was damaged due to electromigration. (c) Circular devices. (d) Close-up view of the rectangular structures. (e) Close-up view of the C1 (left) and C2 (right) structures.

Table 4.3 Testing Conditions (All Units in Volts)

Test \	$V_d(i)$	$V_g(j)$
1	0:0.1:8, 9:1:29, 29.1:0.1:50, 51:1:70	0:1:5
2	5:5:25	0:0.1:8

4.5.1 Breakdown Voltage

The breakdown voltage for a single device at a specific gate to source voltage was defined as the drain to source voltage where the maximum slope in the I-V characteristic occurs [92]. This was determined from measured data according to

$$V_{di}(k) = (V_d(k) + V_d(k+1))/2\,, \tag{4.5}$$

$$S_n(k) = (I_{d,n}(k+1) - I_{d,n}(k))/(V_d(k+1) - V_d(k))\,, \tag{4.6}$$

$$V_{BD,n} = V_{di}(\arg\max_k(S_n)) \tag{4.7}$$

where *n* represents different gate to source voltages. First, the center voltages between two adjacent drain voltages were found using Equation 4.5. Next, the instantaneous derivative was computed using Equation 4.6. Finally, the breakdown voltage was found using Equation 4.7. The computations were performed on data obtained under conditions specified by Test 1. An example waveform from Test 1 for each structure is illustrated in Figure 4.6 – 4.8. The flat lines for C2 and C3 reflect the compliance level of the source measure unit. The different curves in each plot are for V_{gs} ranging from 0 to 5 V in increments of 1 V, arranged from bottom to top.

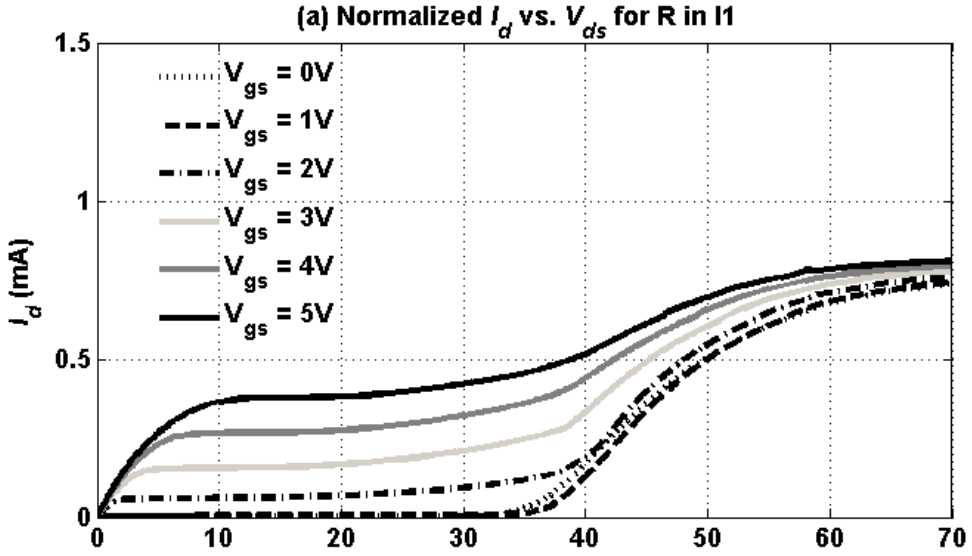


Figure 4.6. Measured I-V characteristic of R structure. The current is normalized by the *W*/*L* ratio. The measured device has L_{g2} of 3.675 μm and L_{gd} of 1.05 μm.

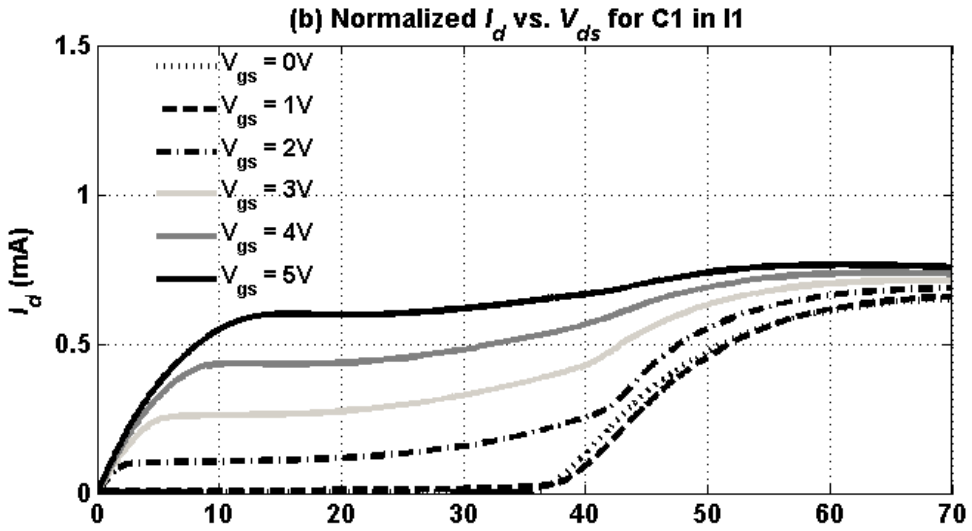


Figure 4.7. Measured I-V characteristic of C1 structure. The current is normalized by the *W*/*L* ratio. The measured device has L_{g2} of 3.675 μm and L_{gd} of 1.05 μm.

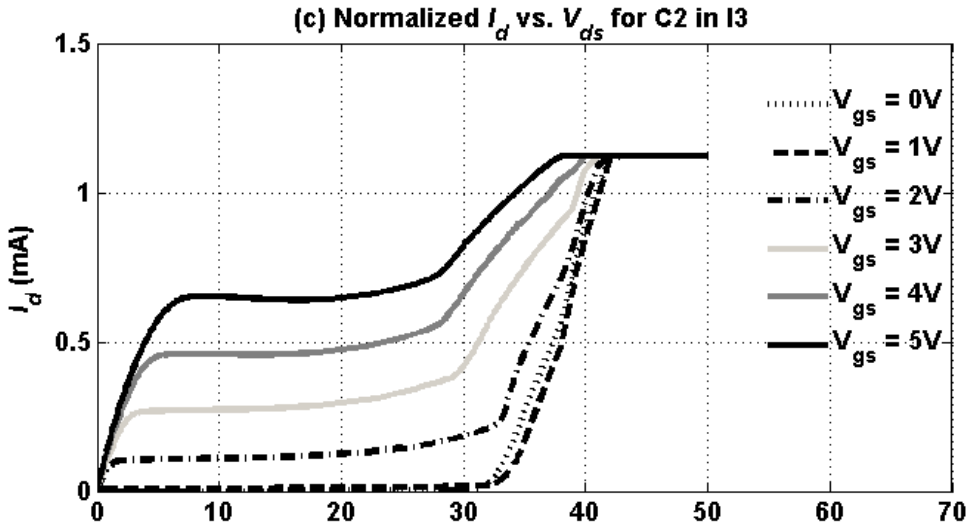


Figure 4.8. Measured I-V characteristic of C2 structure. The current is normalized by the *W*/*L* ratio. The measured device has L_{g2} of 5.075 μm and L_{gd} of 1.75 μm.

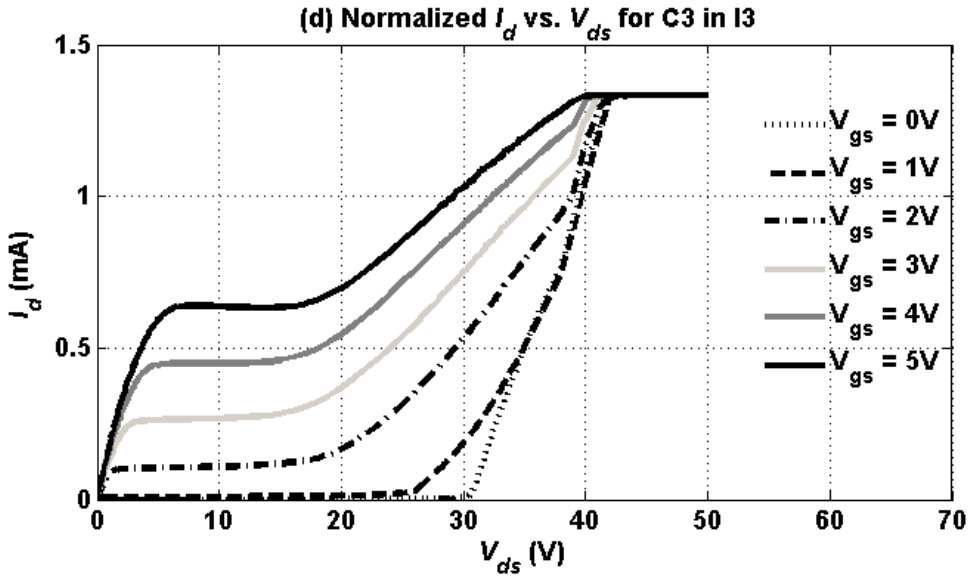


Figure 4.9. Measured I-V characteristic of C3 structure. The current is normalized by the *W*/*L* ratio. The measured device has L_{g2} of 5.075 μm and L_{gd} of 1.75 μm.

The computed breakdown voltages for I1 and I2 are shown in Figure 4.10. Breakdown voltages were found to be the highest for C1 structures with all observed breakdown voltages above 40 V in comparison with 12.5 V for a regular transistor in I2. C2 and C3 exhibited lower breakdown voltages than the other two structures. Table 4.4 provides a detailed comparison of the OFF breakdown voltage (when gate to source voltage is equal to zero) for devices in I1 with different structures and geometry. Table 4.5 summarizes the OFF breakdown voltage for devices fabricated in I2. The results show that there are no consistent trends between L_{g2} and L_{gd} and the resulting breakdown voltage. Details of the breakdown behavior will be discussed in section 4.6. In I3 C2 and C3 have similar OFF breakdown voltage of 38 V. In previously reported work using similar techniques without controlling the doping profile, breakdown voltages were <30 V [82, 103]. The breakdown voltage achieved here was much higher than our goal of 20 V. There are many benefits associated with this higher breakdown voltage as we discussed in the second paragraph of Section 4.2.

Furthermore, the self-regulated breakdown behaviors in R (Figure 4.6) and C1 (Figure 4.7) structures were not conventionally considered as "breakdown," in which current increases dramatically after the onset of breakdown, in contrast to the current behavior shown in Figure 4.8 and Figure 4.9 which is clearly indicative of breakdown. In one reliability experiment we randomly selected one device from both R and C1 structures, biased the devices at a V_{ds} of 60 V for more than 10 hours, and did not observe damage or current degradation of the devices. Therefore, we can potentially operate R and C1 devices beyond the breakdown voltages we reported. In another reliability test to intentionally damage the device by increasing V_{ds}, the metal wire connected to the top left device in Figure 4.5 (b) failed at V_{ds} larger than 85 V due to electromigration before the device was damaged.

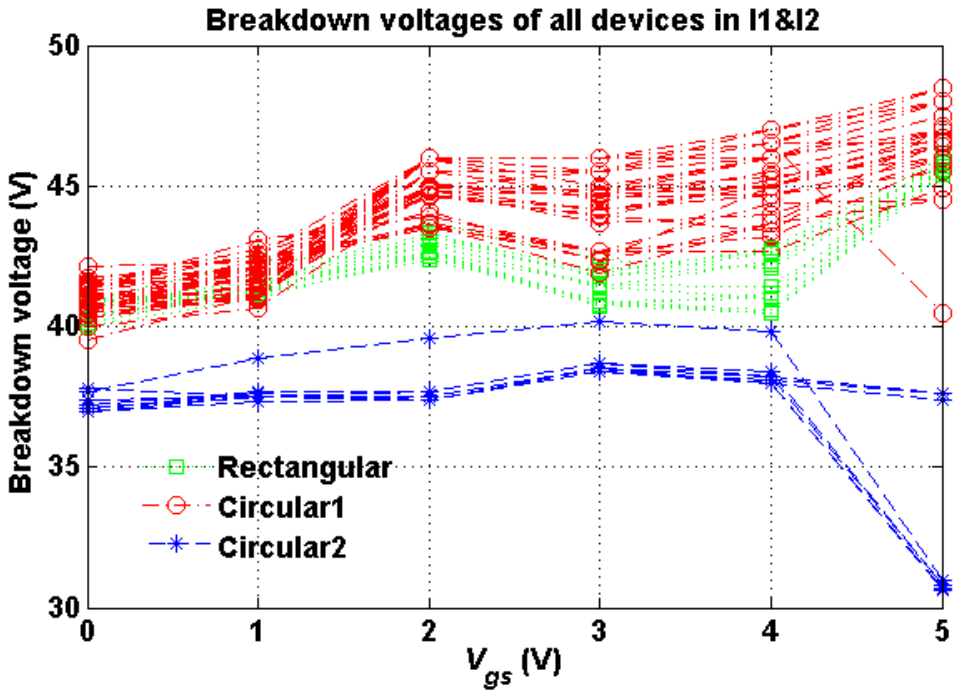


Figure 4.10. Breakdown voltages for circular and rectangular structures in I1 and I2. The C1 structure achieves the highest breakdown voltages.

Table 4.4 Summary of OFF Breakdown Voltages (V) for I1. Rectangular Structure; Circular Structure (Shaded Value Is from C2).

L_{g2} \ L_{gd} (μm)	1.05	1.40	1.75	2.10
3.675	40.9;41.3	40.1;37.8	40.3;37.4	40.0;40.9
4.375	40.7;37.0	40.9;40.7	40.3;40.5	40.4;37.7
5.075	40.8;40.6	40.8;40.6	41.0;37.1	39.2;37.1
5.775	40.8;37.2	40.8;37.0	40.7;40.5	40.9;40.4

Table 4.5 Summary of OFF Breakdown Voltage (V) / Specific ON-Resistance (mΩ-cm^2) of Devices in I2.

L_{g2} \ L_{gd}	0.70	1.05	1.40	1.75	2.10
2.275	-	41.4/3.3	42.2/3.0	41.6/3.0	41.8/3.1
2.975	41.6/3.1	41.5/3.2	41.5/3.3	41.7/3.4	41.0/3.5
3.675	41.3/3.3	41.1/3.4	41.2/3.5	41.1/3.6	40.3/3.6
4.375	41.4/3.4	41.0/3.5	41.0/3.7	40.6/3.9	40.7/4.0
5.075	40.8/3.8	40.8/3.9	40.6/3.9	40.3/4.1	40.5/4.2
5.775	40.6/4.0	-	40.7/4.2	-	40.5/4.4
6.475	-	39.5/4.6	-	40.0/4.5	-

4.5.2 Specific ON-Resistance

The ON-resistance was measured at a gate to source voltage of 5 V and a drain to source voltage of 0.1 V. The area of each device was calculated according to the equations in Table 4.2. The specific ON-resistance is defined as the ON-resistance multiplied by the device area. The specific ON-resistance for devices from I1 is listed in Table 4.6. These results indicate that R and C2 have an average specific ON-resistance of 4.65 mΩ-cm2, while the specific ON-resistance of C1 is 3.78 mΩ-cm2, which is 20% lower than the previous two. In comparison, the specific ON-resistance of a regular transistor in I2 is 4.2 mΩ-cm^2. The results also indicate that specific ON-resistance increases with both L_{g2} and L_{gd}. These correlations are illustrated clearly in the measurements for C1 devices from I2 as shown in Figure 4.11. This figure shows that, with the exception of one data point, when one of the parameters L_{g2} and L_{gd} is

fixed, increasing the other one causes an increase in the resistance. This makes sense because in either case (fix either L_{g2} or L_{gd} and increase the other) the length of the high resistance N-well is increased and, thus, the overall resistance increases. Moreover, the resistance remains relatively constant when the sum of L_{g2} and L_{gd} remains fixed. For I3 the specific ON-resistances for C2 and C3 are 4.8 and 3.6 mΩ-cm2 respectively.

Table 4.6 Summary of Specific ON-Resistance (mΩ-cm^2) for I1. Rectangular Structure; Circular Structure (Shaded Value Is from C2).

L_{g2} \ L_{gd} (μm)	1.05	1.40	1.75	2.10
3.675	3.5;3.7	4.9;4.4	4.6;4.1	4.8;3.7
4.375	3.8;4.0	4.0;3.4	4.8;3.6	5.1;4.5
5.075	4.3;3.6	4.5;3.6	4.6;4.9	5.0;5.1
5.775	5.1;5.0	4.9;5.2	5.1;4.3	5.2;4.3

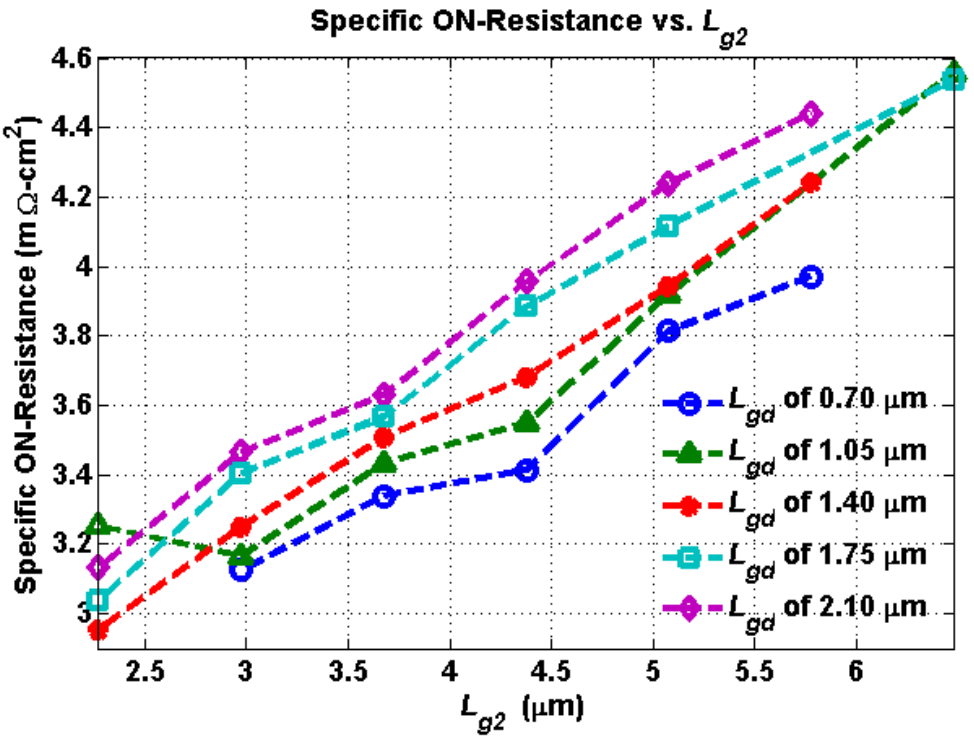


Figure 4.11. Measured specific ON-resistance for C1 devices in I2. The devices on the same curve have the same L_{gd} as indicated in the legend.

4.5.3 Transconductance

The transconductance is defined as the derivative of drain current with respect to gate to source voltage and was calculated on data obtained under Test 2 conditions. An example I-V characteristic from Test 2 for each structure is illustrated in Figure 4.12. The transconductance was normalized to a square transistor (width equals to length) and is shown in Figure 4.13. The trend of the curve is consistent with previously published results by Ouyang [103] and Bazigos [106]. The data reveals that the C1 devices not only have the highest breakdown voltages but also about twice the transconductance as the rectangular devices. We believe this is because the circular structures are more efficient on conducting current compared to square structures. Compared to the transconductance of a standard square transistor in I2 with drain to source voltage of 3 V as shown in the thick dashed line in Figure 4.13, the high voltage devices have comparable performance.

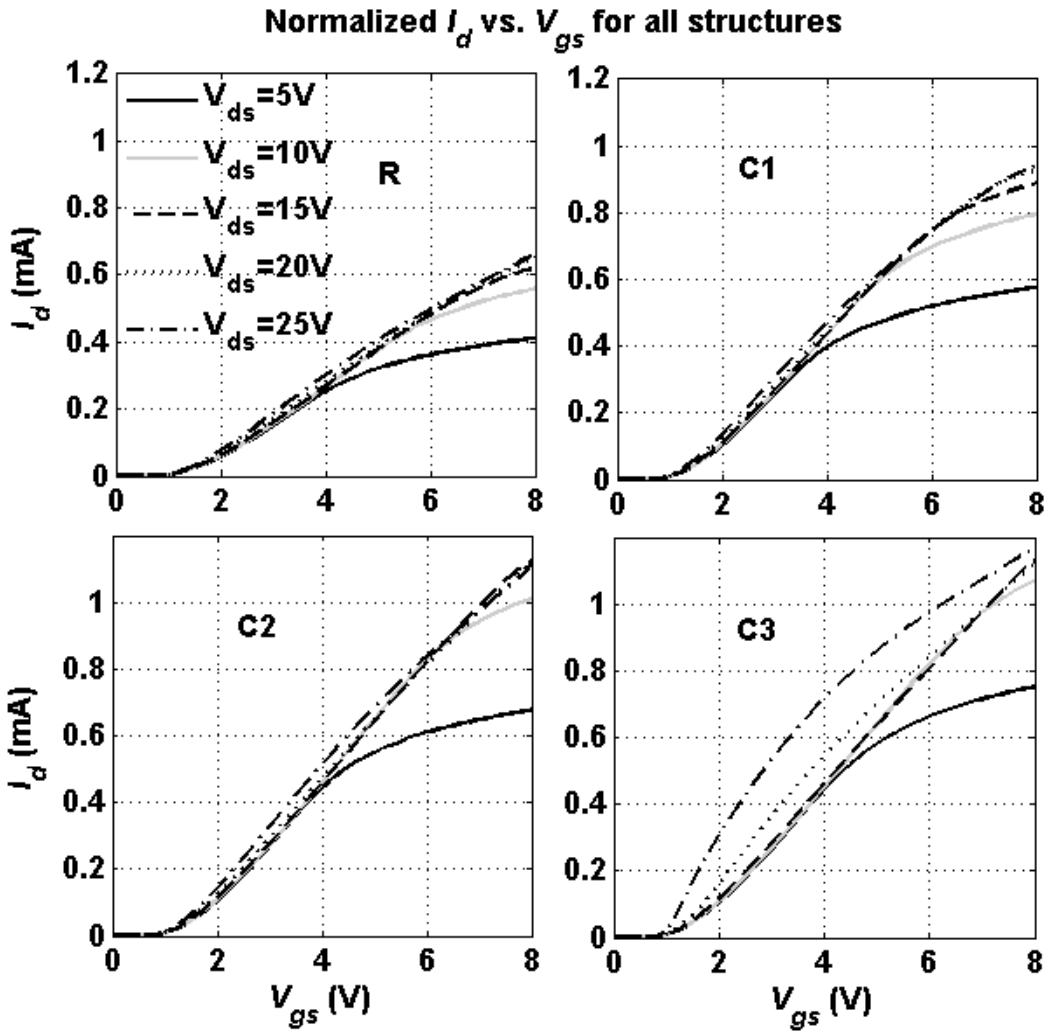


Figure 4.12. Measured I-V characteristic for structures as labeled. The current is normalized by the *W*/*L* ratio. R and C1 devices have L_{g2} of 3.675 μm and L_{gd} of 1.05 μm; C2 and C3 devices have L_{g2} of 5.075 μm and L_{gd} of 1.75 μm.

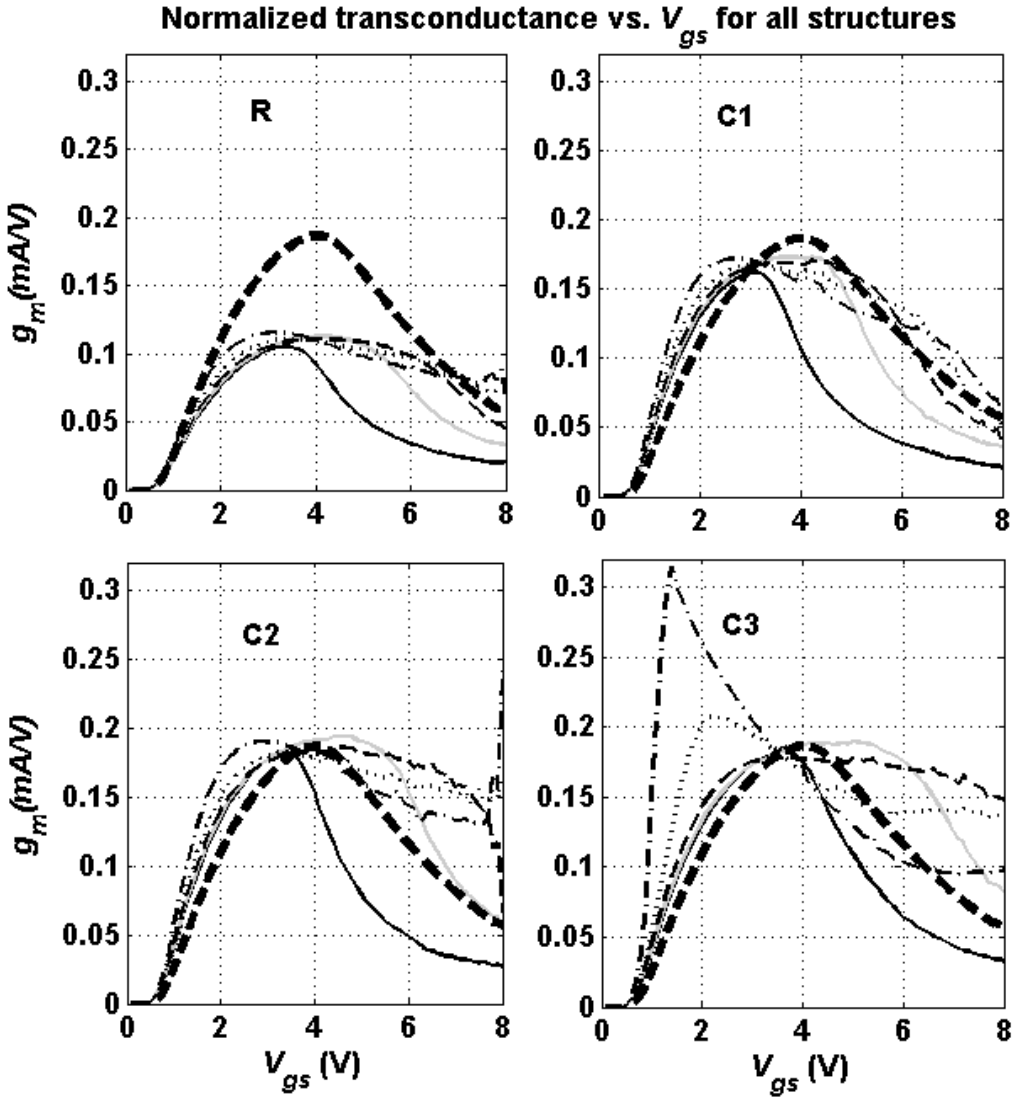


Figure 4.13. Transconductance calculated from the data shown in Figure 4.12 for structures as labeled in plots. The thick dashed line is the transconductance of a standard transistor at a drain to source voltage of 3 V. Drain to source voltages of other traces are as the legend in Figure 4.12.

4.5.4 Modeling and Extracted Parameters

Measurement data from R and C1 devices with L_{g2} of 3.675 μm and L_{gd} of 1.05 μm were fit to model equations and used to extract relevant device parameters. The data was inspected carefully, and only data from transistors operating in saturation were used in the fits. The saturation regime was estimated by assuming that

the threshold voltage is less than 1.5 V. These selected data were then used for regression of the saturation current equation

$$I_d = \frac{K_n}{2}\frac{W}{L}(V_{gs} - V_t)^E\left(1 + \frac{V_{ds} - V_{gs} + V_t}{V_A}\right). \quad (4.8)$$

We assumed the current does not follow the square-law equation and that it follows a power-law with exponent *E* instead. Table 4.7 shows the regression results for parameters of all four device structures. In general the error is low, and the parameters are physically plausible and consistent with other estimation techniques. When the Early voltage was calculated directly from the flat region shown as the middle section of the curves in Figure 4.6 – 4.8, the result was more than 300 V but for the regression results it is ~100V. When the threshold voltage was calculated directly from the data in Figure 4.13, it was in the range of 0.6 V to 0.7 V. This result is more consistent with the expected value. The high-voltage devices should have the same threshold voltage level as standard transistors because their channel regions remain similar.

Table 4.7 Summary of Regression Results for High Voltage NMOS Structures

Parameters	R	C1	C2	C3
# of Points	180	180	180	180
K_n (μA/V^2)	127	251	249	216
V_t (V)	1.097	1.147	1.127	1.080
V_A (V)	100.38	132.47	121.15	32.34
E	1.230	1.104	1.147	1.151
Error (%)	3.78	3.34	3.40	12.97

4.5.5 Yield

Five chips were fabricated and tested for each implementation, for a total of 315 test structures over 15 test chips whose results are reported in this chapter.

During testing, failure in some of the devices was observed. Table 4.8 shows the number of functional devices in chip I1, for a total of 3 failures overall out of 160 structures tested. In addition, there were no failures out of 145 structures tested for I2 and 10 for I3. This gives an overall yield of 99%. The observed failures exhibited a pattern of failing to turn on properly when a large gate voltage was applied. Since the failures were all from the same chip, we believe they were caused by process variations on this particular chip.

Table 4.8 Yield for Devices Tested in Chip I1. Rectangular Structure; Circular Structure (Shaded Value Is from C2)

$L_{g2} \setminus L_{gd}$ (μm)	1.05	1.40	1.75	2.10
3.675	5 ; 5	4 ; 5	5 ; 5	5 ; 5
4.375	5 ; 4	5 ; 4	5 ; 5	5 ; 5
5.075	5 ; 5	5 ; 5	5 ; 5	5 ; 5
5.775	5 ; 5	5 ; 5	5 ; 5	5 ; 5

4.6 Discussion

This chapter reports implementation results for a family of high-voltage NMOS devices. Two techniques were utilized to suppress the avalanche and surface breakdown in the transistor: N-well and field oxide buffer regions in the drain. Majority and minority carrier guard rings were introduced to minimize parasitic effects and isolate the devices. A total of 63 separate devices in four configurations were fabricated in three different runs of a 0.5μm standard 5V CMOS technology with a variety of geometries. Measurement results showed that a circular structure with a central drain has the highest breakdown voltage as well as the highest transconductance which is comparable to standard transistors in the same run. Other

parameters of the devices including threshold voltage and Early voltage were also characterized.

The rectangular and drain-centered circular structures exhibit higher breakdown performance than the source-centered circular structure. We hypothesize that the rectangular and drain-centered circular structures undergo breakdown according to a gate-induced drain leakage (GIDL) mechanism, while the source-centered circular structures exhibit punchthrough and/or avalanche. GIDL breakdown mechanism was rarely observed in high voltage transistors using similar high voltage techniques. One distinct feature of the R and C1 devices is the self-regulating breakdown behavior (see Figure 4.6 - Figure 4.7): the current does not grow fast after breakdown but instead saturates. To investigate the breakdown behavior, we estimated the body current by measuring the drain current for I1 and I3 with source floating. The measurement results show that the breakdown is dominated by the body current; examples for C1 and C2 are given in Figure 4.14. Moreover, section 4.5.1 showed that L_{g2} and L_{gd} have little effect on the breakdown voltages. These together imply that the breakdown is not related to the channel region or the N-well region defined by L_{g2} and L_{gd}. For C2 and C3, the breakdown is most likely caused by punchthrough and/or avalanche breakdown occurring in the reversed-biased junction located at the outer regular N-well sidewall of the lightly doped drain (in contrast to the inner N-well sidewall interfacing with the channel). We estimate that the punchthrough voltage for this junction is 31.5 V using an abrupt junction approximation. This voltage agrees approximately with our observations. The doping concentrations used in the calculation were extracted by fitting parameters of a spice

model; the substrate and N-well concentrations 0.15 μm under the surface were found to be 1.78×10^{16} cm^{-3} and 1.85×10^{16} cm^{-3} respectively. On the other hand, the drain-centered circular structure does not have a regular N-well sidewall, and the regular N-well sidewall of the rectangular structure (upper and lower edge of the N-well region) is better protected using larger L_{dn} and L_{ngr}. The trend of the body current in these structures matches the GIDL observed in [107, 108] although over a different range of voltages. We believe the quantitative difference between prior GIDL results and this work are due to the dissimilarity between the structures studied. Therefore, we attribute breakdown in these two structures to band-to-band tunneling occurring in the N-well region under the gate. We also hypothesize that the regular N-well sidewall is weaker than the one interfacing with the channel and the bottom plate. Therefore, avalanche breakdown occurs for the source-centered circular structures before GIDL is observed. If the L_{dn} (N-well edge to the contact edge) for these two structures can be extended, the breakdown voltage might be improved because punchthrough would occur at larger drain voltage. However, the source-centered circular structures still have larger area compared to the drain-centered structure.

Although the techniques used in this work for implementing high voltage NMOS devices are not novel to this work, we report the highest known breakdown voltages that have been achieved using these techniques. Additionally, this work provides the first direct comparison between drain-centered and source-centered circular devices fabricated in the same technology. An uncommon breakdown mechanism, GIDL, was identified in the rectangular and drain-centered structures. Drain-centered circular devices exhibit breakdown voltages of 40V or higher in all

instances, and rectangular devices exhibit similar breakdown voltages but significantly lower transconductance. They can both operate at a drain voltage beyond the breakdown voltage for hours without visual damage to devices or degradation in output current. Source-centered circular devices exhibited lower breakdown voltages than drain-centered circular or rectangular structures. Although we hypothesize that the breakdown voltage can be improved by extending the distance from outer N-well edge to contact edge, drain-centered circular devices are more efficient than the other devices in terms of area usage. We conclude that drain-centered circular devices with N-well and field oxide buffer regions are the best option for achieving high voltage NMOS devices in a standard CMOS technology.

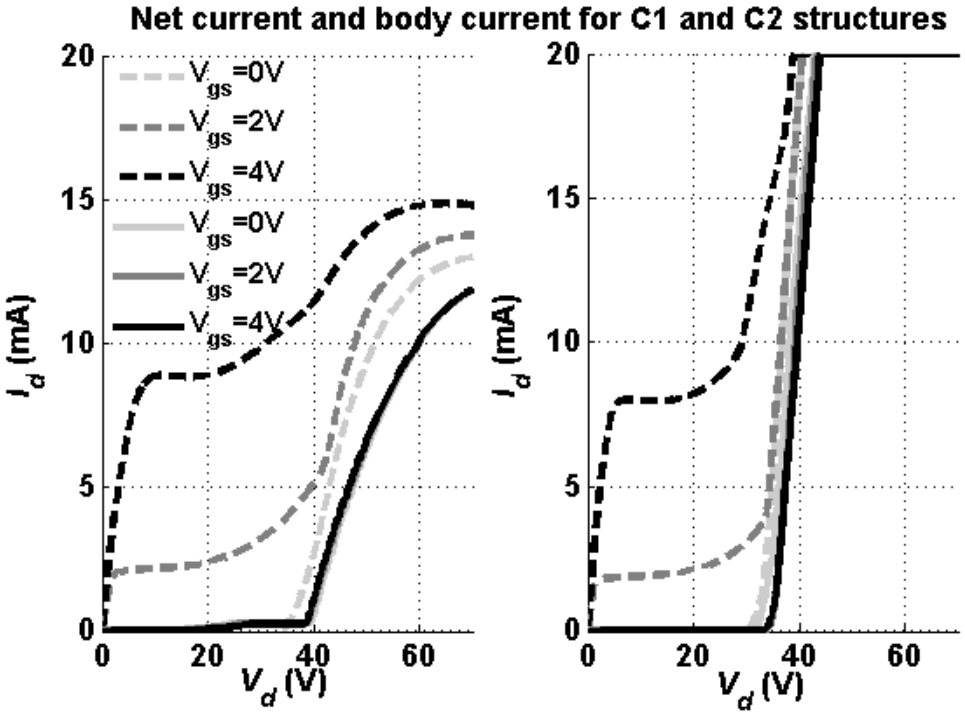


Figure 4.14. Measured I-V characteristic of C1 (left) and C2 (right). Dashed curves and solid curves are net current and body current respectively.

Chapter 5: Post-Fabrication of Actuators on CMOS Chips*

5.1 Motion Mechanisms and Actuators at Tiny Scale

At centimeter scales, motion mechanisms for robots can be assembled using COTS components; wheeled [11-14] and legged [10] motion mechanisms can also be fabricated without extraordinary effort. However at tiny scales, COTS mechanisms are not available [109]. They have to be custom designed and fabricated. Some possible locomotion principles include jumping, walking, stick-slip movement, motion through asymmetric friction forces, and inch worm movement [58]. The Jumping Microrobot used jumping as the locomotion principle and the corresponding actuation is chemical propulsion. Although jumping is an efficient mechanism [18, 19], it is currently uncontrollable at this scale. Design of reusable chemical actuations is difficult and currently cannot be achieved by the Jumping Microrobot; however, multiple actuators could potentially be integrated as suggested by Churaman et al. [18, 19]. The I-SWARM moves by using a stick-slip movement and utilizes piezoelectric actuation. The stick-slip movement principle requires complicated control to achieve manageable movement [8]. The Walking Silicon and Silicon Robot both move by walking. The actuations are thermal, and electrostatic, respectively. However, the Silicon Robot failed after 250 cycles [27]. Moreover, electrostatic actuators require a high bias voltage (hundreds of Volts) to produce reasonable displacement (tens of % of the actuator length). This in turn requires a specialized CMOS process for the chip

* The experiments showing the controller driving the actuators in this chapter and the measurement results in Chapter 3 are in preparation for submission to IEEE Transactions on Circuits and Systems II.

to sustain that high voltage and is against our desire for compact integration. The HV devices with compact integration we described in Chapter 4 could only operate at 60 V (beyond the breakdown voltage). The Walking Silicon featured thermal actuators and moved at a reasonable speed (0.43 BLPS). These existing tiny robots showed that walking principle is promising for locomotion.

Walking principle has an advantage that the whole motion mechanism can consist of only the actuators without other parts. Therefore, the design and fabrication will be easier due to fewer number of mechanical components, which also achieves a smaller size. This work chose thermal actuators because they offer a high force of hundreds μN, which is required to lift the robots, and produce practical displacement of tens of % of the actuator length [110, 111]. The advantage of using the thermal actuators is that this technology is relatively mature. However, thermal actuators consume large currents which are required to create the desired temperature difference (three orders of magnitude higher than the power consumed by the CMOS circuits); these current requirements are impractical since tiny robots have limited power source. The actuator design will have to be optimized to reduce the power consumption so that it can be useful in tiny robot applications. A work combining thermal and electrostatic actuators was reported by Suh et al. to generate the high force and displacement by using thermal actuation and transition to the low-power electrostatic actuation to hold the actuator, so that the power can be reduced while the actuators are holding their positions [112, 113].

5.2 Actuator Design

There are two main types of thermal actuators: thermal bimorphs (or multimorphs) and homogenous actuators [110]. Thermal bimorphs consist of two overlapping materials that have different coefficients of thermal expansion (CTEs). When temperature increases, the mismatch of expansion between these two materials produces out-of-plane displacement. Homogeneous actuators consist of a single material with well-designed structures. Two possible structures are two parallel arms [110] and buckle-beam or chevron [114, 115]. Displacement of the parallel arms structure is caused by the temperature difference between two parallel but connected arms that have different widths; the wide arm has a lower temperature than the narrow arm when current is applied. The displacement is on the same plane as the film. Buckle-beam structure has two head-to-head beams that are at an initial angle. When the beams are heated and expand, the connected part moves forward. It is difficult to convert the in-plane actuation generated from the homogeneous actuators to out-of-plane motion. It is also difficult to use the in-plane actuation on legs. On the other hand, thermal bimorphs demonstrated straightforward integration and out-of-plane actuation. We would like the actuators to be straight (perpendicular to the substrate) in order to maximize displacement at a given bending angle. Vertical actuators cannot be micro-fabricated.

We used an electrode directly as one of the structural layers, which thus reduces the number of required layers and the process complexity compared to having an additional metal layer as a heater sandwiched by two other layers [111-113, 116].

However, the electrode material has to be inert since it is exposed to etchants during fabrication and moisture in air during being heated in normal operation.

The tiny actuators were fabricated using MEMS techniques in the Maryland Nanocenter FabLab at the University of Maryland. Since the actuators were designed to be post-fabricated on the CMOS ASIC directly, the process needed to be CMOS compatible which means that the CMOS chips must not be physically damaged or have their characteristics altered during post-fabrication of the actuators. Any processing steps used in the fabrication should not damage the CMOS chips; the processing temperature should not exceed the allowed temperature for the CMOS chips (such as melting temperatures of materials used on the chips). During actuation, the temperature of the CMOS should not exceed the typical operation temperature range -55 ~ 125 ℃ [117].

5.2.1 Actuator and ASIC Co-design

The design of the actuators necessitates overall planning for both the ASIC and the actuators so that the integration of these two heterogeneous components can be made possible. In order to perform the walking sequence (gait) shown in Figure 2.1, there need to be four rows of legs (each row might have more than one leg). These thermal actuators require a large current to pass through the metal layer (electrode) and accordingly generate heat to raise the temperature. This current was estimated to be 100 mA to produce enough displacement and force. This current is a large load for the CMOS chip, especially if there are multiple actuators that need to be actuated simultaneously (two rows of legs need to be actuated simultaneously). Therefore, we configured the actuators in series so that they shared the same current

source and did not risk exceeding the maximum current that interconnections of the CMOS chips can carry. The maximum current depends on specific designs, we found that our chips could safely source a current up to 200 mA. Exceeding this current level will result in destruction of the circuits, most likely through electromigration of metal wires or vias, similar to what we observed during testing the HV devices (discussed in Section 4.5.1).

The layout of the chip is shown in Figure 5.1. The approximate pad positions and space left at the periphery (pad frame design) were decided by the whole Antbot team. This design reused part of pad frame drawn by Mr. John Turner for the first version of the chip. Two groups of control signals can be produced as outputs at pads *A1* and *B1*. An electrical connection from *A1* (or *B1*) to *A2* (or *B2*) is created by one row of legs (not showing here). Two A2 (or *B2*) pads are internally connected. Pads *A2* (or *B2*) and A3 (or *B3*) are connected by another row of legs. Given this configuration only two output signals are required and the net current can be reduced because multiple legs share the same control (current source).

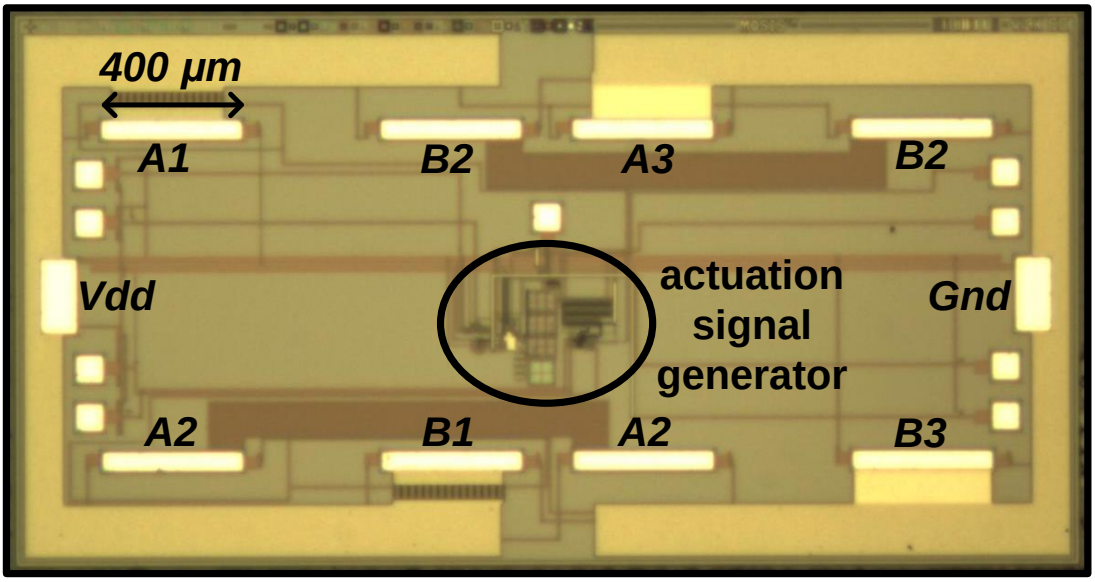


Figure 5.1. Chip photomicrograph showing the layout of the ASIC. The center is the actuation signal generator circuitry. The bright rectangles and squares are the exposed metal layer, elsewhere is covered by passivation material. The two rectangles on either end are *Vdd* and *Gnd* pad, respectively. The rectangles labeled with *A* and *B* are designed to be connected to legs. The squares are testing and programming pads.

5.2.2 Surface Materials of CMOS Chip

It is highly desired to understand the surface materials of the CMOS chip so that the fabrication procedures can be designed to be compatible. The surface materials were given by the foundry as listed in Table 5.1. This information was verified with an elemental analysis technique called energy dispersive spectrometry (EDS). It excites the sample with high energy particles and measures the emission spectrum where each element has unique set of peaks and can be identified accordingly. The analysis was performed in the Maryland Nanocenter AimLab at the University of Maryland assisted by Juan Pablo Hurtado Padilla. Figure 5.2 and Figure 5.3 show the elemental analysis of the exposed metal and passivation layer, respectively. They agree with the information provided by the foundry.

Table 5.1 Surface materials of the CMOS chip from top to bottom.

Passivation	500 nm silicon nitride 1200 nm plasma enhanced tetraethylorthosilicate
Exposed Metal	25 nm TiN ARC 675 nm $AlCu_5$ 10 nm Ti 80-100 nm TiN 30 nm Ti

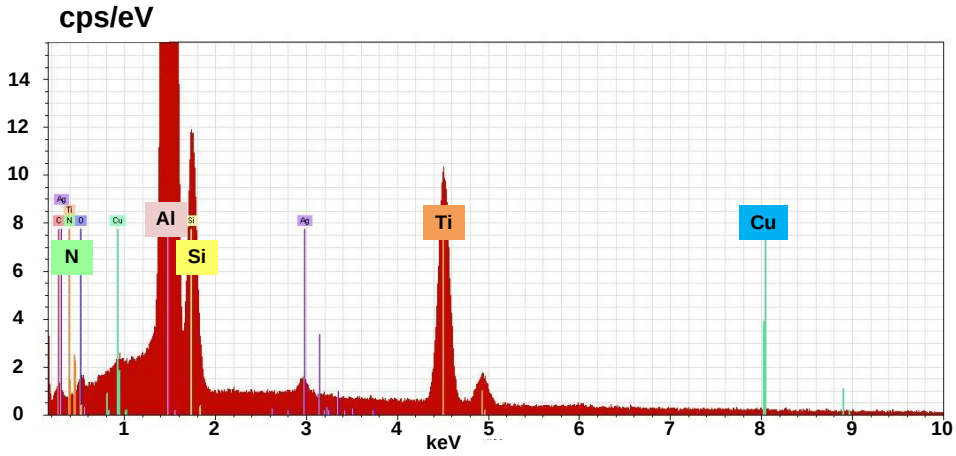


Figure 5.2. EDS analysis of the exposed metal. The elements provided by the foundry (Table 5.1) are highlighted which are Al, Ti, Si, Cu, and N. In the y-axis cps is count per second.

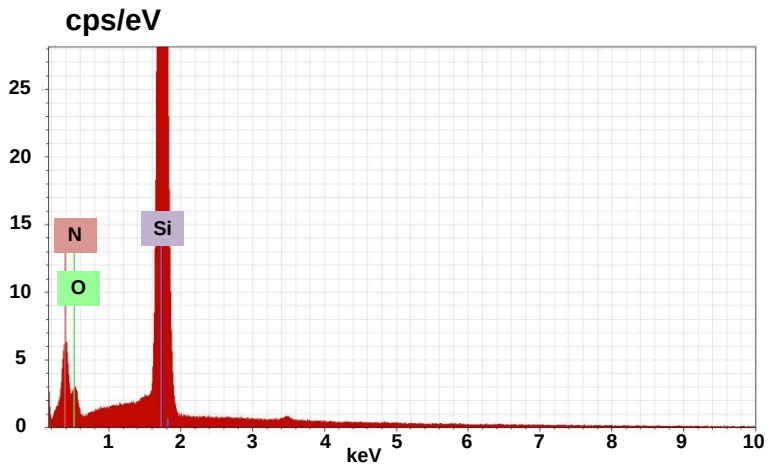


Figure 5.3. EDS analysis of the passivation layer. The elements provided by the foundry (Table 5.1) are highlighted which are Si, N, and O.

5.2.3 Optimization for Bending Force and Radius

The actuators need to have enough force and displacement to move the robot [113]. Therefore, actuation force and displacement are both important factors in the actuator design. The thermal actuators were optimized using a Matlab program developed by Dr. Balakrisnan et al. [118]. Curvature and blocking force were calculated by specifying device geometries, Young's moduli, and strains of each layer. One thing to note here is that the program assumes all layers are uniform, which is different from our structure (the electrode does not cover the whole leg).

We would like to maximize the leg length (in the x direction) because longer legs produce larger displacement for a fixed strain or bending angle. The maximum allowed leg length from the layout of our chip is 480 μm (see Figure 5.1). The aspect ratio of the legs should be larger than 1.3 to ensure that the actuators curve along the longer edges [119]. Given that the length of the legs are 480 μm, the maximum leg width is 370 μm and the corresponding number of legs per row is 2. We designed

three configurations with 2, 3, and 4 legs per row; their leg widths were 375, 250, and 180 μm, respectively.

The CMOS chips received from the foundry were 3 mm × 1.5 mm × 300 μm, which yields a calculated weight of 3.14 μg, assuming the chip is a uniform silicon object and weight of the legs is negligible. To overdesign the leg force by a factor of ten, the desired net force from the legs actuated simultaneously should be 308 μN. The minimum bending angle was chosen to be 30° (assuming the leg curves into a perfect arc of a circle).

The material properties and actuator geometry used in the simulation are summarized in Table 5.2. Device geometry was based on the configuration of 3 legs per row; 6 out of the 12 legs are actuated simultaneously. The CTE of SU-8 was provided by the manufacturer. The reported Young' modulus of SU-8 range from 2 – 7 GPa [120-122]. This results from different measuring techniques, processing conditions, and thicknesses of the SU-8 [120]. One measurement result reported by Kristof et al. matched our application and SU-8 thickness range. Their average Young's modulus was 2.3 GPa [121]. The Poisson's ratio was reported by other work [123]. We used a three layer simulation mode in the Matlab program to optimize the actuator; the three layers were Au, Cr, and SU-8. The strain in each layer is $\Delta T \times CTE$ where ΔT is the temperature difference before and during actuation. We arbitrarily chose the temperature difference to be 100 ℃.

Table 5.2 Actuator simulation conditions for 3 legs per row.

Layers \ Property	Length (μm)	Width (μm)	Thickness	CTE (ppm/K)	Young's modulus (GPa)	Poisson ratio
Gold	480	250	100 – 1000 nm	14.2	79	0.4
Chromium	480	250	100 nm	4.9	279	0.21
SU-8	480	250	1 – 10 μm	52	2.3	0.26

The simulated net force and bending angle for 6 legs are listed in Table 5.3 and Table 5.4, respectively. Grey shaded cells indicate thickness combinations that do not meet the requirement (<308 μN or < 30°). Parameters meeting both requirements are highlighted. The cell marked with a star indicates our final design choice. Force is more important than bending because walking slowly is better than not being able to walk. Therefore, we chose the parameters that produce the highest force among the valid candidates.

Table 5.3 Simulated net force for 6 legs (μN). Grey: specification not meet. Yellow: both force and bending specifications meet. Star: final choice.

SU-8 (μm) \Gold (μm)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	25	22	18	15	11	7	4	0	-4	-8
2	96	94	92	90	88	86	85	83	81	79
3	203	203	204	204	205	205	205	206	206	207
4	340	345	349	353	357	360	363	366	369	372
5	503	515	525	534*	542	549	556	562	568	574
6	689	710	728	743	757	769	780	791	801	811
7	894	928	955	979	999	1018	1035	1051	1065	1079
8	1116	1165	1205	1239	1268	1294	1318	1339	1360	1379
9	1353	1420	1475	1520	1560	1595	1627	1656	1682	1707
10	1604	1691	1763	1822	1874	1920	1960	1998	2032	2064

Table 5.4 Simulated bending angle for legs (degree). Grey: specification not meet. Yellow: both force and bending specifications meet. Star: final choice.

SU-8 (μm) \Gold (μm)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1	137	90	54	30	16	8	3	0	-2	-3
2	83	74	64	54	46	38	31	26	21	17
3	56	53	49	46	42	39	35	32	29	26
4	42	41	39	37	35	33	32	30	28	26
5	34	33	32	31*	29	28	27	26	25	24
6	28	27	26	26	25	24	24	23	22	22
7	24	23	23	22	22	21	21	20	20	20
8	20	20	20	19	19	19	19	18	18	18
9	18	18	17	17	17	17	17	16	16	16
10	16	16	16	15	15	15	15	15	15	15

5.3 Fabrication Procedures

Dr. Bavani Balakrisnan designed initial fabrication procedures and some masks for photolithography. Her initial design did not include metal protection (fabrication step 1). She intended to use silicon oxide and copper as the sacrificial layer (step 2) and the electrode (step 4), respectively. For the other layer of the actuator (step 3), she chose SU-8 instead of polyimide, which is also a popular material for thermal actuators [112, 113], because of concern that the polyimide would be damaged at high temperatures. SU-8 remained in the final process. Ms. Deepa Sritharan helped with the fabrication process. Dr. Elisabeth Smela, Dr. Pamela Abshire, Ms. Deepa Sritharan, and I had regular meetings to discuss the fabrication so the design was a group effort. I did most of the experiments to test and verify different ideas of materials, photolithography recipes, and fabrication steps.

We experimented with using different materials for some of the steps, as summarized in Table 5.5. This summary lists the materials used in the fabrication procedures that we tried experimentally. At the same time two different mask designs were used in the fabrication. Discussion of the material selection and the difference between the two masks is given in Appendix A.1. The fabrication procedures for the actuators are summarized in Figure 5.4; the procedures were for fabricating the actuators in the center of a quarter 4" wafer. Difficulties of fabrication directly on tiny chips are discussed in Appendix B. There are five main steps: 1) deposit and pattern the protection layer for the exposed CMOS aluminum pads, 2) deposit and pattern a sacrificial layer, 3) deposit and pattern leg layer 1 (SU-8), 4) deposit and pattern leg layer 2 (electrode), and 5) release the legs so they can bend. Problems that we encountered in the fabrication are discussed in Appendix A. The layout of the unprocessed chips is shown in Figure 5.1.

Table 5.5 Description of Materials Used in Different Fabrication Sequences

Process \ Step	Protection	Sacrificial	Leg layer 1	Leg layer 2
Dr. Balakrisnan's design	N.A.	SiO_2	SU-8	Cu
Process 1	Cr/Au	SiO_2	SU-8	Cu
Process 2	Cr/Au	Al	SU-8	Cu
Process 3	Cr/Au	Cu	SU-8	Al
Process 4	Cr/Au	Al	SU-8	Cr/Au

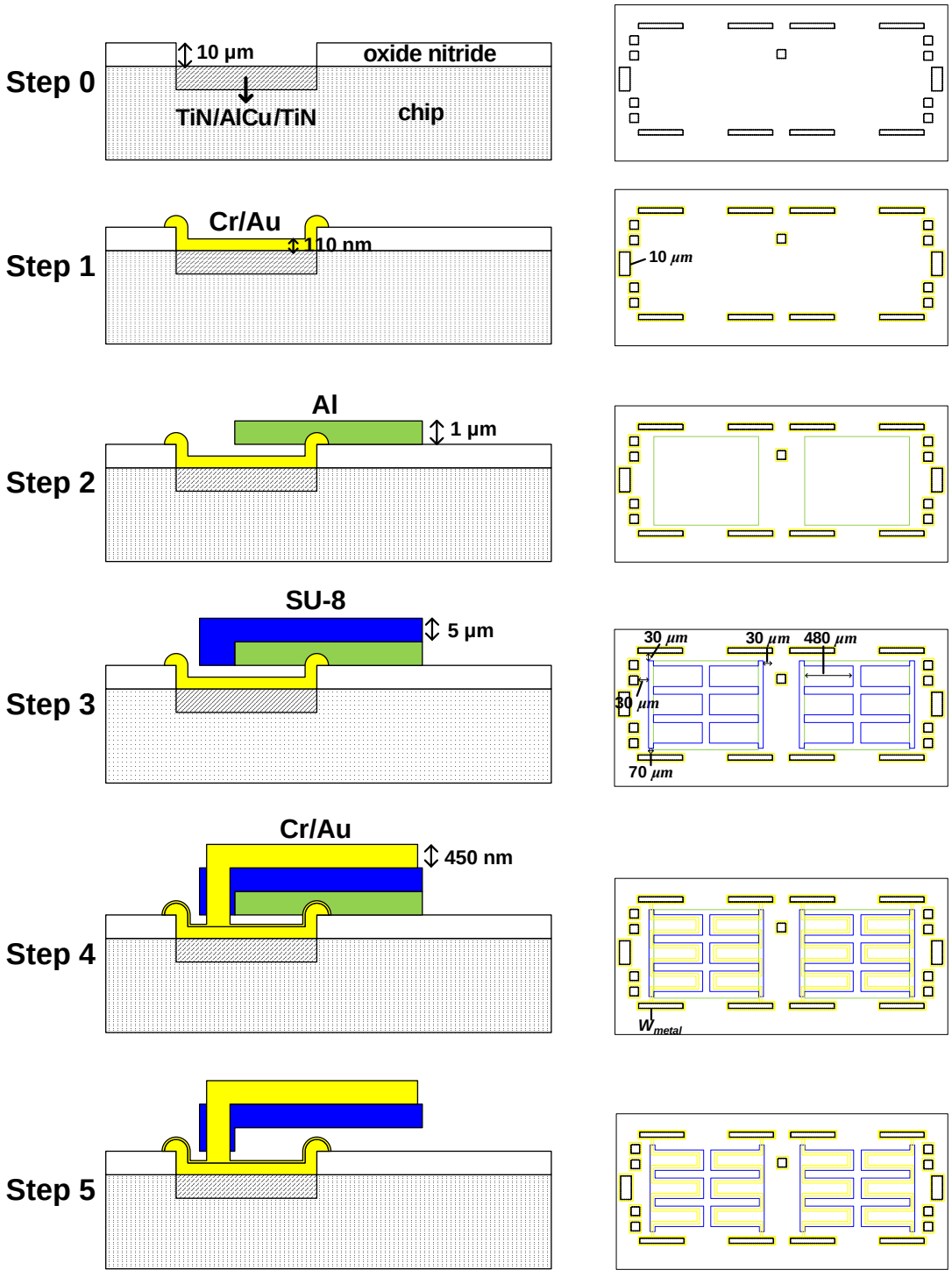


Figure 5.4. Fabrication procedures on the chip. On the left is the cross-section views of the devices; on the right is the top view or masks. Figures on the left column were not drawn to scale. The rectangular outlines on the right are 3 mm × 1.5 mm. The top row shows the schematic of a chip received from the foundry.

Step 1: Exposed Metal Protection

The exposed metal pads on the chips might be damaged because the pads consist of mostly aluminum which does not withstand most etchants. Therefore, we added a metal layer with low reactivity to cover the exposed pads but still maintain its conductivity. We chose Cr/Au because Au is inert and we are familiar with techniques for processing it. This step is not required for processing on a bare quarter wafer. However, the processing done in this step will result in pattern that can emulate the pads of the CMOS chip.

1.1: Cr/Au deposition

Deposit a thin adhesion layer of chromium (10 nm) and then gold (100 nm) by sputtering using an AJA Sputtering unit. Sputter deposition, which is non-directional, was chosen to make sure the Cr/Au covered the side walls of the exposed pads as shown in step 1 of Figure 5.4. The sputter deposition is not required for processing on the quarter wafer, instead thermal evaporation deposition can be used.

1.2: Pattern 1813 photoresist

Dehydrate the chip by baking at 115 ℃ on a hotplate for 10 minutes. Apply positive photoresist Shipley 1813 using a pipette. Spin for 40 seconds at 4000 RPM. Bake at 95 ℃ on a hotplate for 1 minute. Rehydrate for 1 minute at room temperature at a relative humidity of 60%. Align mask 1. A Karl Suss MJB-3 mask aligner was used for all the photolithography steps. Expose for 7.5 seconds at 8 mW/cm^2. Develop for 75 seconds. Rinse with de-ionized (DI) water and dry with a nitrogen air gun. Hard bake at 115 ℃ on a hotplate for 2 minutes.

1.3: Pattern Cr/Au

Put the sample in Transene Gold Etchant TFA to etch the exposed gold. Rinse with DI water and dry with nitrogen. Put the sample in Transene Chromium 1020 Etchant to the chromium. Rinse with DI water and dry with nitrogen. Both etching times were determined by visual examination by bare eye for uniform color of the patterns.

Step 2: Sacrificial Layer

A sacrificial layer was used for releasing the legs later. The selection of aluminum is discussed in Appendix A.1. One thing to note here is that the 1813 photoresist was not removed after patterning the chromium/gold in step 1. The photoresist acted as a physical isolation layer for the first layer gold and the second layer aluminum. The isolation is desired because aluminum and gold form an aluminum-gold intermetallic alloy [124]. We have observed that this intermetallic alloy is less reactive to the aluminum etchant (etch rate is lower than pure aluminum) and sometimes cannot be fully removed with the aluminum etchant. This issue introduces irreproducibility in the fabrication process, such as etching time for aluminum; the aluminum should be avoided.

2.1: Al deposition

Deposit 1 μm of aluminum with a Metra Thermal Evaporator at 18 Å/sec. Thermal evaporation deposition was used because this layer does not need to cover any side walls and its deposition rates are generally higher than sputtering.

2.2: Pattern 1813 photoresist

Follow the same procedures as step *1.2* but use mask 2.

2.3: Pattern Al

Put the sample in Transene Aluminum Etchant – Type A in a beaker on a 50 ℃ hotplate and stir every 20 seconds for 5 minutes. Etching at 50 ℃ is suggested by the manufacturer. Rinse with DI water and dry with nitrogen.

2.4: Remove 1813

Put the sample in an acetone bath and use an ultrasonic agitation system to strip off the 1813 from both step 1 and step 2. Use a March Jupiter III O2 plasma system to strip the 1813 (50 W). Examine the sample for residual 1813 under the microscope. These two methods might have to be repeated until the 1813 is completely removed.

Step 3: Leg Layer 1 (SU-8)

We used negative photoresist SU-8 as the structural layer of the leg. SU-8 was deposited on top of the whole sample, partly over the sacrificial layer and partly on the wafer surface. The portion on the sacrificial layer will be free to bend after fabrication; the portion on the wafer surface will anchor the legs to the wafer. Wet etch via holes (not shown in Figure 5.4) were designed to accelerate the release by increasing the area of the sacrificial layer exposed to the etchant to allow faster undercutting. We used MicroChem SU-8 2005 photoresist for its 5 μm film thickness.

3.1: Pattern SU-8 photoresist

Apply SU-8 with a pipette. We followed the recipes suggested by the datasheet provided by MicroChem except using a lower exposure dose, a gradient temperature profile for post-exposure bake, and a longer development. Spin profile is as follows. Spin at 500 RPM for 10 seconds with acceleration of 100 RPM/sec followed spin at 3000 RPM for 30 seconds with acceleration of 300 RPM/sec. Bake at 95 ℃ on a hotplate for 2 minutes. Align mask 3. Expose for 11 seconds at 8 mW/cm^2, which is slightly less than the dose suggested by the manufacturer, 90-105 mW/cm^2 for 3-5 μm thickness. Bake at 60 ℃ on a hotplate for 1 minute, at 95 ℃ for 3 minutes, and at 60 ℃ for 1 minute (the manufacturer suggested baking at 95 ℃ for 2-3 minutes). We experimented with gradient temperature profiles for both soft bake and post-exposure bake and found that this gradient in temperature profile used in post-exposure bake helped to improve the adhesion of SU-8 to the sacrificial aluminum layer and the wafer surface. While it was suggested that soft bake is more critical to SU-8 adhesion [125], our result differed. Progressive temperature is typically used for soft baking [126]. Develop for 3 minutes with MicroChem SU-8 Developer (the manufacturer suggested 1 minute but we found longer time was needed to remove the unexposed SU-8 completely). Rinse with isopropyl alcohol (IPA) and then DI water. Dry using nitrogen. An additional step of “descum” to remove residual photoresist was sometimes required after development. Descum was done by using the March Jupiter III O2 plasma system with a radio frequency power of 20 W and oxygen pressure of 5 mTorr for 10 seconds.

Step 4: Leg Layer 2 (Electrode)

We again chose Cr/Au for the electrode because it is not reactive with Al etchant. The mask pattern for this layer not only includes the shapes of the electrode but also the same exact shapes for the metal protection layer (see Figure 5.4 step 4). The latter was for protecting the gold deposited in step 1.

4.1: Deposit Cr/Au

Deposit an adhesive layer of chromium (200 nm) and then gold (430 nm) using sputtering. The reason for sputtering is to make sure the Cr/Au is continuous over the steps introduced by the sacrificial layer and the SU-8, as shown in step 4 of Figure 5.4.

4.2: Pattern AZ 9260 photoresist

Dehydrate the quarter wafer by baking at 115 ℃ on a hotplate for 10 minutes. Apply the viscous positive photoresist AZ 9260 using a small beaker, pouring close to the sample. Spin at 2000 RPM for one minute with an acceleration of 100 RPM/sec. Use Q-tips soaked with propylene glycol monomethyl ether acetate PGMEA (main ingredient of MicroChem SU-8 Developer) to remove the edge bead of the sample. Use Q-tips soaked with water to remove the PGMEA (because we observed it changing the color of the developer). Bake at 110 ℃ on a hotplate for 3 minutes. Rehydrate for 60 minutes at room temperature at a relative humidity of 60%. Align mask 4. Expose for 1.2 minutes at 8 mW/cm^2. Develop for 11 minutes using AZ400K mixed with water 1:4 in volume. Rinse with DI water and dry with nitrogen. Descum was done by using the March Jupiter III O2 plasma system with radio frequency power of 20 W and oxygen pressure of 5 mTorr for 10 seconds.

4.3: Pattern Cr/Au

Use the same procedures as step *1.3*.

4.4: Remove AZ 9260

Put the sample in acetone bath with an ultrasonic agitation system to strip off the AZ 9260 photoresist. Use the March Jupiter III O2 plasma system to strip the AZ 9260 (100 W). These two methods might have to be repeated until AZ 9260 is completely removed, especially in the wet etch via holes since the area of the 9260 photoresist exposed to the developer inside these holes is limited.

Step 5: Releasing

5.1: Etch sacrificial layer

Etch sample with Al etchant mixed with water 1:1 in volume at 50 ℃. This step might take up to 20 hours. The reason that it took such a long time to release the legs was because the etching relies on the etchant accessing the sacrificial layer only from the side (undercut). With the aid of wet etch vias, the sacrificial layer still has to be etched horizontally for 50 μm.

5.4 Fabrication and Testing Results

We fabricated actuators on a dummy quarter 4" wafer. Devices on two quarter wafers were tested on a probe station. The resistance of the electrodes forming top of 27 rows of legs was measured: average, median, and standard deviation of the resistance were 28.91 Ω, 20.80 Ω, and 27.85 Ω, respectively. Actuators were actuated by connecting a power supply to probes which touched the electrodes. Top and side views of the actuators before and during actuation are shown in Figure 5.5 and Figure

5.6, respectively. All the horizontal rectangular metal pads shown in the photos are 400 μm wide. The actuation current was set to a value close to the maximum allowed current for the actuators so that the actuators were not damaged while still creating a large enough bending that could be observed. Most actuators were damaged above 100 mA, as shown in Figure 5.7. This 100 mA is lower than the current that the controller chip can supply so damaging to the CMOS chips was not a concern. From the dashed lines in Figure 5.6 we measured the bending angle at the tip to be 32°. This angle was observed from an oblique perspective and so, theoretically, the actual bending angle is higher when observed horizontally. This angle meets our specification of 30° and should provide enough displacement when the devices are ready to be flipped to walk. We attributed the difference between the simulation (31°) and the measurement to two settings in the simulation that do not reflect the real experiment. First, the simulation assumes the two leg layers are uniformly overlapping, but the electrode layer only overlaps part of the SU-8 layer. Second, we assumed the temperature difference caused by actuation is 100 ℃, which might be off from the difference in the real experiment.

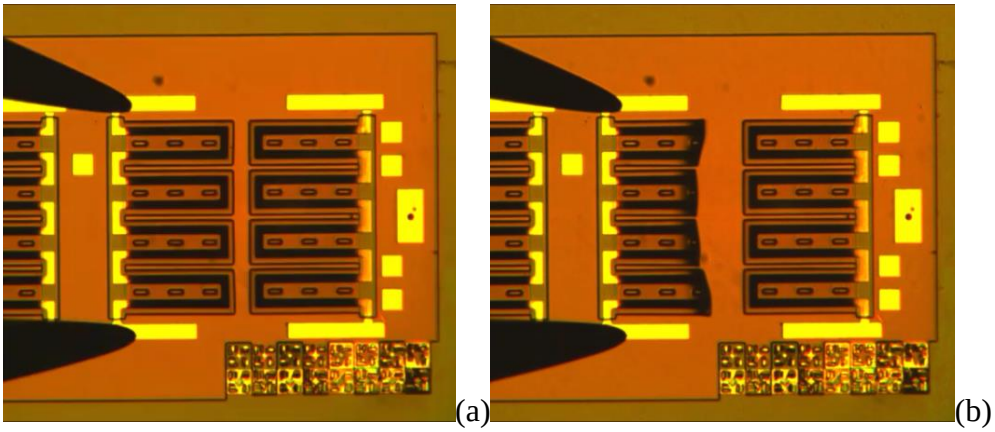
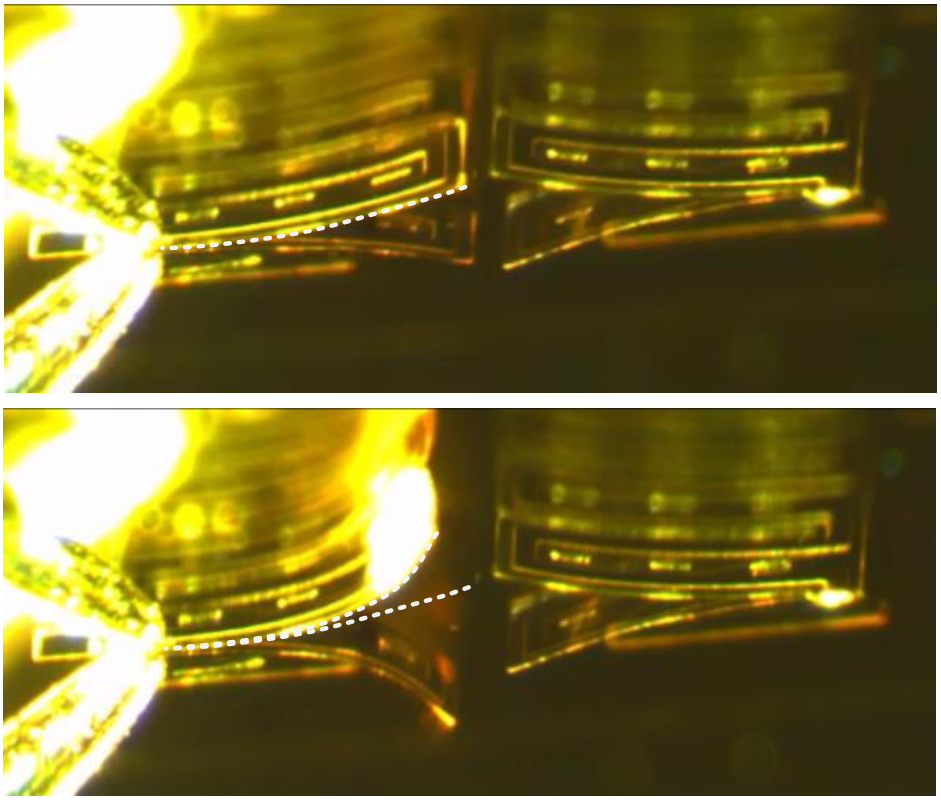

(a) (b)

Figure 5.5. (a) Top view of actuators before actuation. (b) Top view of actuators during actuation at 90 mA. The electrode on the anchor part of the legs is gold color as expected. However, the electrode on the releasing part of the legs is dark because that part of the legs is not flat and does not reflect light back vertically.


(a)

(b)

Figure 5.6. There are two images in each photo. The bottom one is reflection of the actuators on the flat mirror substrate. White dashed lines indicate the edge of the legs. (a) Side view of actuators before actuation. (b) Side view of actuators during actuation at 90 mA. The lower white dashed line is a copy from the top figure for reference.

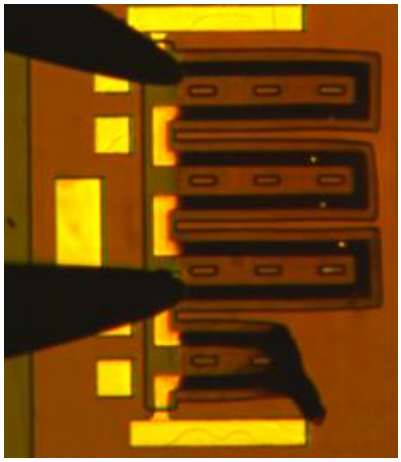


Figure 5.7. The fourth leg was already completely damaged by passing a too large current in a test before the photo was taken. The top three legs also later deformed due to passing a large current in a different experiment.

We also performed a loading test of the actuators. Bare silicon chips that are the same size as our CMOS chip were used as loads. Side views of the actuators lifting three silicon chips before and during actuation are shown in Figure 5.8. This experiment shows that the actuators can provide enough force to lift the CMOS chip.

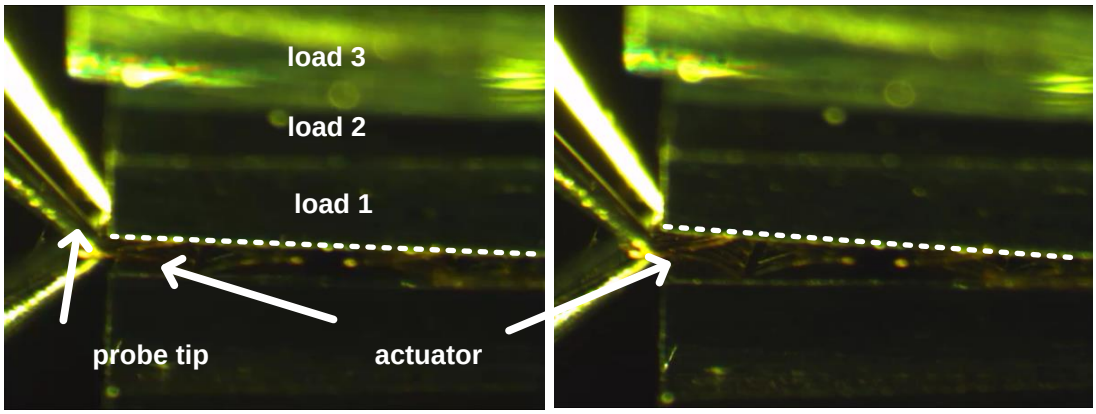


Figure 5.8. (left) Side view of actuators before actuation. (right) Side view of actuators during actuation at 85 mA. White dashed lines indicate the edge of the loading chip and show a change of angle.

We also connected the FGPLL controller chip (see Chapter 3) to the actuators to test if the CMOS chip can control the actuators properly. We first used photoresist

1813 to glue a FGPLL chip to a corner of the cavity on a DIP40 package. The chip was placed at the corner to reduce the required wire length. Then, pads that were required for this experiment were wire-bonded to the pads on the package, as shown in Figure 5.9. The pads wired out were power, ground, two control outputs for the actuators, and two control inputs for the output mode. The large pads had two wires to increase reliability and reduce connection resistance. Ivan Penskiy (Micro Robotics Lab at the University of Maryland) provided training on a wire bonder (Westbond 7476E) and did all the wire bonding. We used 1 mil diameter aluminum wire for bonding. The purpose of packaging the chip was for easier handling of the chip and connections to other devices.

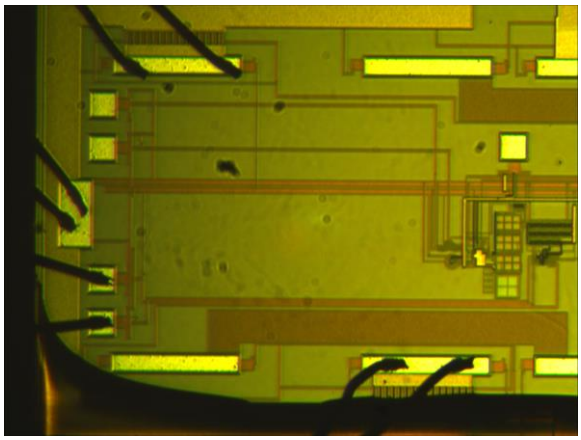


Figure 5.9. Pads on the CMOS chip are wire-bonded to pads on the package (not shown in the photo). The wired pads are (counterclockwise from top) first control output, power, two control signals to adjust the output duty cycle and control signal overlap, and second control output. Ground pad was also wire-bonded but it is not shown in the photo. On the bottom left corner is the extra photoresist glue (use as a glue) overflow to the top of the chip.

In order to connect the chip to the legs, we needed to determine the supply voltage. There was an issue associated with the determination of the supply voltage. The issue was that the supply voltage to the controller chip determines the output

current to the actuators and also modulates the output frequency. However, both output current and the output frequency need to be set to a desired value. Therefore, we had to first test the output current of the controller at different supply voltages. We found that the output current was most appropriate (not to damage the actuators but large enough to provide significant actuation) at 3.6 ± 0.1 V supply. We then programmed the controller chip (in the package) to 1 Hz on the probe station using a 3.6 V supply. Lastly, the chip was connected to the actuators through the probe station and to the power supply to test the actuation control. Two different modes of control signal driving the actuators were tested. The first control mode had 25% duty cycle and 12.5% overlap between two control signals and is shown in Figure 5.10. The second control mode had 25% duty cycle without overlap between two control signals and is shown in Figure 5.11. Since the two control signals had no overlap, the two rows of legs were not actuated simultaneously. These two experiments again verified the controller chip can be programmed to a desired frequency and can provide signals for leg control.

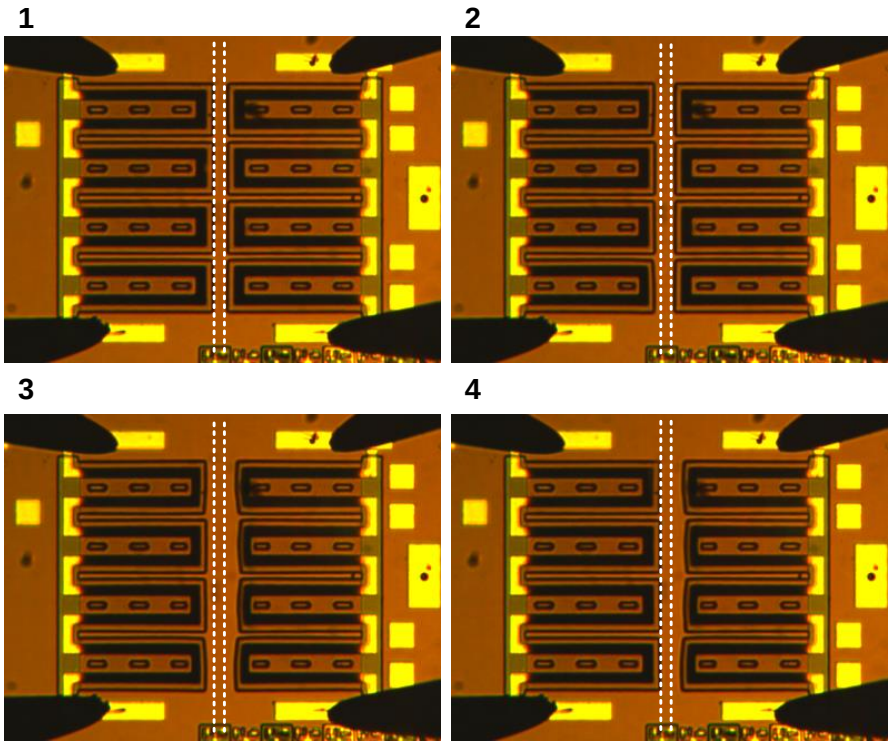


Figure 5.10. Four photos demonstrate leg control of the controller chip. White dashed lines represent the leading edges of the legs at their relaxing state and help identify if the legs are actuated. The control signal had 25% duty cycle and 12.5% overlap between the two control signals. The control signal for the left row of legs was leading the other signal. The first photo has two control signals at low level and the actuators were not actuated. The second photo had the control signal for the left row of legs high and the left legs were actuated. The third photo had both controls high. The left legs remained at previous position and right legs were actuated. The fourth photo had the control for the left legs low and the control for the right legs high. The left legs returned to the relaxed position and the right legs were actuated. After that, all the legs returned to the original position as in the photo 1.

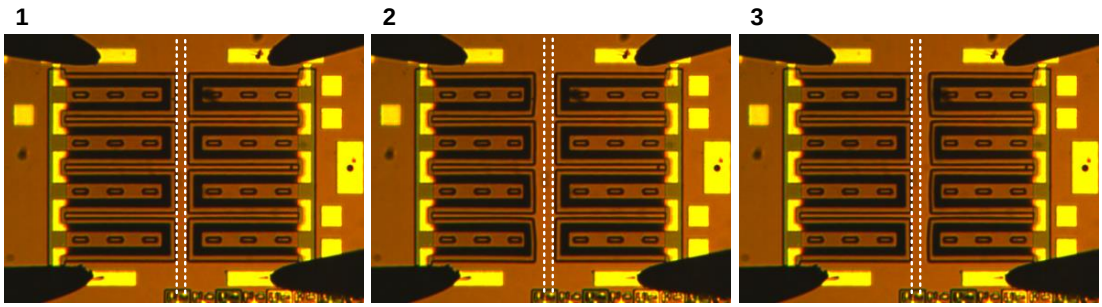


Figure 5.11. Three photos demonstrate leg control of controller chip. White dashed lines represent the leading edges of the legs at their relaxing state and help identify if the legs are actuated. The control signal had 25% duty cycle and zero overlap between the two control signals. The control signal for the left row of legs was leading the other signal. The first photo had two control signals at low level and the actuators were not actuated. The second photo had the control signal for the left row of legs high and the left legs were actuated. The third photo had the control for the left legs low and the control for the right legs high. The left legs returned to the relaxing position and the right legs were actuated. After that, all the legs returned to the position as in the photo 1.

Comparing the bending of the actuators driven by the controller chip (experiments of Figure 5.10 and Figure 5.11) with that driven by the controlled current source (experiments similar to Figure 5.5), we estimated the output current of the controller chip delivered to one row of the actuators to be 80 mA. Given the 3.6 V supply voltage, the actuation power for one row of the legs is 288 mW. Assuming a constant current of 80 mA during actuation, the maximum instant power for the chip is 576 mW when the controls have overlap and 288 mW when the controls have no overlap. The average power would have to account for the duty cycle of the control signals, for example 72 mW for 12.5% and 144 mW for 25% duty cycle, respectively. The voltage drop across one row of the legs estimated using median resistance 20.80 Ω for the legs is 1.66 V (80 mA × 20.80 Ω). This voltage indicates that the voltage drop across the drivers on the chip is 1.94 V. Therefore, in order to drive two rows of legs in series the supply voltage has to be 5.26 V (1.94 V + 2×1.66 V) assuming the

driver needs the same 1.94 V to function. This supply voltage is still within the safe operating voltage (7-8 V) for the chips. However, it shows that the resistance of the legs has to be well controlled in a range according to the controller design, so that the output current of the chips is large enough to generate reasonable actuation while not too large to damage the legs.

Chapter 6: Low Power Computation for Robotic Control

6.1 Motion Planning Using Randomized Receding Horizon Control

For robots to autonomously complete the tasks that they are assigned, they must have an adequate level of control. This control is normally based on performing designed computations and algorithms. There are many robotic functions that can potentially be implemented like deduction, planning, perception, learning, motion, and creativity [127]. Not all of them are possible or necessary to be implemented on a tiny robotic platform which has strict size and power constraints. Therefore, more efficient algorithms and implementation methods are required to achieve the desired level of autonomy. There are a few functions that are more essential than the others, for example motion (localization, mapping, or planning) and perception (hazardous gas or defects detection). Among them we especially consider motion planning as one of the most important traits because the tiny robots must "move" to some specific locations to perform their tasks.

Motion planning of tiny robots usually involves a multivariable nonlinear dynamic system, constraints on the inputs and the outputs, and online processing. Receding horizon control (RHC) or model predictive control [128] is able to handle changes in system parameters and constraints as well as be easily applied to large and multivariable processes [129, 130]. It is also highly adaptable to changing environments [131]. These advantages make it a promising control method for tiny robots.

RHC is an iterative control strategy minimizing a cost function over finite time horizon. At time step k the current state is observed and a finite horizon optimization problem from k to $k + N$ (N is number of steps) is solved. The computation of this problem requires a behavior model for the system dynamics to predict the new system state caused by changes in the input. Only the first step of the solution for control inputs is executed. Then, the horizon advances one step to $k + 1$. Again, the state is observed and the optimization is computed yielding a new solution whose first step is executed. The repetitive operation is shown in Figure 6.1.

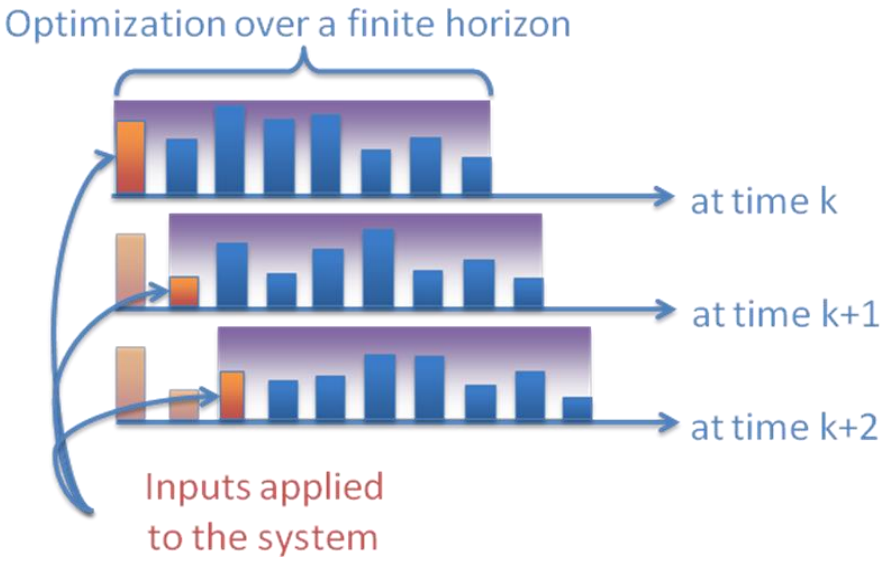


Figure 6.1. Demonstration of RHC operation. Only the first control step is implemented. Source: Eduardo Arvelo and Nuno Martins, UMD.

More detailed formulation is as follows [131]. The system behavior can be described as

$$X_{k+1} = f(X_k, U_k) \tag{6.1}$$

where X_k is the system state vector at step k, U_k is the control vector at step k, and f is the model to predict next system state. X_k and U_k are subject to constraints $\bar{X}_k$ and

$\bar{U}_k$ respectively. A control sequence $\mathbf{U}_k = [U_{k|1}, \cdots, U_{k|N}]$ is computed. After the system states have been predicted for *N* steps $\mathbf{X}_k = [X_{k|1}, \cdots, X_{k|N}]$, the cost function is given by

$$C_N(k) = \sum_{i=1}^{N} g(X_{k|i}, U_{k|i}) \tag{6.2}$$

where $C_N(k)$ is the accumulate cost, *g* is the cost at a specific step, $X_{k|i}$ and $U_{k|i}$ are the *i*-th next predicted state and computed control at step *k*.

For a system with complex dynamics, solving the finite horizon optimization in real-time is impossible or impractical. Many methods have been developed to solve this issue but only few of them can be handled by a state of the art μC [132, 133]. Randomized RHC (RRHC) has been proposed by Tanner and Piovesan to offer an attractive tradeoff between performance and computation time [134]. Instead of trying to find an optimal control, it generates multiple control sequence candidates $\mathbf{U}_k^j$ and picks the best among them

$$\mathbf{U}_k = \arg\min_{\mathbf{U}_k^j} \sum_{i=1}^{N} g(X_{k|i}^j, U_{k|i}^j) \tag{6.3}$$

where *j* represents the *j*-th candidates. Although the randomized search approach can only find a feasible solution, the stability requirement can be satisfied by adding an extra constraint to the cost function [135].

A hardware architecture has been reported to perform RRHC for a differential drive robot by Kuhlman et al. [131]. Based on this architecture we developed a modified system architecture as shown in Figure 6.2. The control generator maps the random samples ***R*** generated by random number generator (RNG) to useful control

commands. A system dynamics simulator computes future states based on control candidate $\mathbf{U}_k^j$ that satisfies the constraints $\bar{U}_k$ and observed current state X_k. Predicted states $\mathbf{X}_k^j$ and control candidate $\mathbf{U}_k^j$ are used to compute cost and determine the suboptimal control $\mathbf{U}_k$. The system observer provides observation of the current state X_k. A sensor is responsible for providing constraints according to environmental changes.

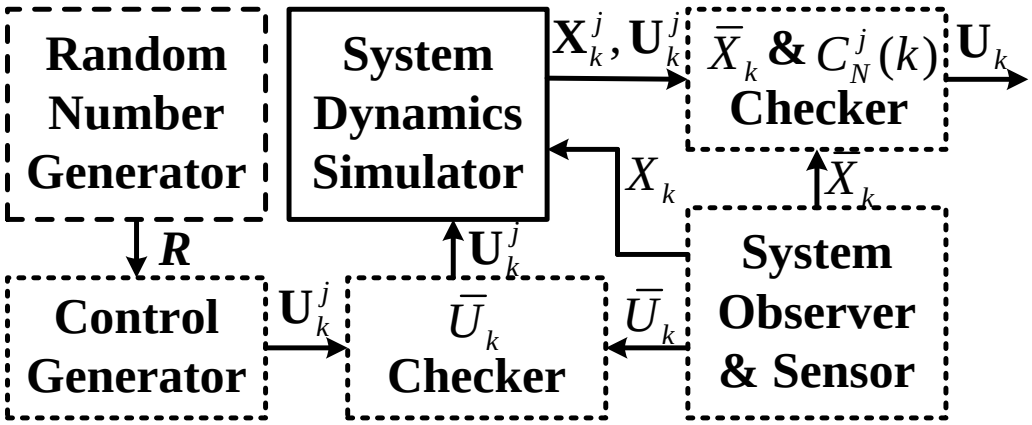


Figure 6.2. System architecture for RRHC. The block with solid lines is the focus of this research.

Analyzing the functional blocks in the RRHC system, the system dynamics simulator is the most complex and requires the highest computational power. We implemented this circuit which will be discussed in the following sections. The circuit utilizes a mixed-signal approach to meet strict size and power constraints for tiny robots. In our point of view, successful implementation of this example is representative enough to show the suitability of mixed-signal circuit for demonstrating designed intelligent behavior of a tiny robot.

6.2 *Implementation Example: System Dynamics Simulator – Odometry*[*]

6.2.1 Introduction

The odometry circuit was originally developed by Mr. Michael Kuhlman, who defined the problem, designed the preliminary mixed-signal circuits, and performed the error analysis. He initiated this circuit as his final project in a graduate circuit class instructed by Dr. Timothy Horiuchi at the University of Maryland. I later modified the circuit architecture, improved the performance (specifically reduced the error to an acceptable level), implemented a digital counterpart for comparison, and performed data analysis to understand the experiments.

Odometry is a methodology that uses robot motions to estimate the change in position and orientation. Autonomous operation generally requires local modeling of system dynamics, since many control strategies require knowledge of the system state and direct real-time sensing of position is not always possible due to size and power constraints. RRHC requires many predictions of the robot's future state before executing a given command sequence because the quality of the solution depends on the number of candidates tested. Theoretically, the more candidates evaluated the more likely the solution is close to the optimal one. Digital implementation of odometry using either general purpose or specific circuits is relatively large, power-hungry, and slow. To alleviate this challenge we proposed a mixed-signal architecture implementing an odometry function that maps motor commands to estimated changes in position using a kinematic model.

[*] Most of the material in this section was originally published as " M. J. Kuhlman, T.-H. Lee, and P. A. Abshire, "Mixed-signal odometry for mobile robotics," in *Proc. SPIE8725, Micro- and Nanotechnology Sensors, Systems, and Applications V*, 2013." © SPIE.

6.2.2 System Overview

The odometry circuit is designed to support control of a differential-drive robot as shown in Figure 6.3. The system state vector X is $(x(t), y(t), \theta(t))$ and control vector U is $(v(t), \omega(t))$ or $(u_R(t), u_L(t))$ which are linearly related assuming a kinematic model. The output and input relationship is defined by nonlinear differential equations of the kinematic model

$$\begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{\theta}(t) \end{bmatrix} = \begin{bmatrix} v(t)\cos\theta(t) \\ v(t)\sin\theta(t) \\ \omega(t) \end{bmatrix}. \tag{6.4}$$

As RRHC has piecewise constant control $(v(k), \omega(k))$, the discretized closed form solution can be found

$$\begin{bmatrix} x(k+1) \\ y(k+1) \\ \theta(k+1) \end{bmatrix} = \begin{bmatrix} \frac{v(k)}{\omega(k)}(\sin\theta(k+1) - \sin\theta(k)) + x(k) \\ \frac{v(k)}{\omega(k)}(\cos\theta(k) - \cos\theta(k+1)) + y(k) \\ \omega(k)T + \theta(k) \end{bmatrix}. \tag{6.5}$$

where T is the step size.

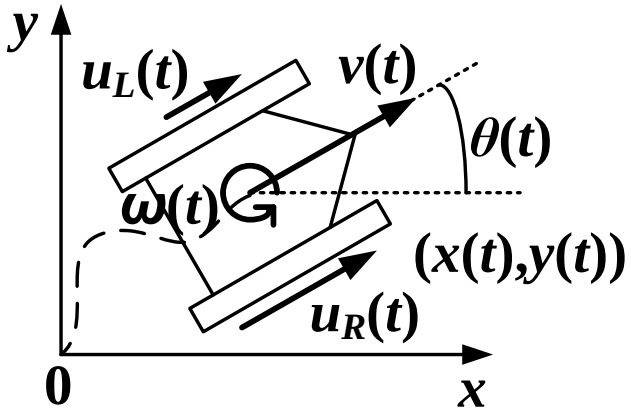


Figure 6.3. Robot state space (x,y,θ) and control space (v, ω) (or (u_R,u_L)). x, y: position in Euclidean space; θ: bearing of the robot; v: robot velocity; w: robot angular velocity; u_R, u_L: control commands to right and left actuator respectively. This figure was modified from Mr. Kuhlman's figure [48].

We designed a mixed-signal circuit directly implementing Equation 6.4. In comparison with Equation 6.5, it has much fewer computations and does not need a divider that is relatively difficult in circuit implementation. The computation requires three integrators, one cosine function, one sine function, and two multipliers. Signal flow and configuration of these components are shown in Figure 6.4.

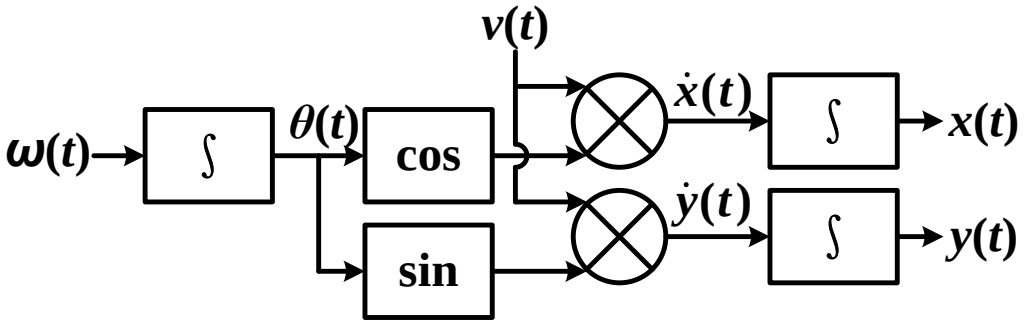


Figure 6.4. Block diagram of components computing Equation 6.4. This figure was modified from Mr. Kuhlman's diagram [48]. Blocks with the ∫ symbol are integrators.

6.2.3 State Control for $\theta(t)$ Integration

Due to the limited dynamic range of state variables in a real circuit, the representation of $\theta(t)$ needs to be carefully considered. For $\theta(t)$ we define -2π to 2π as mapping linearly to the range 1 V to 3 V where two complete periods can avoid discontinuities at the boundaries. Mr. Kuhlman's original design used two modulus circuits to bring the voltage at $\theta(t)$ back to 2 V (0 radian) when it exceed a defined range in either direction. However, sudden raising or dropping of voltage caused strong voltage coupling to adjacent nodes and resulted in significant error. Instead, we designed an integrator with two states whose system diagram is in Figure 6.5.

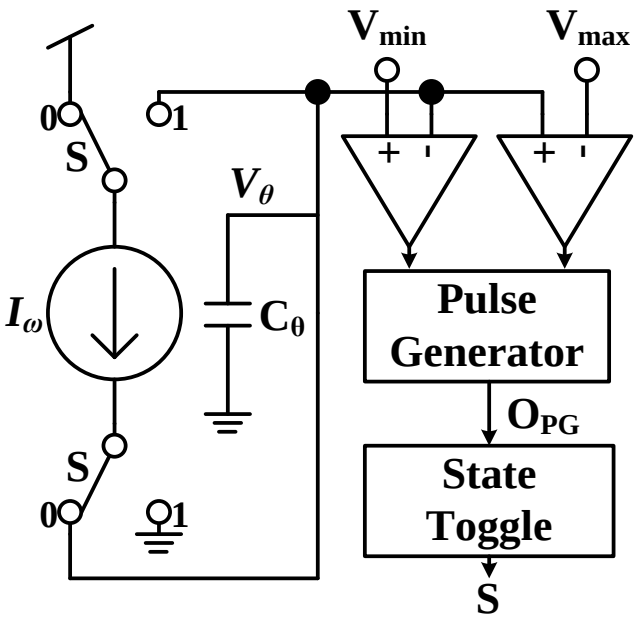


Figure 6.5. System diagram of the integrator and state machine. The state machine consists of two comparators, a pulse generator, and a state toggle circuit.

The two states of operation are summarized in Table 6.1. In state S = 0, both switches (controlled by state) flip to the left and I_ω is sourced to capacitor C_θ. Here we assume that the robot can only rotate in one direction. Therefore, I_ω is always greater than or equal to zero and V_θ is monotonically increasing and decreasing in state S = 0 and state S = 1, respectively. In state S = 1, both switches flip to the right and I_ω is drained from capacitor C_θ. In this state, a mapping is required to compute the correct θ; the mapping is done by fliping the differential output current of the multiplier circuit (see Table 6.1).

Table 6.1 States of operation (θ': angle translated from V_θ; N: integer)

State	Desired mapping	sine	cosine
S = 0	$\theta = 2N\pi + \theta'$	$\sin\theta = \sin\theta'$	$\cos\theta = \cos\theta'$
S = 1	$\theta = 2N\pi - \theta'$	$\sin\theta = -\sin\theta'$	$\cos\theta = \cos\theta'$

The state machine works as follows. The integrated voltage V_θ is compared to two thresholds using comparators. When V_θ crosses a threshold (V_{min} or V_{max}), the output of the corresponding comparator becomes high. This rising signal is fed into a pulse generating circuit, converting the rising signal into a pulse. The rising edge of the pulse toggles the state as shown in an example waveform in Figure 6.6.

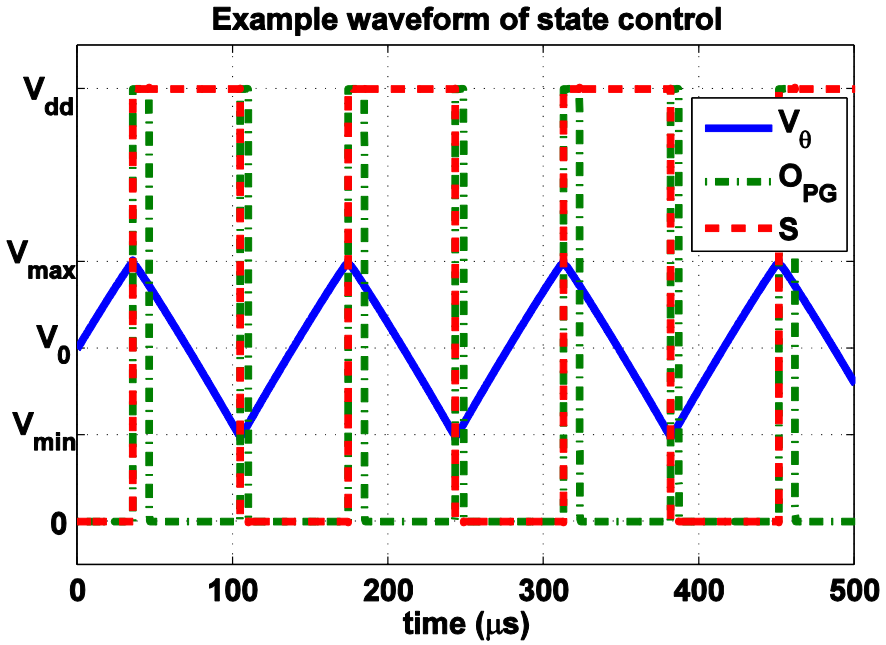


Figure 6.6. Example waveform of state control. I_ω remains constant and V_θ increases or decreases linearly.

6.2.4 Sine and Cosine Function Circuits

The sine function circuit performs $I_o = I_B \cdot \sin(V_\theta)$ where I_o is the differential output current, I_B is a amplitude constant, and V_θ is the input voltage. A sine function has a series expansion using hyperbolic tangent [136]

$$\beta \sin(\frac{\pi x}{\alpha}) \approx \lim_{m\to\infty} \sum_{k=-m}^{m} (-1)^k \tanh(x + k\alpha) \tag{6.6}$$

where α and β are constants. A tanh function can be obtained in a CMOS circuit having a differential pair operate weak inversion region and the output current is

$$I_{diff} = I_B \tanh(\lambda(V_{i1} - V_{i2})). \tag{6.7}$$

We defined α in Equation 6.6 as $\lambda/2$. For the input range to be 1 to 3 V, x and m were designed to be $\lambda(V_\theta - 2)$ and 2 respectively. Then, Equation 6.6 can be rewritten as

$$\sum_{k=-2}^{2} (-1)^k \tanh(\lambda(V_\theta - (2 - \frac{k}{2}))) \approx \beta \sin(\frac{\pi\lambda(V_\theta - 2)}{\lambda/2}) = \beta \sin(2\pi V_\theta). \tag{6.8}$$

Compared to Equation 6.7, we have V_{i1} as our input voltage V_θ and also V_{i2} as 1, 1.5, 2, 2.5 and 3 V (k =-2,-1,0,1,2) for each differential pair respectively. The circuit design is based on a reported implementation by Fried and Enz [137] with two major modifications. First, the biasing voltages were generated by a resistive voltage divider. Second, source degeneration was used in each differential pair to increase linear range. These modifications were made by Mr. Kuhlman and remained in the final design.

The final design of the sine shaping circuit is shown in Figure 6.7. Simulation results using a BSIM3.3 model of a commercially available 0.5 μm 2P3M CMOS technology in PSPICE are shown in Figure 6.8. The calculated mean and root mean square of the absolute error compared to an ideal sine wave are 4.50 nA and 5.29 nA respectively.

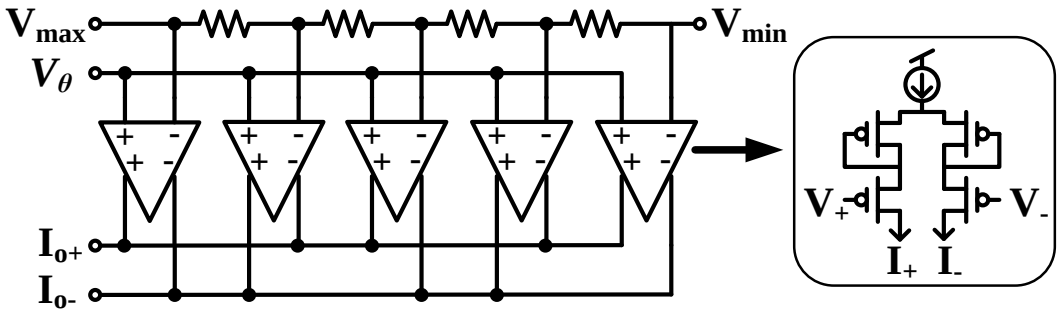


Figure 6.7. Five differential pairs performing Equation 6.8. A resistive network provides the biases. This figure was modified from Mr. Kuhlman's figure [48].

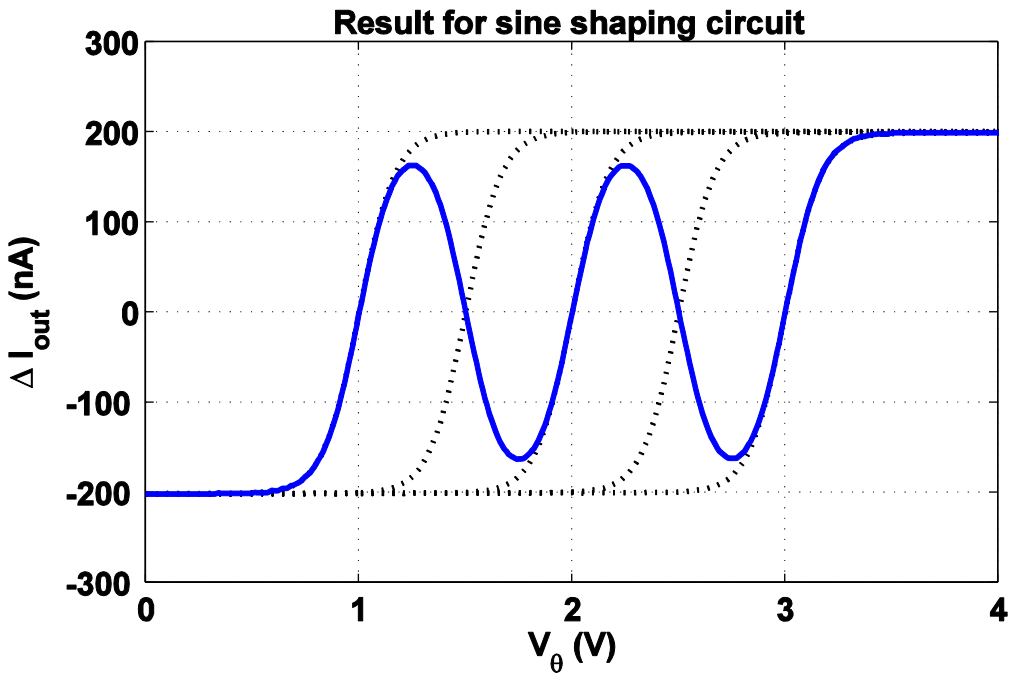


Figure 6.8. Simulation results of the sine function circuit. Blue trace is the differential output of the circuit and black dashed traces are outputs of individual differential pairs.

The cosine function circuits can be designed similarly with a phase shift of the sine function. We defined m, x and α to be 3, $\lambda(V_\theta - 1.75)$ and $\lambda/2$ respectively, yielding

$$\sum_{k=-3}^{3}(-1)^k \tanh(\lambda(V_\theta - (1.75 - \frac{k}{2}))) \approx \beta \sin(2\pi(V_\theta - 1.75)) = \beta \cos(2\pi V_\theta)\,. \qquad (6.9)$$

6.2.5 Multiplier Circuits and *x*, *y* Integrators

The multiplier circuit design is implemented using a four quadrant translinear Gilbert cell as shown in Figure 6.9 [138]. The choice of using Gilbert cell as the multiplier was made by Mr. Kuhlman. The differential current output $I_+ - I_-$ equals $\Theta \cdot S_P \cdot I_B$ if all NMOSs operate in weak inversion region and second order effects like channel length modulation are ignored. We defined S_P as $2 \cdot I_v/I_B$ so S_P is a ratio

and normally takes values between -1 and 1. S_P is proportional to the speed $v(t)$ (or I_v) and Θ is proportional to the output of trigonometric function circuits. We took V_B in Figure 6.9 to be 0.75 V which was optimized through simulations to match the source voltages of those NMOSs with their gates connected together.

One modification was made to the output stage of the multiplier on the sine path to accommodate the different operations of the two state controls summarized in Table 6.1. The output stages of both multipliers are shown in Figure 6.10. Figure 6.10 (a) is the output stage for the cosine path. Cascode PMOS and NMOS current mirrors are used to reduce the mismatch when duplicating the current. For Figure 6.10 (b) the current injected into the capacitor is $I_+ - I_-$ or $I_- - I_+$ depending the state S equaling to 0 or 1, respectively; the figure shows the status when S=1. Four unity gain buffers are used to minimize the error due to state change. The error is caused by charge sharing between node O and either O_1 or O_2. For instance, when S=1, O is connected to O_2 and their voltages are equal as shown in Figure 6.10 (b). However, if the unity gain buffer does not exist, voltage at O_1 depends on the I_+ and I_- current sources on that branch and can be a few volts different from the voltage at O. When the state S changes to 0, O is disconnected from O_2 and connected to O_1. The voltage difference between O and O_1 will change the voltage at O instantly and causes error. The unity gain buffer equalizes their voltages while they are not connected. When state S changes, only minimum charge sharing would occur. The unity gain buffer is biased at a current that is low to save power but is larger than the difference between any possible I_+ and I_- so it is still capable of controlling the voltage.

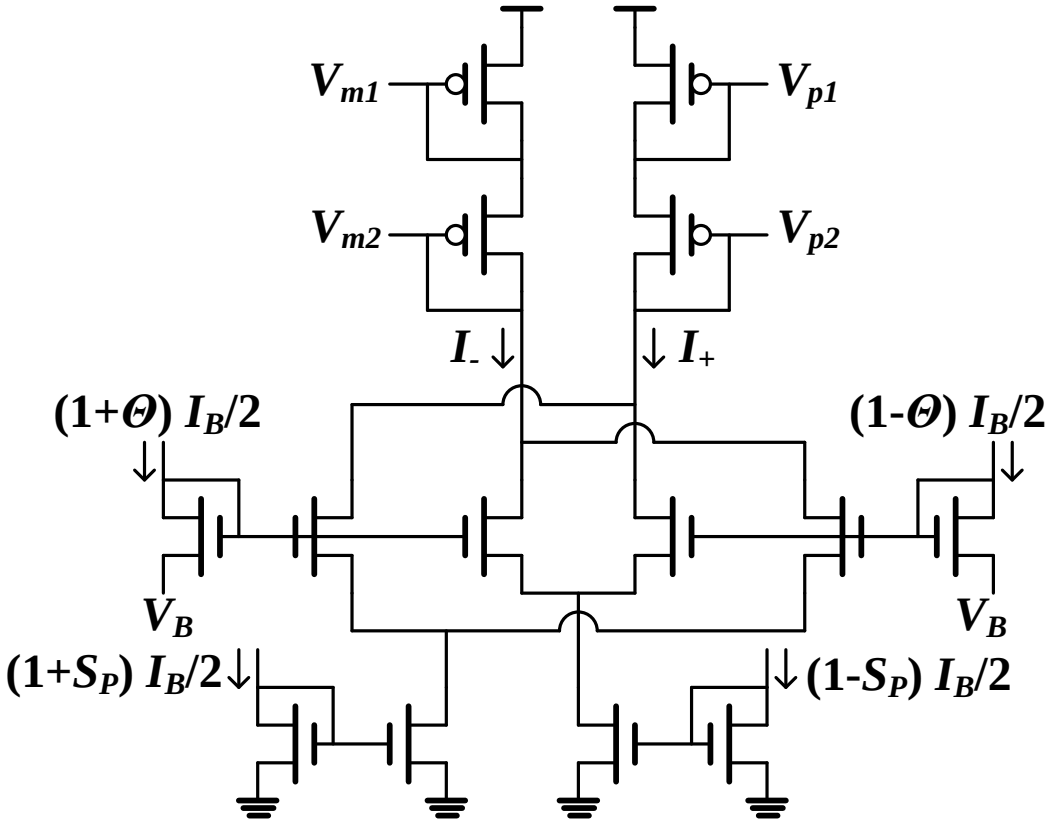
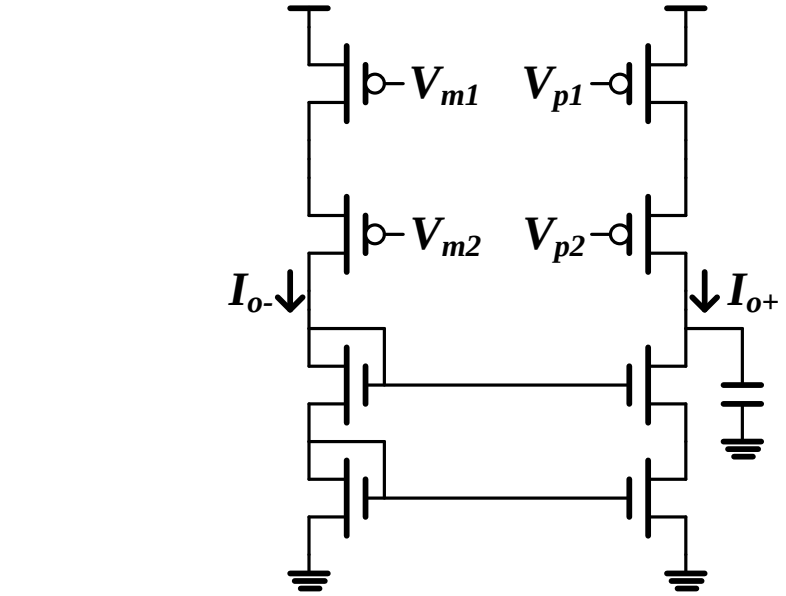
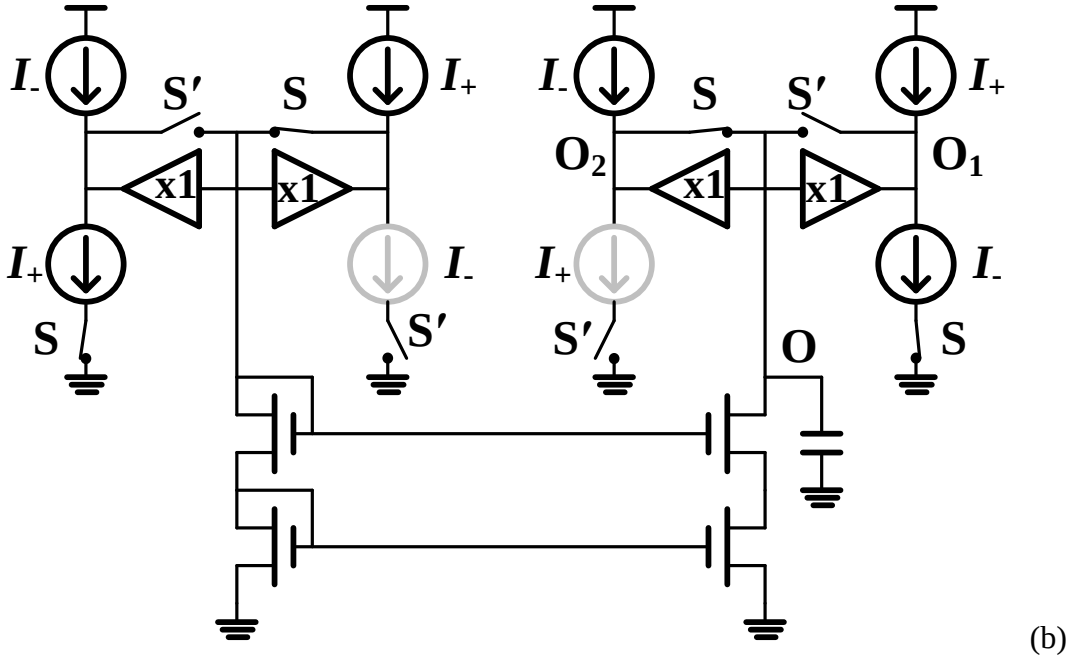


Figure 6.9. Circuit topology of Gilbert multiplier cell. S_P is proportional to $v(t)$ and Θ is proportional to the output of trigonometric function circuits.



(a)



(b)

Figure 6.10. (a) Output stage of one multiplier producing differential output current on the cosine path. (b) Output stage of the other multiplier for the sine path when S=1. Faded current sources are disconnected at this control S. Four NMOSs and one capacitor represent the integrator.

6.2.6 Digital Application Specific Integrated Circuit Implementation

To perform comprehensive comparisons between different types of implementations, we also designed a digital ASIC which computes the closed form odometry solutions in Equation 6.5. The number of bits for inputs, outputs, and internal nodes were chosen based on the requirements of dynamic range and precision for this application; v, ω, x, y, and θ are 8, 8, 12, 12, and 8 bits respectively. The sine and cosine functions are generated using a lookup table which contains the sine and cosine values from 0 to $\pi/2$ (64 elements for each in this case). The other values are calculated by exploiting the symmetry of these functions. Since the requirement for operation time is not strict here, hardware sharing technique was applied to reduce area and hopefully the leakage power. As a result, the computations in Equation 6.5 can be carried out in two clock cycles with only one multiplier. The design was implemented in Verilog HDL. Synthesis was done by Cadence Encounter RTL Compiler using 0.5 μm 2P3M technology. The reported area is 0.70 mm^2 and the highest speed is 61.8 MHz. These values could be improved if the system had access to read-only memory (ROM) to implement the lookup table instead of using combinational logic circuits.

6.2.7 Simulation Results

We simulated the circuit using BSIM3.3 model of a commercially available 0.5 μm 2P3M CMOS technology in PSPICE. Three different types of velocity and angular velocity conditions were tested. Speed $v(t)$ is controlled by lower differential input current in the multiplier (see Figure 6.9) where we defined S_P as $2 \cdot I_V/I_B$. Angular velocity $\omega(t)$ is controlled by I_ω in the θ integrator (see Figure 6.5). The

origin of Euclidean space (x, y) = (0,0) is defined as (2.5 V, 2.5 V) and the $\theta(t)$ range of $[-2\pi,2\pi]$ as [1 V, 3 V].

The first simulation is constant v and zero ω. In this simulation the robot is supposed to move linearly with time in one direction (x direction in this case because initial θ is zero radian). The result is shown in Figure 6.11 demonstrating that the travel distance is linearly proportional to the velocity. This shows that the integrator is working nearly perfectly with zero angle. The second simulation has both constant non-zero velocity and angular velocity. The robot should ideally circle at constant radius. The results are shown in Figure 6.12 – 6.14. In the last simulation we used zero I_v and zero I_ω. Ideally V_θ, V_x, and V_y should stay constant but we found that after 2 ms only V_θ stays unchanged, V_x, and V_y drifted -58.1 mV and 0.3 mV, respectively. One thing to note here is that the capacitor used in the simulation was ideal. In a physical implementation, capacitors would have finite resistance and introduce leakage. This leakage will depend on the materials and the structures of the capacitors.

A few issues still needed to be resolved and investigated including drifting and spikes in the figures. The spikes were due to the toggles of state S. Quick state change is necessary to reduce error due to the state ambiguity during transition. Unfortunately, this quick transition couples to other nodes in the circuit and causes problems. The errors also came from the non-ideal sine/cosine shaping circuits and integrators. An improved sine function circuit will be discussed in the next section. One non-ideality of the integrator is imperfect current mirrors. When we copy the currents from the multipliers to the current sources in the integrators and from the left NMOS branch to the right NMOS branch, second order effects like channel length

modulation make the copy imperfect. The other non-ideality is the undesired current injection from the unity-gain buffers. In upper right half of Figure 6.10 (b), while the I_- is charging node O (O_2 connects to O), the second unity gain buffer from the right, although its input tied to its output, still sinks or sources some current from node O and introduces error.

Although the simulation results shown in Figure 6.12, Figure 6.13, and Figure 6.14 seem inaccurate, this simulation condition might not be practical and might not be used in a real case. If the warping factor is set to 10^6, each simulation represents 400 seconds in real time. This might be much longer than required. If the robots can occasionally acquire their own accurate position through communicating with a base station or onboard sensors, the positions can be updated, say, every 5 seconds. In that case, between each update we are only looking a small portion of the circles in the figures and the errors do not accumulate as drastically.

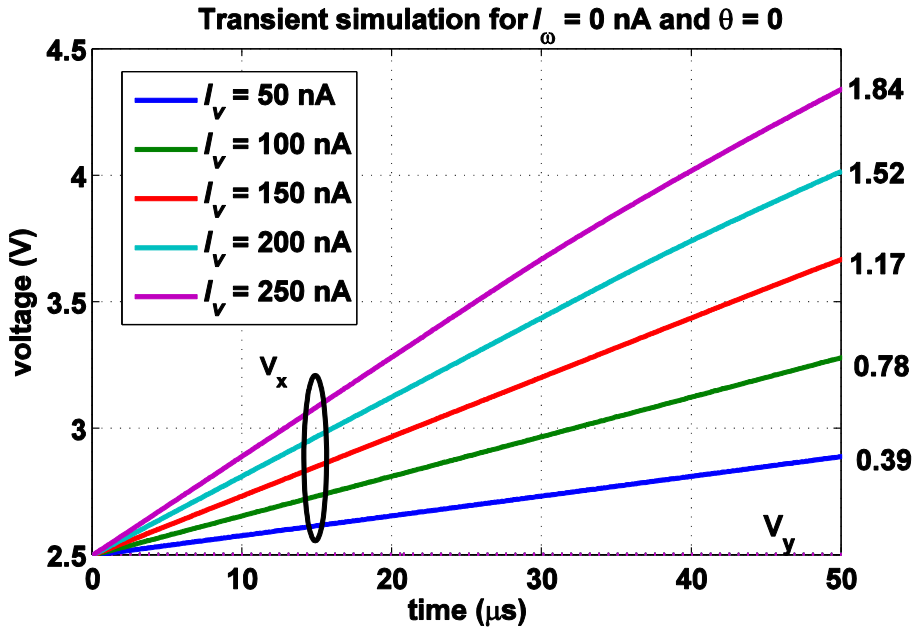


Figure 6.11. Simulation results for x and y with zero I_ω and five constant I_v. Numbers on the right indicate differences between the initial and final point.

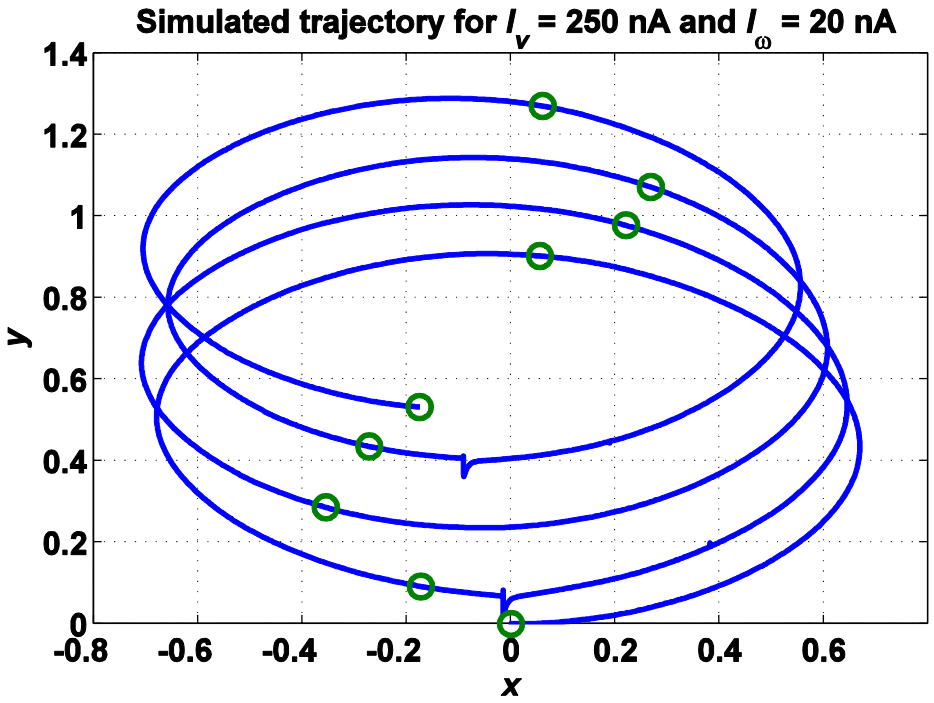


Figure 6.12. Simulated trajectory for I_v and I_ω of 250 nA and 20 nA respectively. Simulation starts at initial position (0,0). Green circles mark every 50 μs.

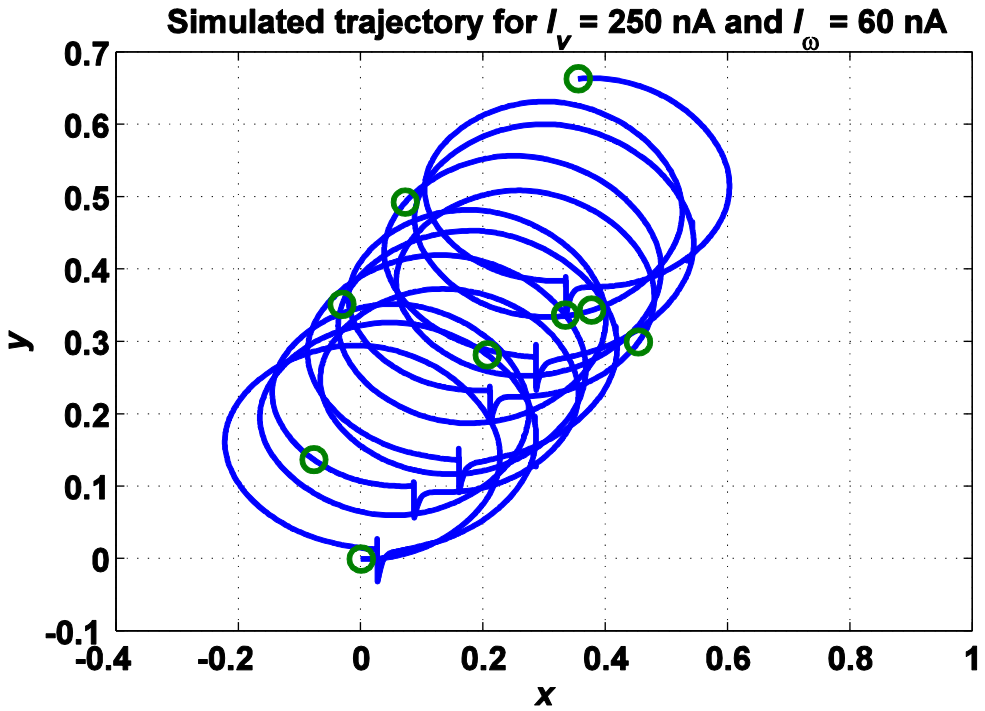


Figure 6.13. Simulated trajectory for I_v and I_ω of 250 nA and 60 nA respectively. Simulation starts at initial position (0,0). Green circles mark every 50 μs.

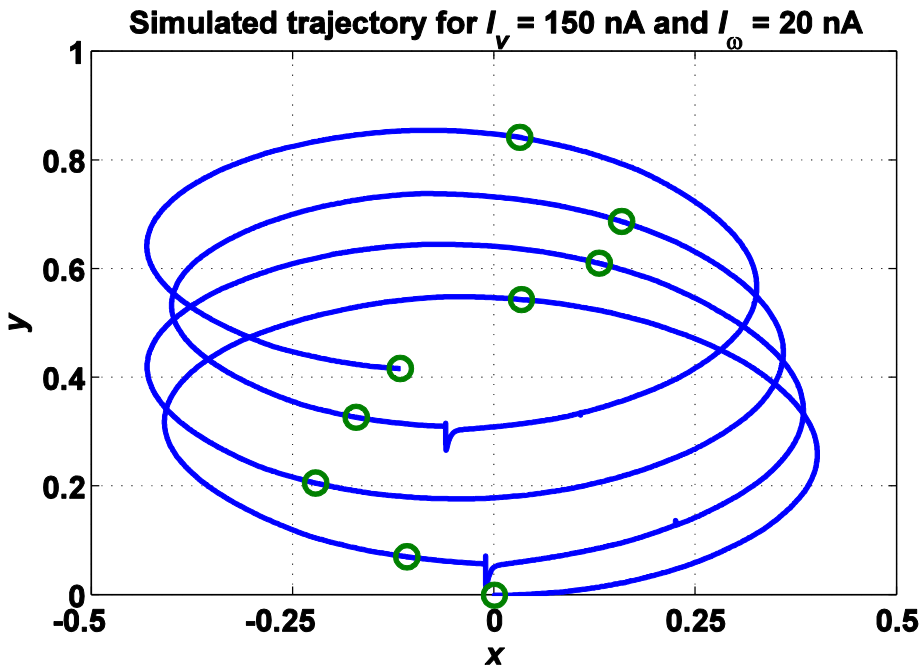


Figure 6.14. Simulated trajectory for I_v and I_ω of 150 nA and 20 nA respectively. Simulation starts at initial position (0,0). Green circles mark every 50 μs.

We examined the power consumption of the analog circuit for each circuit component. The static current for each component in odometry is listed in Table 6.2. The dynamic current that occurs when the state changes depends greatly on operation and is not included here. We compare this with the power required to compute Equation 6.5 using a digital microcontroller and a custom designed digital ASIC. We assume that computations are implemented on a microcontroller comparable to the TI MSP430 with a hardware multiplier. Equation 6.5 requires 1 division, 3 multiplications, 6 additions, and 4 trigonometric functions. Approximating the trigonometric functions with piecewise linear functions accounts for one multiplication, one addition, and one condition checking for either sine or cosine function. These computations sum up to 1 division, 7 multiplications, 10 additions, and 4 condition checking operations. We assume that all the operations take 100

clock cycles ignoring memory access costs. The energy required to perform the computation is independent of the clock, but assuming 100 μA/MHz (similar to TI MSP430), V_{DD} = 3V and a 32 MHz clock, the equations of motion can be solved in 3.2 μs consuming 30 nJ of energy. The power comparison is summarized in Table 6.3.

Table 6.2 Static current draw by each component in the odometry circuit

Component	*N* elements	Current (A)
Gilbert multiplier	2	2 μ, 3.8 μ
sine function	1	1.2 μ
cosine function	1	1.4 μ
state controller	1	10 μ
Total	5	18.4 μ
Power at 5 V		92 μW

Table 6.3 Power comparison between different implementations and design settings

Implementation	Design setting		Energy / operation (nJ)
Mixed-signal	Warp	10^4	9.2
		10^5	0.92
		10^6	0.092
Digital ASIC	Clock rate (MHz)	1	5.42
		10	14.8
μC (TI MSP430)	Power (μA/MHz)	100	30 (estimation)

6.3 *Improved Sine Shaper*[*]

6.3.1 Introduction

The sine shaper we have discussed in section 6.2.4 was experimentally optimized without using a systematic approach. However, we have found that its accuracy greatly affects the overall performance of the odometry circuit (Figure 6.12 - Figure 6.14) and its input range is limited (only a few hundred mV). This finding motivated us to develop a systematic approach to better design the sine shaper so that its accuracy can be improved and its input range can be selected arbitrarily.

Sine shaping circuits are important building blocks for many applications. Their uses can be divided into two major categories. First, they are often used to map an angular input signal to a corresponding sine output. The applications include analog computation [48, 139] and numerically controlled oscillators (NCOs) used in baseband communications systems [140]. Second, they are used with a triangle wave or modulated signal input in order to generate a sine wave output in applications like function generators [141], phase-locked loops, direct digital frequency synthesizer (DDFS) [142], radio frequency phased array [143], and on-chip testing [144].

There are many previous works implementing sine shaping circuits using digital and analog approaches. The most popular digital techniques are: lookup table, lookup table plus interpolation, coordinate rotation digital computer (CORDIC) algorithm, and infinite impulse response filter [145]. These approaches suffer from low accuracy, large power consumption, and large chip area. On the other hand,

[*] The material in this section was originally published as " T.-H. Lee and P. A. Abshire, "Accurate, wide range analog sine shaping circuits," in Proc. *IEEE International New Circuits and Systems Conference (NEWCAS)*, 2014, pp. 285-288." © 2014 IEEE.

analog sine shaping circuits usually offer better tradeoffs among power, area, and accuracy [48, 142].

In analog circuits there is no component whose transfer function is naturally a perfect sine function. Fortunately, some techniques have been proposed to approximate a sine function so that it can be implemented in both complementary metal-oxide-semiconductor (CMOS) and bipolar junction transistor (BJT) technologies. Four techniques that produce nearly perfect approximations are: 1) The polynomial approximation $\sin(x) \approx x(1-x^2)/(1+x^2)$ is implemented using the translinear principle by having metal-oxide-semiconductor field-effect transistors (MOSFETs) operating in weak inversion region or by using BJTs [142, 146]. However, this approximation is only valid for input ranges within half a cycle (-0.5π~0.5π) and cannot be extended easily; 2) A sine function is approximated by the first few terms of its Taylor series expansion [141, 147]. The largest input range achieved with this technique was half a cycle, the same as the first technique; 3) A sine function is approximated by using the sum of alternating sign hyperbolic tangent functions [48, 136, 137]; 4) A sine function is approximated by sum of parabolically spaced exponential functions [139]. The last two techniques are implemented with BJTs or MOSFETs operating in weak inversion. They offer an easy extension of input angular range and are preferred in some applications because the exact angle is sometimes more important than just the modulus angle.

To achieve arbitrary input scaling (i.e., arbitrary voltage range and angular range) requires some additional considerations. Some implementations are able to achieve arbitrary input angular range using the third approach, but the input voltage is

fixed to 300 mV per cycle [136, 137]. This results in incompatibility with other circuits and difficulties in system design. An input range of one volt per cycle was demonstrated in [48] but no systematic approach or analysis was given as a design guide. The fourth technique is able to achieve arbitrary input scaling [139]. However, this differential input design is impractical because it requires a negative copy of the input voltage and a negative power supply. In this work we adapted techniques for linearizing differential pairs in order to realize arbitrary input scaling and highly accurate sine shaping which is a novel contribution. We also presented detailed analysis on accuracy and efficiency. Furthermore, we generalized this approach to provide a design methodology.

6.3.2 Sine Approximation with Hypertangent

The sine function circuit in this paper is a single-ended input, differential output design. It produces an output current that is proportional to $\sin(2\pi \cdot V_\theta)$, where V_θ is the input voltage representing the turn information. The input voltage V_θ ranging from V_o-V_r to V_o+V_r is defined to represent $-0.25 \cdot N$ to $0.25 \cdot N$ turns where N is assumed to be a positive integer for simplicity. As a result, the input voltage ranges over $0.5 \cdot N$ cycles.

A. Sum of alternating sign hyperbolic tangent functions

Gilbert first introduced an approximation of the sine function using the sum of alternating sign hyperbolic tangent functions in a BJT circuit implementation [136] as Equation 6.6. A general circuit architecture that implements this approximation is shown in Figure 6.15 [48] where the transconductance amplifiers (TCAs) are numbered from $-M$ to M and the net output current $I_o = (I_{o+} - I_{o-})$ is expressed as

$$I_o = \sum_{m=-M}^{M} (-1)^m \cdot I_T(V_\theta - V_o - mV_s) \tag{6.10}$$

where I_T is the differential output current function of the TCA. Hyperbolic tangent functionality can be obtained from a CMOS differential pair operating in the weak inversion region [48, 137]. If the TCA is a simple differential pair with input p-type metal-oxide-semiconductor (PMOS) where its body is connected to its source to eliminate body effect (Figure 6.18 (a) without two resistors), the output current is

$$I_{T1}(V_i) = I_B \cdot \tanh\left(\frac{\kappa}{2U_T} V_i\right) \tag{6.11}$$

where I_B is the bias current of the TCA, κ is gate coupling coefficient, U_T is thermal voltage, and V_i is the differential input voltage. Then, substituting Equation 6.11 into Equation 6.10 and defining x and α in Equation 6.6 to be $\kappa \cdot (V_\theta - V_o)/2U_T$ and $\kappa/4U_T$ respectively, we have

$$\frac{I_o}{I_B} \approx \eta_o \sin(2\pi[V_\theta - V_o]) \tag{6.12}$$

where η_o is the amplitude of the sine function. The valid input angular range is determined by M. This sine approximation is valid within M complete cycles.

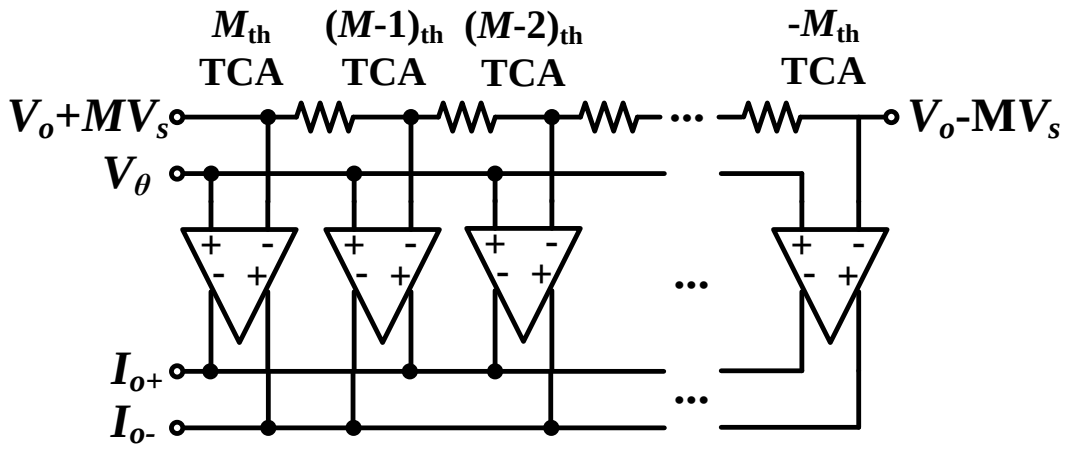


Figure 6.15. A generalized circuit architecture of Figure 6.7 to implement sine approximation using sum of alternating sign hyperbolic tangents. In the figure M is assumed to be an odd integer. If M is even, the output polarity needs to be inverted. A resistive network provides voltage references.

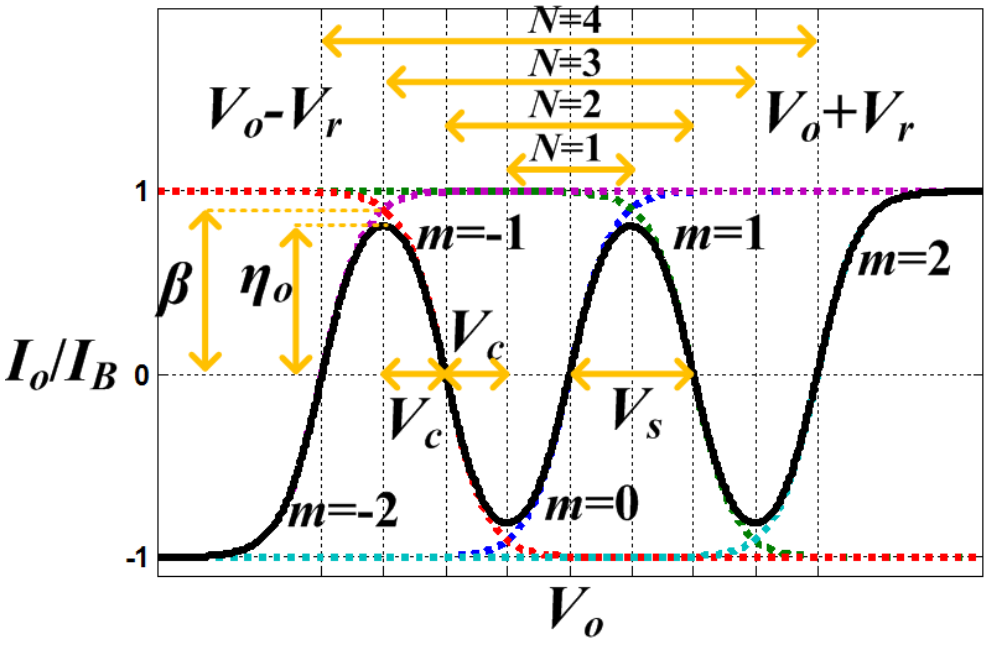


Figure 6.16. Illustration of sine approximation using five hyperbolic tangent functions (M=2) and related parameters. All currents are normalized to I_B. The solid line is sum of hyperbolic tangent functions (dashed lines).

B. Accuracy for approximation and range limitation

An illustration of the approximation is shown in Figure 6.16. Neighboring TCAs have output current crossing at $\pm\beta \cdot I_B$ with input voltage shifted by V_c from their center voltage. $V_s = 2V_c$ is the difference of center voltages between adjacent TCAs. It is apparent that β is an important factor affecting the accuracy of the approximation and taking values from 0 to 1. Utilizing relative errors to evaluate accuracy causes problems when the sine reference approaches zero. Therefore, we chose coefficient of determination (r^2) to quantify the approximation quality. The evaluation was performed over one complete cycle. For each β, we found the corresponding I_o and calculated r^2 (see Figure 6.17 for the trace with R of zero). For r^2 computation, M is assumed to be one unless otherwise specified. Making M larger than one would only affect r^2 for small β, but in this case the approximation is inaccurate and would not be used anyway. In our observation, r^2 has to be larger than 0.999 so that there is no noticeable deviation between the approximation and the reference. We set 0.9998 as our criterion for high accuracy purposes. This greatly limits the ability to cover an arbitrary input voltage range. V_c is the voltage range for 0.25 turns and can be found by solving $I_{T1}(V_c) = \beta \cdot I_B$

$$V_c = \frac{2U_T}{\kappa} \operatorname{arctanh}(\beta) \tag{6.13}$$

where U_T is 25.8 mV and κ is assumed to be 0.6. The result is shown in Figure 6.17. r^2 does not reach 0.9998 when R=0. We conclude that a simple differential pair TCA cannot achieve arbitrary input scaling.

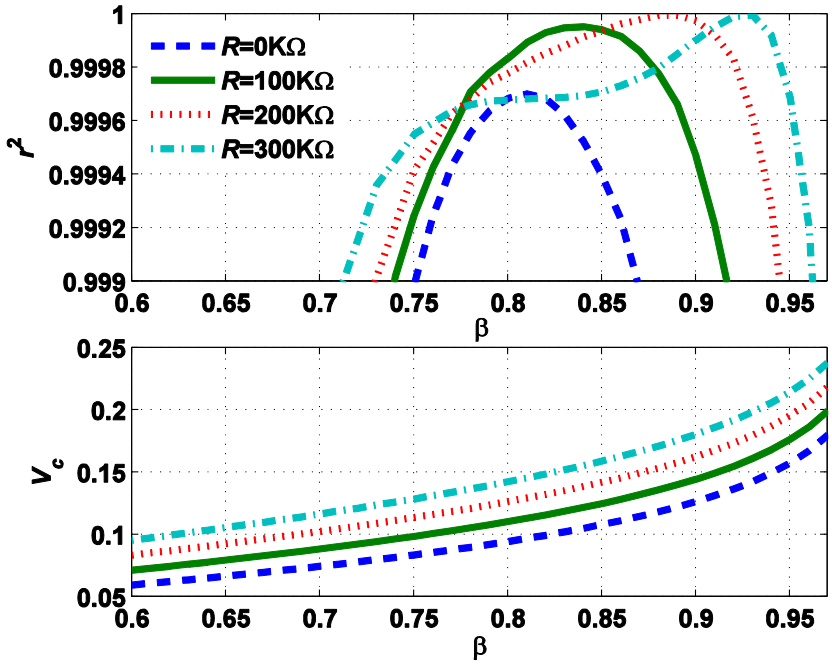


Figure 6.17. Calculated accuracy r^2 (top) and V_c (bottom) as functions of β over the range 0.6 to 0.97. r^2 exceeds 0.999 for a large range of β values when the source degeneration resistor R is larger.

6.3.3 Resistive Source Degeneration

Because a simple differential pair offers limited input voltage range, source degeneration resistors (R) are introduced in the differential pairs (Figure 6.18(a)). The output current function of this TCA becomes

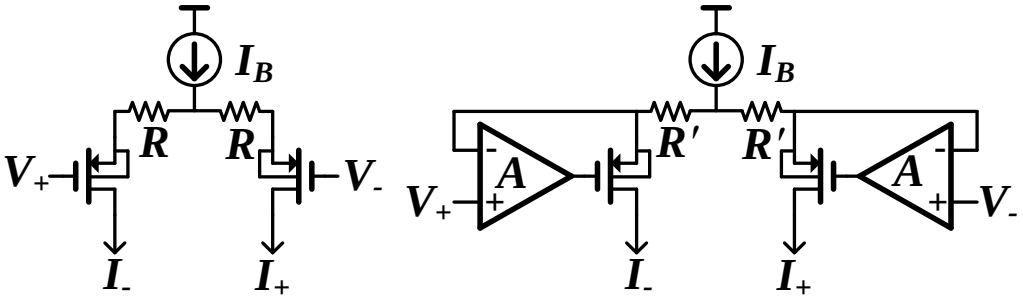
$$I_{T2}(V_i) = I_B \cdot \tanh\left(\frac{\kappa}{2U_T}[V_i - R \cdot I_{T2}(V_i)]\right). \tag{6.14}$$

I_{T2} here was solved numerically for a possible set of V_i and is not exactly a hyperbolic tangent function. I_B is selected carefully so that the PMOS does not operate in strong inversion region even if all the bias current flows in one side of the TCA. In this analysis I_B was set to 200 nA. After I_o is obtained from Equation 6.10, r^2 is calculated.

For specific R and β, the resulting V_c can be found by solving $I_{T2}(V_c)/I_B = \beta$ using Equation 6.14

$$V_c = 2U_T \cdot \operatorname{arctanh}(\beta) / \kappa + \beta \cdot I_B \cdot R \,. \tag{6.15}$$

Both r^2 and V_c are shown in Figure 6.17. The range of V_c where r^2 is larger than 0.999 is very limited. To achieve large V_c, the required resistances would be impractically large.



(a) (b)

Figure 6.18. (a) TCA with source degeneration. (b) TCA with a tunable gain buffer.

6.3.4 Transconductance Attenuation

In order to realize an arbitrary input voltage scaling, we use a circuit modified from a design that was originally proposed to increase the linearity of the TCA, shown in Figure 6.18 (b) [148]. A voltage buffer with gain A is inserted between the input and the PMOS gates. The value for A is chosen based on the input range requirements and is less than one for most cases. The output current of this TCA is solved from

$$I_{T3}(V_i) = I_B \cdot \tanh\left(\frac{\kappa}{2U_T} [A \cdot V_i - (A+1) \cdot R' \cdot I_{T3}(V_i)] \right). \tag{6.16}$$

If we set $(A+1)\cdot R'=R$ and compare Equation 6.16 with Equation 6.14, we find that $I_{T3}(V_i)=I_{T2}(A\cdot V_i)$. This means that Equation 6.16 is a scaled version of Equation 6.14 on the input voltage axis. From Equation 6.15 gain A can be designed to achieve arbitrary V_c by

$$A=\frac{V_c}{2U_T\cdot\operatorname{arctanh}(\beta)/\kappa+\beta\cdot I_B\cdot R}. \tag{6.17}$$

From r^2 calculation based on Equation 6.10 and Equation 6.14, having R between 100 KΩ to 250 KΩ yields high accuracy over a wide range of β. We chose R to be 200 KΩ. In this case, β can be chosen from 0.81 to 0.92 for r^2 larger than 0.9998 (see the top figure in Figure 6.17). In conclusion, source degeneration is used to shape I_o into a nearly perfect sine function while transconductance attenuation is used to achieve arbitrary input scaling.

Current efficiency η of the circuit is defined as the ratio of maximum output current to the sum of all bias currents [136]

$$\eta=\frac{\eta_o}{2M+1}\approx\frac{2\beta-1}{2M+1} \tag{6.18}$$

where η_o approximately equals to $(2\beta\text{-}1)$ for large β. Since $(A+1)\cdot R'$ is fixed, I_{T3} functions for different A and R' are scaled versions of the same curve versus V_i. Therefore, they all have the same approximation accuracy and efficiency at a fixed β.

Instrumentation amplifier architecture was used to implement the variable gain voltage buffer as shown in Figure 6.19. This circuit isolates the negative terminal with a unity gain voltage buffer to increase its input impedance so it hardly loads the TCA circuit and draws negligible input current; TCA only has litter bias current and is sensitive to the input impedance of the buffer. Without this isolation R2 has to be

chosen impractically large to serve the same purpose. Variable gain can be enabled by using variable resistors and its output voltage V_o is

$$V_o = A \cdot (V_+ - V_-). \tag{6.19}$$

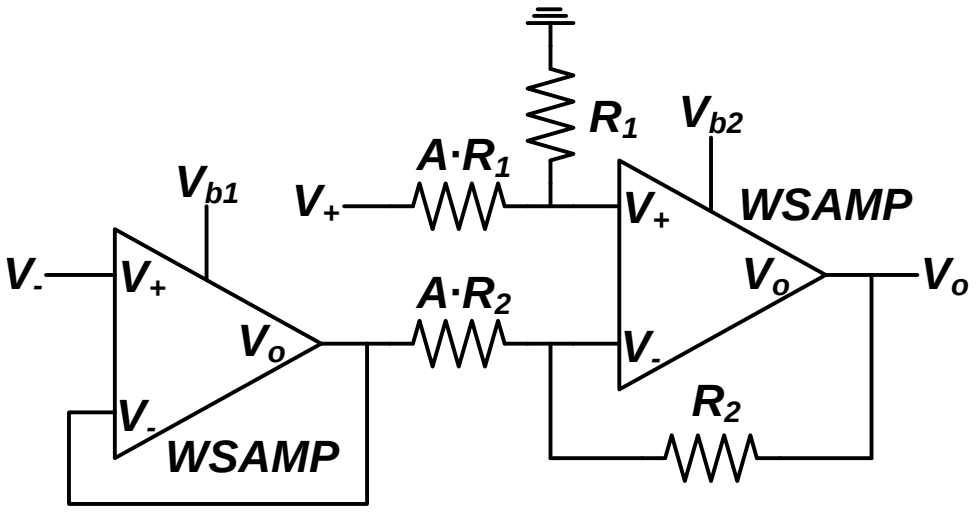


Figure 6.19. Architecture for the instrumentation amplifier. Wide-swing amplifiers were used for both amplifiers inside the instrumentation amplifier.

In order to achieve arbitrary input scaling the TCA needs to have a wide input common mode range. To accommodate this requirement wide-swing amplifier (WSAMP) [73] (as shown in Figure 6.20) is used for the two amplifiers in the instrumentation amplifier. The technique to increase input common mode range is to use complementary input stages in parallel. One drawback for this circuit is that the gain is not constant over different input common mode. Gain at transition among having single input stage on and both input stages on is not smooth. However, this effect can be minimized by making the conduction parameters of the input transistors the same by adjusting the MOSFET size according to mobility of both types of MOSFET. A compensation capacitor was added to stabilize the amplifier.

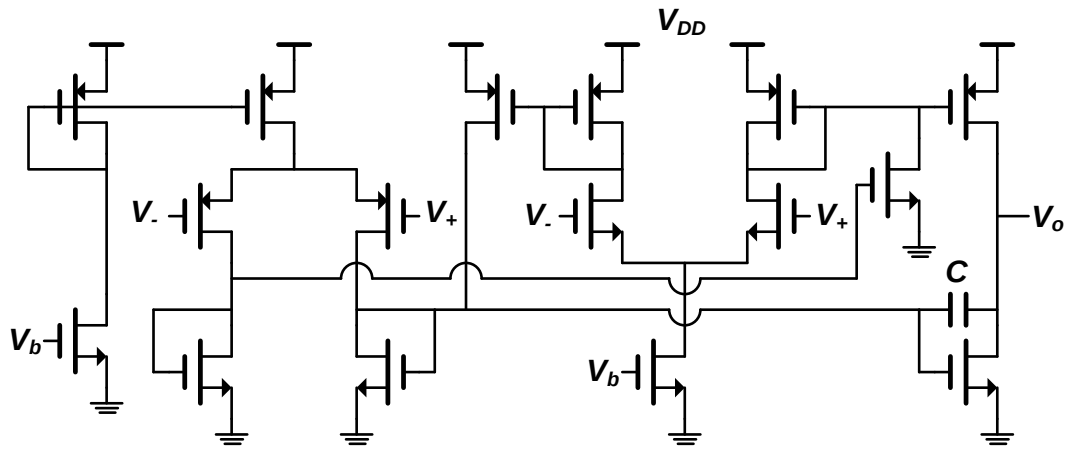


Figure 6.20. Architecture for the WSAMP.

6.3.5 Design Procedure

The design procedure is summarized in Table 6.4. First the input specification (V_o, V_r, and N) is determined. Next the size of the input PMOS and I_B are chosen to ensure operation in weak inversion and, possibly, to meet speed requirements of the circuit. Parameter κ can be determined from simulation or measurement results. Then, V_c, V_s, and M (see Figure 6.16 for the relationship between M and N) are found at step 3. Step 4 computes I_{T3} for a set of possible values of R, and, then, picks an R value which yields acceptable area and a wide range of β where r^2 is close to one (hereafter defined as Φ). There are two tradeoffs for the selection of R. The first tradeoff is area to Φ width and accuracy (compare the curves with R of 100 KΩ and 200 KΩ in the top figure of Figure 6.17). Normally the wider Φ is, the higher tolerance to design variations the circuit has. The second one is area to efficiency: according to Equation 6.19, valid β should be as high as possible to achieve high efficiency but it requires large R (see the top figure in Figure 6.17). At step 5 β is selected from Φ, preferably

close to the middle of Φ for better error tolerance. At the last step A is determined by Equation 6.17 and $R' = R/(A+1)$.

Table 6.4 Summary of Design Procedure

Step	Design
1	Determine system specification V_o, V_r, and N
2	Determine size of the input PMOS, I_B, and κ
3	$V_c = V_r/N$, $V_s = 2V_c$, M = ceiling(N/2)
4	Pick R which yields a wide range of β, Φ, for high accuracy
5	Choose β from the middle of Φ
6	Find A using (9) and $R' = R/(A+1)$

6.3.6 Implementation Results

The circuit was simulated and fabricated using commercially available 0.5μm 3M2P CMOS technology. The chip photomicrograph with the functional blocks is shown in Figure 6.21. We provide simulation results for two design examples and measurement result for the first example. Design parameters and results are summarized in Table 6.5. In the simulation V_{dd}, I_B and R were set to 5 V, 200 nA and 200 KΩ respectively. A and R' were set to approximated values in the simulation assuming we do not have precise control of these components. The overall results agree with our analysis very well. However, I_B cannot be measured directly in our design and so cannot the individual current function of the TCA (I_T). Therefore, β and η are difficult to determine so these values are not provided.

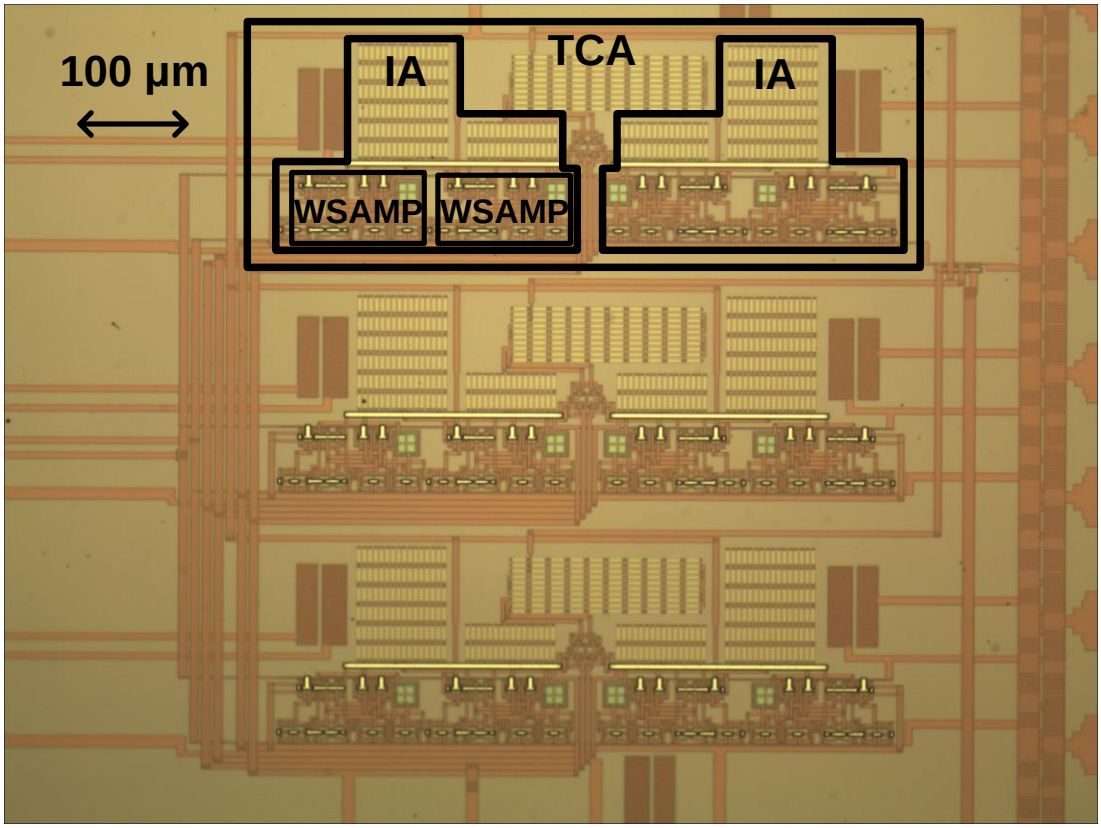


Figure 6.21. Chip photomicrograph. The chip has three identical TCAs and one is highlighted in the image. Each TCA has two instrumentation amplifier (IA) as highlighted. Each IA has two WSAMP as highlighted.

Table 6.5 Summary of Design Parameters

Parameters \ Design	**Design 1**			**Design 2**	
	Ideal	***Simulation***	***Measurement***	***Ideal***	***Simulation***
V_o / V_r / V_c / V_s (V)	2.5 / 1 / 1 / 2			3 / 0.8 / 0.2 / 0.4	
N / M	1 / 1			4 / 2	
β	0.9	0.8827	N.A.	0.9	0.8979
A	0.1626	0.17	0.17	0.8131	0.82
R' (KΩ)	172.0	150	156	110.3	100
r^2	1.0000	0.9993	0.998	1.0000	0.9994
η	0.2667	0.2493	N.A.	0.16	0.1626

In the implementation we have fixed resistors instead of variable resistors. We chose R_1 of 16 KΩ, R_2 of 48 KΩ, and A of 0.17. All the NMOSs in WSAMP have the

same size of 7 μm/3.5 μm; all PMOSs have the same size of 21 μm/3.5 μm. In the simulation the resistive dividers providing voltage biases were replaced with diode-connected PMOSs while in the real chip the three biases where provided outside form three pins. Current biases in the TCA were implemented using current mirrors with one external bias voltage. Both examples demonstrated nearly perfect sine functions and the DC simulation is shown in Figure 6.22. Measurement of design 1 also showed a nearly perfect r^2 value of 0.998. This circuit has not been integrated with the odometry circuit so no overall performance can be reported yet.

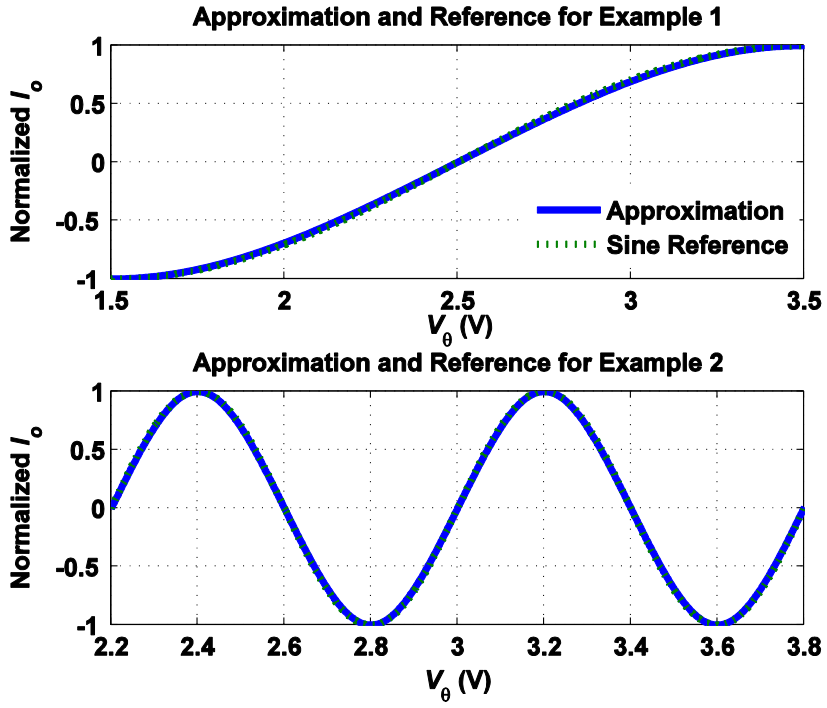


Figure 6.22. Simulation output showing nearly perfect sine function within defined input range.

We also simulated and tested design 1 functioning as a triangle-to-sine converter (TSC). Total harmonic distortion (THD) and spurious-free dynamic range (SFDR) were used to evaluate the quality of the output sine waveform. Because the

quality of the signal depends heavily on the match between devices, we ran Monte Carlo simulation of the TSC for 1000 runs to investigate the influences of mismatch between devices fabricated on the same run. The simulation conditions are summarized in Table 6.6. Because mismatch model from the foundry was not available, three parameter that we considered most representative were chosen to vary statistically in the simulation: threshold voltage of the MOSET, size of the MOSFET, and resistance. Threshold voltage and size were chosen because most mismatch of the MOSFETs can translate to these two parameters. The variations were assumed to be normal distribution with zero mean and standard deviation as specified in Table 6.6. Simulation results showed that the average, standard deviation, maximum, and minimum SFDR at 1 KHz are -51.3, 2.76, -56.88, and -43.93 dBc respectively; those metrics of THD at 1 KHz are 0.40, 0.08, 0.68, and 0.25 %, respectively. The distribution of SFDR and THD is in Figure 6.23. In the physical layout special attention was paid on the matching between devices. Techniques like common centroid and interdigitized structures were widely used (see Figure 6.21).

Table 6.6 Monte Carlo Simulation Conditions for the TSC

Parameters	Nominal	Distribution	Standard Deviation
NMOS threshold voltage	597 mV	Normal	10 mV
PMOS threshold voltage	915 mV	Normal	10 mV
MOSFET width	depends	Normal	2 nm
MOSFET length	depends	Normal	2 nm
Resistance	depends	Normal	3 %

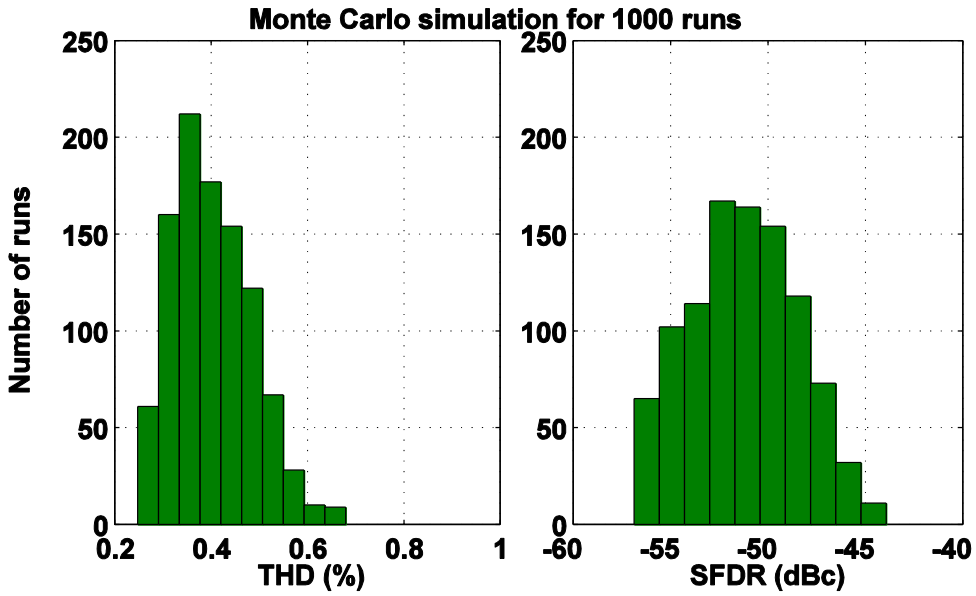


Figure 6.23. Monte Carlo simulation for design 1. A total of 1000 runs was performed. Figures show the distribution of THD and SFDR.

Table 6.7 summarizes the comparisons with previous works of sinusoidal waveform generator. Results show that our design produces nearly perfect sine waveform at a wide frequency range and requires the least power. We believe the deviation of measurement results from simulation was due to the mismatch between devices was severer than we have set in the simulation. The lowest operating V_{dd} is limited by the input requirement which sets the bias V_o+MV_s (see Figure 6.15). We experimented with several different combinations of V_o and V_r for N and M both equal to one. We found that the circuit can always generate nearly perfect sine waveform as long as V_{dd} is not lower than 1.8 V and greater than V_o+MV_s by 200 mV in the simulation. In the measurement we tested with 1.43 V supply and still obtained nearly perfect sine wave.

Table 6.7 Comparisons of Sinusoidal Waveform Generator

Ref	**[149]**[a]	**[150]**[b]	**[142]**[c]	**Simulation**	**Measurement**
Process (μm)	0.35 CMOS	0.18 CMOS	0.35 SiGe BiCMOS	0.5 CMOS	0.5 CMOS
Area (μm^2)	$3.8x10^4$	n.a.	$2.7x10^4$ [d]	N.A.	$6.2x10^5$
Power (μW)	n.a.	> 409	$2.3x10^4$	25.1	115.6
SFDR (dBc)	n.a.	-54.3@125	-45.7 @2.49G	<-52@0.1-10K -39.5@100K	<-44@1-5K <-37@5K-20K
THD (%@Hz)	< 5.11 @8k-21M	0.22@125 1.18@2.5M	n.a.	<0.37@0.1-10K 1.61@100K	<1.2@1-5K 2.15@20K

[a.] Not a TSC design; [b.] Simulation; [c.] Only TSC part; [d.] Estimation

Chapter 7: Dynamic Clock Scaling*

Many computations can be performed with mixed-signal circuits but they cannot completely replace digital circuits. Interfacing between different computational blocks, coordination of the system, and digital components in many cases require a clock signal. Clock generation and its distribution can potentially consume the largest portion of the total power consumption of a system especially when the clock speed is high. Therefore, a low power clock generation is highly desired. This design was inspired during the experiments on the CSVCO circuit (see Chapter 2, square signal can be a clock signal); we observed that the performance exhibited a tradeoff between jitter and power consumption. This tradeoff enables dynamic power management that is highly desired for a power constrained system.

7.1 Introduction

VCOs are fundamental building blocks for many very-large-scale integration (VLSI) systems and are widely used in many timing applications [50, 151, 152]. Current-starved ring oscillators (ROs) are popular as they offer a reasonable balance between area, power, and phase noise, in addition to having the widest tuning range [66, 67]. In comparison with other popular VCO architectures, such as the *LC*-tank oscillator, the compact design and easy integration of ROs make them especially attractive in monolithic VLSI systems [153]. Moreover, these benefits of ROs scale with technology, making them even more attractive in advanced technologies [154].

* Most of the material in this section was accepted for publication "T.-H. Lee and P. A. Abshire, "Frequency-boost jitter reduction for voltage-controlled ring oscillators," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*" in February, 2016. © 2016 IEEE.

However, ROs are generally considered to have poor phase noise and jitter performance that affect system performance adversely. This paper presents a jitter reduction technique to allow more VLSI systems to utilize ROs, which in turn should provide better performance for monolithic systems.

Most existing techniques to reduce jitter and phase noise in ROs involve putting ROs in a phase-locked loop (PLL) and relying on the PLL to correct and confine timing and phase error. These techniques focus on minimizing the ring oscillator induced noise in the loop but not on reducing intrinsic noise. A slew rate balancing circuit equalizes the rising and falling slew rate of the oscillator and, thus, improves the symmetry of the output waveform so that the up-converted flicker (1/f) noise is minimized [153]. A dual control VCO and a 4^{th} order PLL were used to stabilize VCO gain over process variations and frequency shifts, which in turn maintained the PLL bandwidth to improve jitter performance [155]. Similarly the dual control technique was used to minimize VCO gain and maximize loop bandwidth to reduce jitter while maintaining a wide frequency range [65, 156]. However, use of a PLL requires a reference signal which may not be available in all systems. There are other techniques that do not require a PLL but can potentially be incorporated with a PLL to further improve jitter performance. One such method is a phase-aligning technique that realigns the output to a clean reference periodically so that the jitter does not accumulate over a long period [157]. Another approach maximizing the output amplitude of the oscillator was found useful to improve phase noise and suppress noise current [158]. A similar study showed that fast rail-to-rail switching reduces phase noise [159]. Jitter can also be reduced by minimizing the

number of active devices so as to minimize noise sources [160]. Supply regulation is also important for ROs because supply noise modulates directly into ROs and induces timing error [156, 161, 162]. Design parameters can be optimized to reduce the intrinsic noise-induced jitter by understanding their governing relationship [72, 163, 164].

This work describes a novel frequency-boost technique to reduce jitter in ROs. Following the last set of techniques above, we have observed that both thermal noise and flicker noise-induced jitters are inversely proportional to oscillation frequency [72, 164]. The fundamental concept of this technique is that boosting the oscillation frequency achieves low jitter in conjunction with a frequency divider (FD) that adjusts the RO output to the desired output frequency. An oscillator followed by a FD is widely used but rarely are they designed as one unit and their overall performance reported. The FD has to be carefully designed to add minimal jitter to ensure the overall jitter performance is better than that of a comparable non-frequency-boosted design. Expected performance has been confirmed using theoretical analysis [72, 164] and verified with measurement results.

This chapter is organized as follows. Section 7.2 explains the theoretical basis of the frequency-boost technique and its mechanism of jitter reduction as well as the effect of frequency division on jitter. Section 7.3 describes the prototype circuit and discusses frequency sensitivity in its bias generation. Section 7.4 presents measurement results of frequency sensitivity, jitter, power, and a power- and frequency-normalized jitter figure-of-merit (FOM). In Section 7.5 we discuss design guidance based on the measurement results and the jitter model, how technology

influences the performance of current-starved ROs, and applications of this technique. The main contributions of this work are the novel frequency-boost jitter reduction technique, its theoretical and experimental verification, an improved jitter model for the current-starved RO, and presentation of the direct relationship between oscillation frequency and jitter for ROs. For the purposes of lowering jitter and raising FOM, this technique can be applied to any oscillator whose dominant jitter variance is inversely proportional to oscillation frequency raised to a power greater than one and two, respectively.

7.2 Frequency-Boost Technique

7.2.1 Jitter in ROs

We briefly summarize Abidi's model [72] for jitter and phase noise in ring oscillators and introduce the theory underlying the frequency-boost technique. The oscillation frequency f_{osc} in a RO can be estimated as [49, 72]

$$f_{osc} = \frac{1}{2N \cdot t_d} = \frac{I_{INV}}{N \cdot V_{SW} \cdot C_{eff}} \tag{7.1}$$

where N is the number of inverter stages in the RO, t_d is the mean delay for one stage, I_{INV} is the controlled current in inverters, V_{SW} is the peak to peak swing amplitude in the RO, and C_{eff} is the effective loading capacitance for each inverter. This equation is valid under the assumption that the swing is centered within the rails and that currents remain at constant I_{INV} during both charging and discharging. Assuming V_{SW} and C_{eff} depend only weakly on I_{INV}, I_{INV} is approximately linearly proportional to f_{osc}.

For thermal noise-induced jitter, the timing delay uncertainty in one inverter stage $J_{d,w}$ is

$$J_{d,w} = \frac{1}{I_{INV}} \int_0^{t_d} i_w(t)dt \tag{7.2}$$

where $i_w(t)$ is thermal noise current. This integral can be converted into a convolution of $i_w(t)$ and a rectangular unit window of width t_d. The power spectral density (PSD) of $J_{d,w}$ can be found accordingly, then the variance of $J_{d,w}$ can be found using the Wiener-Khinchine theorem as

$$J_{d,w}^2 = \int_0^{\infty} S_{J_{d,w}}(f)df = \frac{t_d}{2} \frac{S_{i_w}(f)}{I_{INV}^2} \tag{7.3}$$

where $S_{J_{d,w}}(f)$ and $S_{i_w}(f)$ are the PSDs of $J_{d,w}$ and $i_w(t)$, respectively. The period jitter of the RO J_w is the summation of $2N$ independent $J_{d,w}$. The thermal noise current PSD of a transistor is proportional to its transconductance g_m, which is proportional to the square root of the transistor current. From this short derivation above, the effect of frequency-boost can be summarized. First, increasing the oscillation frequency linearly decreases the width of the integration window t_d as in Equation 7.1. Therefore, less noise is accumulated over a shorter integration period as can be seen from Equation 7.2. This is also reflected in Equation 7.3. Second, increasing the oscillation frequency (equivalent to increasing current in the RO) increases the signal to noise ratio of the current as can be seen from the term after the second equal sign in Equation 7.3 because the PSD $S_{i_w}(f)$ only has a square root current dependence. The t_d term introduces an inverse dependence, the $S_{i_w}(f)$ term a square root dependence, and the I_{INV} term an inverse square dependence on the oscillation frequency. As a result, the variance of thermal noise-induced jitter is proportional to the oscillation frequency to the -2.5th power. Phase noise can be inferred from jitter noise and f_{osc}: $L_w(f) = J_w^2 \cdot f_{osc}^3 / f^2$ [72]. Note that we have neglected jitter contribution from KTC

thermal noise which will later be shown to be negligible compared to the thermal noise current.

Flicker noise-induced jitter is derived differently because an integral similar to Equation 7.3 cannot be integrated analytically. An alternative method is to find the frequency sensitivity to noise sources, and then, phase noise PSD $L_f(f)$ from the frequency sensitivity according to

$$L_f(f) = \frac{k_{i_f}^2}{4f^2} S_{i_f}(f) \tag{7.4}$$

where $k_{i_f}(f)$ is the frequency sensitivity and is equal to $f_{osc}/(2NI_{INV})$ and $S_{i_f}(f)$ is the PSD of flicker noise current $i_f(t)$ in one inverter in the RO. The net phase noise PSD is simply the summation of all individual phase noise PSDs because they are uncorrelated. $S_{i_f}(f)$ is found using an empirical flicker noise model and is linearly proportional to current (equivalent to oscillation frequency). The variance of flicker noise-induced period jitter J_f^2 is found from the relationship of jitter and phase noise for flicker noise suggested by Liu and McNeill [164, 165]

$$J_f^2 = 25 L_f(f) \frac{f^3}{f_{osc}^4}. \tag{7.5}$$

The $L_f(f)$ term introduces a linear dependence and the f_{osc} term an inverse quartic dependence on the oscillation frequency. As a result, variance of flicker noise-induced jitter is proportional to oscillation frequency to the -3rd power.

In summary, the variances of period jitter due to thermal noise J_w^2 and flicker noise J_f^2 in a voltage-controlled RO are derived from a jitter model for a simple inverter RO as [72, 164]

$$J_w^2 = \frac{\eta_{w1}}{f_{osc}^{2.5}} + \frac{\eta_{w2}}{f_{osc}^2} \tag{7.6}$$

$$J_f^2 = \frac{\eta_f}{f_{osc}^3} \tag{7.7}$$

where η_{w1} and η_{w2} are combinations of frequency independent terms reflecting V_{DD}, design geometry, Boltzmann constant, temperature, thermal noise coefficients, and technology parameters. η_f is also a frequency independent term reflecting design geometry, flicker noise coefficients, and technology parameters. These two equations are simplified versions in order to focus on the frequency dependence. Detailed derivations and equations will be given later. From Equations 7.6 and 7.7, increasing f_{osc} decreases jitter when other design parameters are held fixed.

7.2.2 Jitter in FDs

When the oscillation frequency is boosted, an FD is required to adjust the RO output to the desired output frequency. However, this frequency division has a negative effect on overall jitter performance. In order to estimate the period jitter, we consider the period uncertainty for two adjacent rising edges in the frequency signal. The effect of frequency division is that instead of two adjacent rising edges, we consider two rising edges that are farther away. If the frequency is divided by A, we have to account for period jitters over A cycles. Treating period jitter as a random variable, this is equivalent to summation of A independent and identically distributed random variables. Therefore, if the jitter variance at the oscillator output is J_{osc}^2, the jitter variance at the FD output is $J_{out}^2 = A \cdot J_{osc}^2$, which was observed in N-cycle jitter measurements [165, 166].

7.2.3 Jitter in RO plus FD

The effect of total jitter performance of the RO and the FD after division by a factor of A is summarized as:

$$J_{out}^2 = \frac{\eta_{w1}}{\mathrm{F}_{\mathrm{out}}^{2.5}} \frac{1}{A^{1.5}} + \frac{\eta_{w2}}{\mathrm{F}_{\mathrm{out}}^{2}} \frac{1}{A} + \frac{\eta_f}{\mathrm{F}_{\mathrm{out}}^{3}} \frac{1}{A^2} \tag{7.8}$$

where $\mathrm{F}_{\mathrm{out}}$ is the desired output frequency and equals f_{osc}/A. It is clear that the higher the frequency division A (the higher the oscillation frequency f_{osc}), the better the jitter performance.

7.2.4 Detailed Derivations for Jitter

The thermal and flicker noise-induced jitter variances in a RO consisting of simple inverters are, respectively [72, 164]

$$J_w^2 = \frac{2k_B T}{I_{INV} \cdot f_{osc}} \left(\frac{1}{V_{DD} - V_t} (\gamma_N + \gamma_P) + \frac{1}{V_{DD}} \right) \tag{7.9}$$

$$J_f^2 = \frac{25}{8NI_{INV} \cdot f_{osc}^2} \left(\frac{\mu_N K_{fN}}{L_N^2} + \frac{\mu_P K_{fP}}{L_P^2} \right) \tag{7.10}$$

where k_B is Boltzmann's constant, T is absolute temperature, I_{INV} is the current of the inverters, V_t is threshold voltage of MOSFET (assuming V_t is the same for NMOS and PMOS). Parameters γ_N, μ_N, K_{fN}, and L_N are thermal noise coefficient, carrier mobility, flicker noise coefficient, and channel length for NMOS; γ_P, μ_P, K_{fP}, and L_P are for PMOS. γ_N and γ_P depend on the operation region and are generally taken to be 2/3 for saturation region and 1 for linear region. K_{fN} and K_{fP} are empirically chosen to be 10^{-24} [72]. For a RO consisting of current-starved inverters the jitter model has to be modified. The transconductance g_m is different for the two types of ROs. For simple

inverters $g_m = 2I_{INV}/(V_{DD}-V_t)$ whereas for current-starved inverters $g_m = (2\beta_N I_{CSI})^{0.5}$ where β_N is the transconductance parameter of NMOS transistor *M2* (see Fig. 2) and I_{CSI} is the average current-starved inverter current (the same for charging and discharging). The other difference is that we included the noise contributions from *M1* and *M4* (but ignored *M5* for simplicity), and still assume that noise from all transistors are uncorrelated. These modifications do not change the shape of noise PSDs or the frequency sensitivity used in the derivations for flicker noise. Moreover, the assumption of wide-sense stationary process for thermal noise still holds. Therefore, the derivations used in Abidi's model are still valid. The resulting PSDs of thermal noise current and flicker noise current in the NMOS transistors are respectively

$$S_{i_{w,N}}(f) = 4k_B T\gamma_N \sqrt{2I_{CSI}}(\sqrt{\beta_N} + \sqrt{\beta_{N,CS}}) \tag{7.11}$$

$$S_{i_{f,N}}(f) = \frac{2I_{CSI}\mu_N k_{fN}}{fL_N^2}\left(1+\frac{1}{\beta_L^2}\right) \tag{7.12}$$

where subscript *CS* stands for current-starved transistors and $\beta_L = L_{N,CS}/L_N$. PSDs for the PMOSs (charging phase) can be found using the same method. The resulting jitter variances are

$$J_w^2 = \frac{2k_B T}{I_{CSI}\cdot f_{osc}}\left(\frac{\sqrt{\beta_N}+\sqrt{\beta_{N,CS}}}{\sqrt{2I_{CSI}}}(\gamma_N+\gamma_P)+\frac{1}{V_{DD}}\right) \tag{7.13}$$

$$J_f^2 = \frac{25}{8NI_{INV}\cdot f_{osc}^2}\left(1+\frac{1}{\beta_L^2}\right)\left(\frac{\mu_N K_{fN}}{L_N^2}+\frac{\mu_P K_{fP}}{L_P^2}\right). \tag{7.14}$$

In the derivation we have assumed equal transconductance parameters for both types of corresponding MOSFETs ($\beta_N = \beta_P$, and $\beta_{N,CS} = \beta_{P,CS}$) which is reasonable because

most designs would impose this condition by adjusting the device sizes in order to make the threshold voltage of the RO equal to $V_{DD}/2$ and minimize phase noise.

Next, we replace I_{CSI} with other design parameters so that the frequency dependence in the jitter model can be isolated. According to Equation 7.1 the effective capacitance is

$$C_{eff} = \frac{I_{CSI}}{V_{DD} \cdot N \cdot f_{osc}} \tag{7.15}$$

assuming $V_{SW} = V_{DD}$. On the other hand, assuming the effective capacitance is mainly contributed by the gate capacitance that

$$C_{eff} = \beta_C C_{ox}'(W_P L_P + W_N L_N) \tag{7.16}$$

where β_C depends on the operating region of the MOSFETs of interest, C_{ox}э is oxide capacitance per area, and W_P is the channel width for PMOS. Here we have considered the GCSINV as a simple inverter with only *M2* and *M3* (see Figure 7.2). The previous stage contributes no capacitance assuming the conducting MOSFET is in saturation and the complementary MOSFET is off. Then, the effective capacitance only depends on the next stage; β_C is 1 for MOSFETs in linear and cut-off region and is 2/3 for saturation region [167]. In this work we have treated β_C as a constant assuming the operation does not vary dramatically across ROs. Furthermore we assume $L_P = L_N$ and due to the matched transconductance parameters $(W_N/L_N)/(W_P/L_P) = k_p'/k_n'$ where k_p' and k_n' are process transconductance for PMOS and NMOS, respectively. Equation 7.16 then becomes

$$C_{eff} = \beta_C' C_{ox}' L_N^2 \eta \tag{7.17}$$

where $\eta = k_n'/k_p'+1$and β_Cэ is $\beta_C \cdot W_N/L_N$, which is also a constant assuming that W_N/L_N is the same for all ROs. From Equations 7.15 and 7.17 we have obtained

$$I_{CSI} = \beta_C' C_{ox}' L_N^2 \eta V_{DD} N f_{osc} \tag{7.18}$$

Then, substituting Equation 7.18 into Equation 7.13 we have

$$J_w^2 = J_{w1}^2 + J_{w2}^2 \tag{7.19}$$

$$J_{w1}^2 = \frac{\sqrt{2} k_B T \left(\sqrt{\beta_N} + \sqrt{\beta_{N,CS}}\right)(\gamma_N + \gamma_P)}{f_{osc}^{2.5} \left(\beta_C' C_{ox}' L_N^2 \eta V_{DD} N\right)^{1.5}} \tag{7.20}$$

$$J_{w2}^2 = 2 k_B T f_{osc}^{-2} \left(\beta_C' C_{ox}' L_N^2 \eta V_{DD}^2 N\right)^{-1} \tag{7.21}$$

where J_{w1}^2 is jitter contributed by thermal noise current from the transistors and J_{w2}^2 is jitter contributed by KTC thermal noise. Substituting Equation 7.18 into Equation 7.14, we have

$$J_f^2 = \frac{25\left(1+\frac{1}{\beta_L^2}\right)\left(\mu_N K_{fN} + \mu_P K_{fP}\right)}{8 f_{osc}^3 N^2 L_N^4 \left(\beta_C' C_{ox}' \eta V_{DD}\right)} \tag{7.22}$$

If we take into account the effects of FD as discussed in Section 7.2.2 (with $A = 2^M$) and replace variables with fundamental process parameters and design parameters, we have

$$J_{w1}^2 = \beta_1 V_{DD}^{-1.5} \mu_N^{0.5} \left(\frac{\mu_N}{\mu_P} + 1\right)^{-1.5} t_{ox} L_N^{-3} (\gamma_N + \gamma_P) \tag{7.23}$$

$$J_{w2}^2 = \beta_2 V_{DD}^{-2} \left(\frac{\mu_N}{\mu_P} + 1\right)^{-1} t_{ox} L_N^{-2} \tag{7.24}$$

$$J_f^2 = \beta_3 \left(\frac{\mu_N}{\mu_P} + 1 \right)^{-1} t_{ox} L_N^{-4} \left(\mu_N K_{fN} + \mu_P K_{fP} \right) \tag{7.25}$$

where β_1, β_2, and β_3 are parameters that do not vary with technology and can be expressed as

$$\beta_1 = \sqrt{2} k_B T \varepsilon_{ox}^{-1} \beta_C^{-1.5} \left(\frac{W_N}{L_N} \right)^{-1.5} \left(\sqrt{\frac{W_N}{L_N}} + \sqrt{\frac{W_{N,CS}}{L_{N,CS}}} \right) N^{-1.5} 2^{-1.5M} \mathrm{F}_{\mathrm{out}}^{-2.5} \tag{7.26}$$

$$\beta_2 = 2 k_B T \varepsilon_{ox}^{-1} \beta_C^{-1} \left(\frac{W_N}{L_N} \right)^{-1} N^{-1} 2^{-M} \mathrm{F}_{\mathrm{out}}^{-2} \tag{7.27}$$

$$\beta_3 = \frac{25}{8} \varepsilon_{ox}^{-1} \beta_C^{-1} \left(\frac{W_N}{L_N} \right)^{-1} \left(1 + \frac{1}{\beta_L^2} \right) N^{-2} 2^{-2M} \mathrm{F}_{\mathrm{out}}^{-3} \tag{7.28}$$

7.3 Circuit Description

The circuit consists of a current-starved VCO followed by a FD as shown in Figure 7.1. The oscillator structure is a single-ended current-starved RO with nine stages of gated current-starved inverters (GCSINVs) which can be reconfigured to 3, 5, and 7 stages. The frequency divider has a cascade of 7 divide-by-2 (DB2) circuits. The output can be selected from the oscillator or any of the DB2 circuits. The resulting output frequency is the frequency of the oscillator divided by two to the power of the number of DB2 stages, 0 to 7.

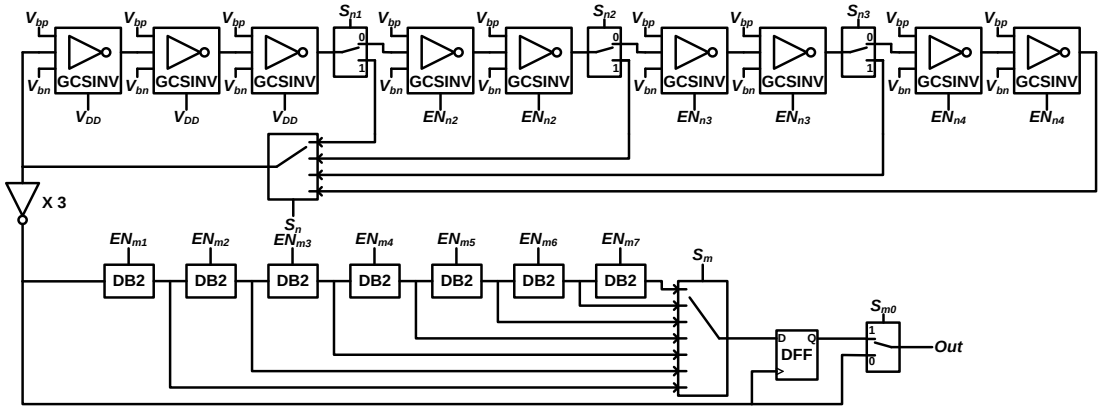


Figure 7.1. System diagram of the current-starved ring oscillator (top) and frequency divider (bottom). Switches are introduced as shown to enable reconfigurability. Three sequential inverters on the left sharpen the signal edge of the oscillator to ensure that the divide-by-2 (DB2) circuit works properly. A re-sampling D-flip-flop (DFF) is used at the output to minimize the jitter introduced by the FD.

7.3.1 Current-Starved Ring Oscillator

A. Gated Current-Starved Inverters

The ring oscillator consists of nine stages of current-starved inverters. Each inverter has an NMOS switch to ground at the bottom to turn it on or off as shown in Figure 7.2 (a). Unused components are disabled to ensure they do not consume power and contribute additional noise to the rest of the circuit.

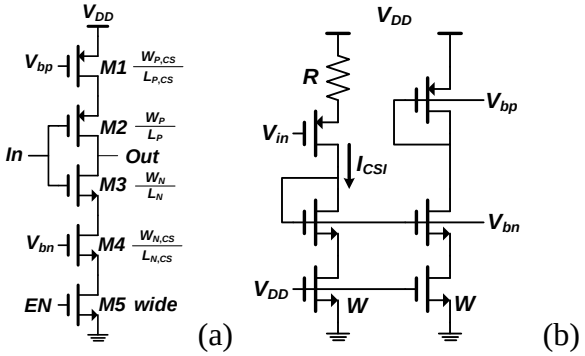


Figure 7.2. (a) GCSINV has an NMOS to ground to enable this stage. *W*/*L* indicates the width over length for the transistors. (b) Schematic of the circuit that generates the bias voltage for the GCSINVs. The two dummy transistors at the bottom with their gate connected to V_{DD} are used to match the gated inverters. The body of PMOSs are tied to V_{DD} and NMOSs to ground.

B. Bias Generation

In this design we use a PMOS input transistor for converting control input voltage to current (see Figure 7.2 (b)). A source-degenerated resistor is used to increase the linearity of the input voltage to control current. PMOS has the potential to achieve lower frequency sensitivity to power supply voltage compared to NMOS. The converted current is

$$I_{CSI} = \left[V_x + \frac{1}{\beta_p R} + \sqrt{\frac{2V_x}{\beta_p R} + \frac{1}{(\beta_p R)^2}} \right] \Big/ R \tag{7.29}$$

where V_x is defined as $V_{DD}–V_{in}–V_{tp}$, V_{tp} is the threshold voltage of the input PMOS, and β_P is the transconductance parameter of the input PMOS. Inserting Equation 7.29 into Equation 7.1 we obtain the oscillation frequency. The frequency sensitivity to supply voltage is defined as $S_p = (\partial f_{osc}/\partial V_{DD})/f_{osc}$ and can be expressed as

$$S_P = \frac{V_{SW}\left(\frac{2}{\beta_p R\sqrt{\varepsilon}} + 1 \right) - \delta\left(V_X + \frac{1}{\beta_p R} + \sqrt{\varepsilon} \right)}{V_{SW}\left(V_X + \frac{1}{\beta_p R} + \sqrt{\varepsilon} \right)} \tag{7.30}$$

where ε is defined as $2V_x/\beta_p R+1/(\beta_p R)^2$ and δ is defined as dV_{SW}/dV_{DD}. In Equation 7.29 we have neglected the dependence of V_{tp} to V_{DD} due to body effect for simplicity.

For most cases the oscillator switches rail to rail and, thus, V_{SW} equals V_{DD}. Following the method above the frequency sensitivity to supply voltage using an NMOS input transistor S_N is $–1/V_{DD}$. Comparing the absolute values of these two sensitivities, S_P is lower than S_N when V_{in} is less than $V_{DD}/2–V_{tp}$ and absolute S_p is the lowest when V_{in} is zero volt. Thus, the frequency sensitivity can be improved by using

a PMOS as an input transistor. The frequency-boost technique benefits from this improved frequency sensitivity because it tends to make the oscillation frequency higher which is equivalent to a higher current and a lower V_{in} using a PMOS input. Therefore, the oscillators will mostly operate in a low sensitivity region using the frequency-boost technique. If jitter is found from frequency sensitivity using the jitter derivation for flicker noise, this lower frequency sensitivity to supply translates to lower jitter due to supply noise. All the derivations above assumed that V_{in} is a ground-referenced control. If a supply-referenced control (V_{in} shifts with V_{DD}) is used, the frequency sensitivities found for the two different input transistors are opposite to what we derived.

7.3.2 Frequency Dividers

The DB2 circuit is a transmission gate based DFF with its inverted output connected to its D input and using CLK and Q as input and output respectively. The DFF has an NMOS switch to ground to enable it. The FD yields the output frequency $F_{out} = 2^{-M} \cdot f_{osc}$, where M is the number of DB2 stages. At the end of the FD is a re-synchronizing DFF. It re-samples the divided signal based on the oscillator signal to minimize the jitter contribution of the divider [168].

7.3.3 Control Generator and Switching Circuit

The control generation circuit uses digital logic gates. Control signals are generated based on two controls, N and M. N has two bits and controls the number of

stages used in the RO; *M* has three bits and controls the number of DB2 stages used in the FD. As a result the circuit has a total of 32 modes with the number of inverter stages and the number of DB2 stages varying. The mapping between *N*, *M*, and internal signals is listed in Tables Table 7.1 and 7.2, respectively.

Table 7.1 Control Generation for Voltage-Controlled Ring Oscillator

N	$S_{n1\text{-}4}$	$EN_{n2\text{-}4}$
3	1000	000
5	0100	100
7	0010	110
9	0001	111

Table 7.2 Control Generation for Frequency Divider

M	S_{m0}	S_m	$EN_{m1\text{-}7}$
0	1	0000000	0000000
1	0	1000000	1000000
2	0	0100000	1100000
3	0	0010000	1110000
4	0	0001000	1111000
5	0	0000100	1111100
6	0	0000010	1111110
7	0	0000001	1111111

Switches are used to reconfigure the circuit based on the control signal. The 7-1 switch has seven 1-1 switches, each controlled by one bit of S_m. In combination with the power gating technique, this approach ensures that unnecessary switching does not occur and unused components doesn't consume power.

7.4 Measurement Results

The circuit was fabricated using a commercially available 0.5 μm 2P3M CMOS technology. The nominal operating voltage is 5V. A chip photo is shown in

Figure 7.3. The active area is 0.15 mm^2. The jitter was measured as the difference between ideal transition time and the observed 50% transition time using a Tektronix MSO4034B oscilloscope. The phase noise was measured using an HP 4396B spectrum analyzer.

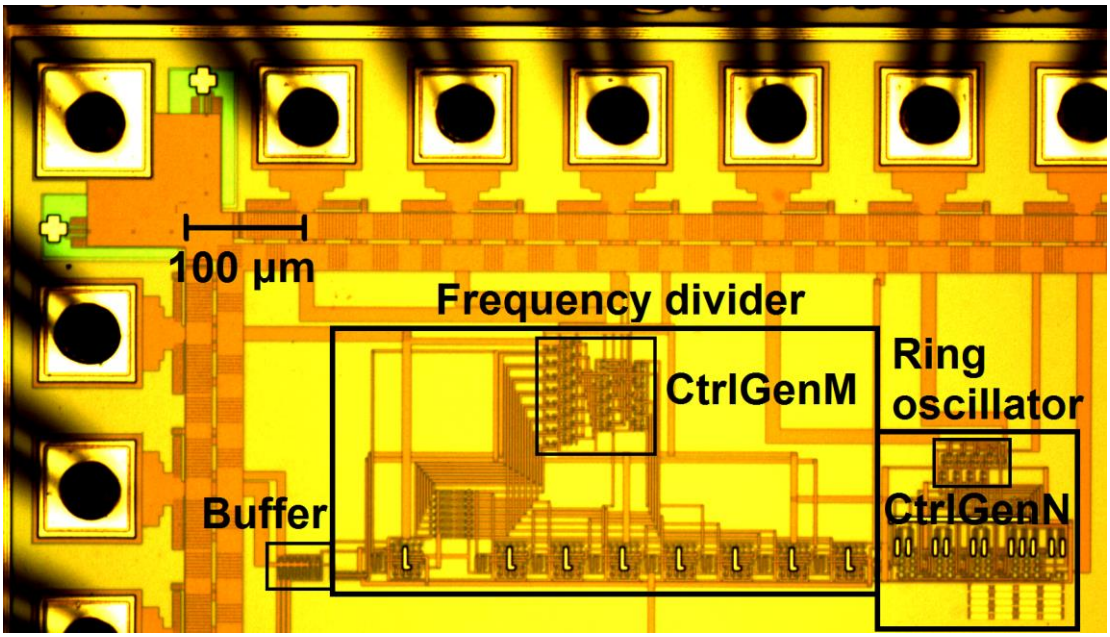


Figure 7.3. Chip photomicrograph

7.4.1 Frequency vs. V_{in}

We measured the output frequency as a function of V_{in} for several modes at 5 V supply as shown in Figure 7.4. The upper four traces are output frequencies measured from the RO output without frequency division. The lower four traces are output frequencies measured after seven DB2 stages which results in a division by 128. Generally the output frequency is linearly proportional to V_{in} from 0 V up to 3.2 V. After that it is no longer linear because the nonlinear terms in Equation 7.29 start to dominate.

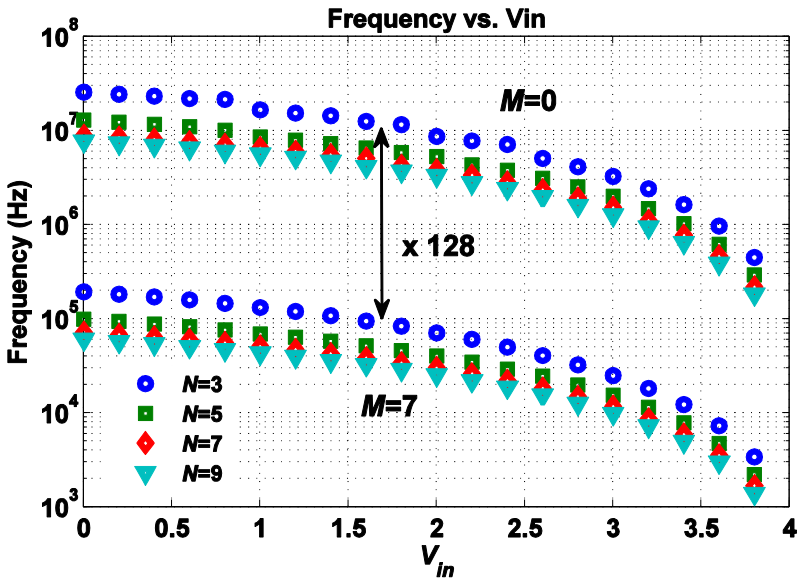


Figure 7.4. Measurement results of the output frequency vs. V_{in}. with all four N control and two M control.

7.4.2 Frequency Sensitivity to Supply Voltage

We measured output frequencies directly from the RO under supply voltages of 4.5, 5, and 5.5 V. Figure 7.5 shows the observed frequency sensitivity to supply voltage as compared to the expected values from Equation 7.30. TheoryP1 was calculated by assuming V_{SW} equals V_{DD}. TheoryP2 assumes V_{SW} is less than V_{DD} by 0.5 V and δ is 1.1. TheoryN is calculated based on the assumption of an NMOS input transistor. The overall measurement agrees with theory very well. For N = 3 and low V_{in}, the data matches with TheoryP2 better because under such conditions the oscillator has very high frequency and does not swing rail to rail. As expected from Equation 7.30, the lowest S_p is 3.5% which was measured when N = 3 and V_{in} = 0 V. Moreover, $S_p < S_n$ when $V_{in} < V_{DD}/2 - V_{tp} \approx 1.5$ V. However, a regulated supply [156, 161, 162] is still required for systems with strict requirements for power supply noise.

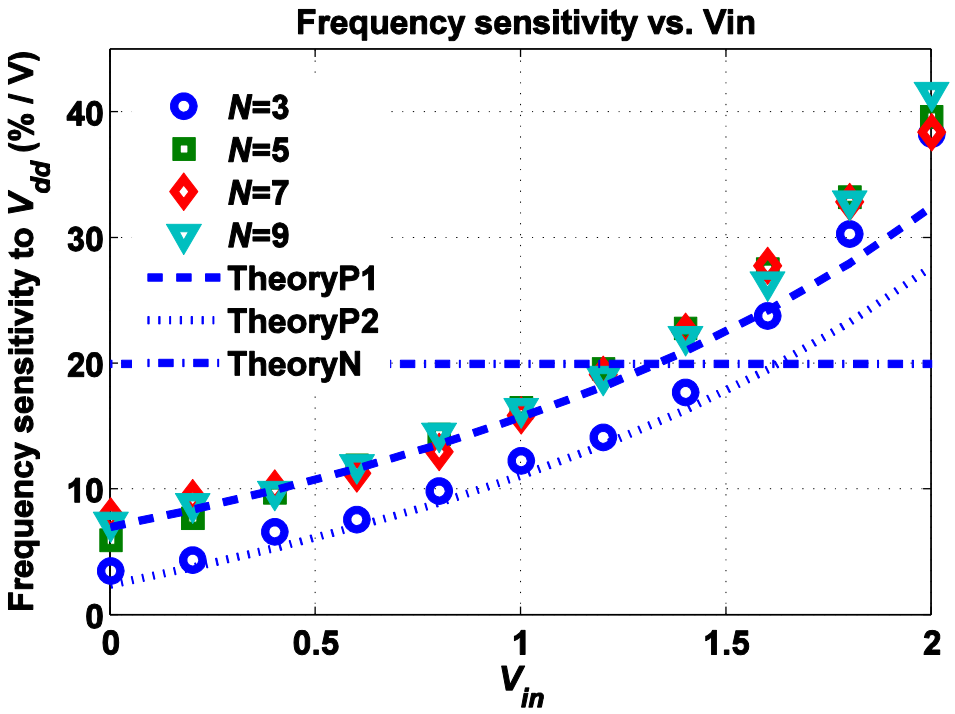


Figure 7.5. Observed frequency sensitivity compared to theoretical analysis.

The decreased oscillation amplitude would have a negative effect on the phase noise due to decreased carrier power. This happens when the ring size is small and current (equivalent to frequency) is high. This equivalent effect can be observed from our jitter model, Equations 7.23 and 7.24, which has inverse relationship to V_{DD}. However, using this design as a specific example the amplitude starts to decrease from V_{DD} (5 V) when V_{in} is less than 1.8 V and N = 3 (see curve with blue circles in Figure 7.5). When the current reaches the highest (V_{in} = 0 V), the amplitude becomes 4.5 V (decreases 10%) as predicted by TheoryP2. However, when V_{in} changes from 1.8 V to 0 V, the oscillation frequency increases by more than two times and it allows a frequency division by at least two. The overall effect is that the shrinkage of amplitude is negligible.

7.4.3 Jitter

A. Measurement Results

We measured jitter in units of ps at different output frequencies. A total of 99 data points across 28 modes was collected. No data were collected for modes (5,6), (5,7), (7,6), and (7,7) (modes are represented with the format of (*N*,*M*)). Data for selected modes is shown in Figure 7.6 (data points represented by markers). Due to the limitation of testing equipment the lowest jitter points (indicated by solid markers) experienced flooring as discussed in Appendix D. Therefore, the circuit has the potential to achieve even lower jitter performance. We didn't measure jitter at the highest frequency for some modes so we have extended the curve based on the slope of the data in order to estimate the performance. All modes have a slope between -1 and -1.5 which agrees with Equations 7.6 and 7.7. The average slope is -1.26 which indicates that the thermal noise of transistors in the inverter (Equation 7.20) dominates the jitter.

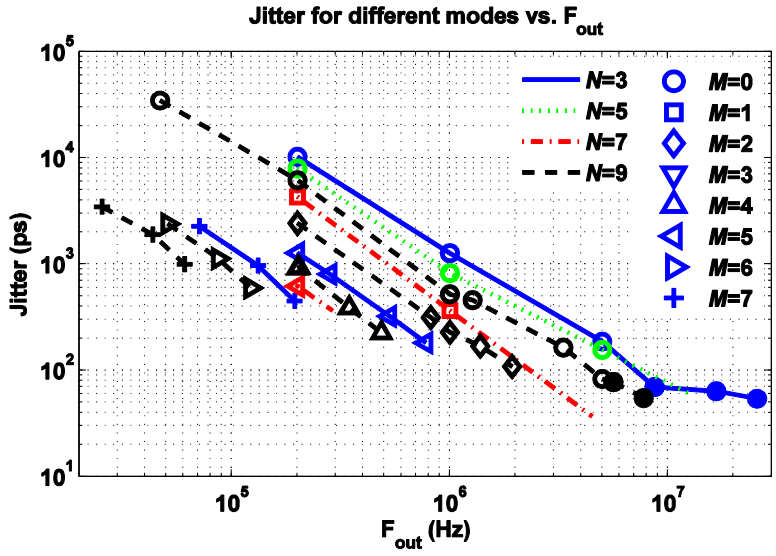


Figure 7.6. Jitter for selected modes as a function of output frequency. The mode is represented by a combination of line style and markers.

B. Model Verification

To verify the modified jitter model in Equations 7.20, 7.21, and 7.22, we have plotted jitter for four modes alongside their theoretical values in Figure 7.7. The four modes are respectively (3,0), (5,0), (7,0), and (9,0). These modes were selected because the oscillator output goes directly to the circuit output without frequency division so we can eliminate any uncertainty introduced by the FD.

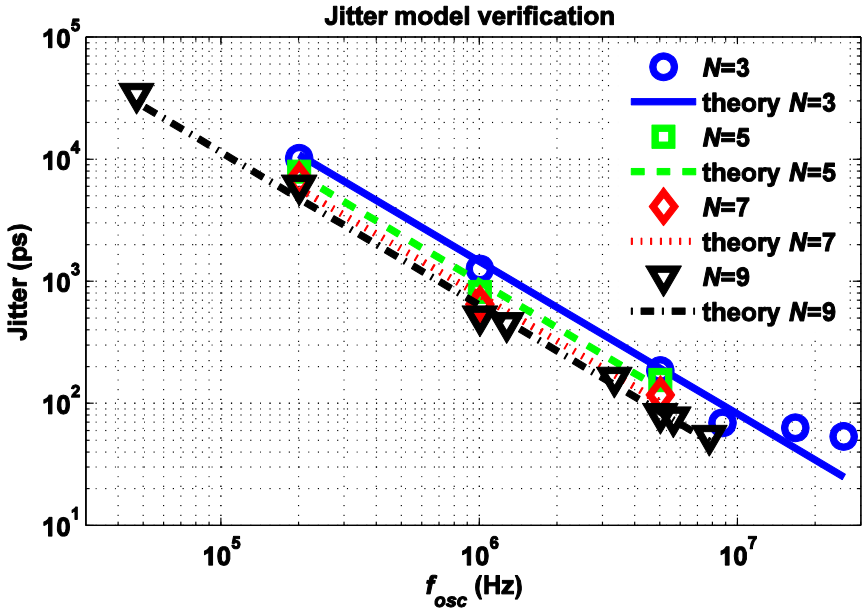


Figure 7.7. Jitter for four modes with theoretical model with fitted thermal noise coefficients plotted alongside measured data.

We observe that the jitter is dominated by the thermal noise of Equation 7.20 as suggested by the -1.26 slope in Figure 7.6. The flicker noise corner (defined as the output frequency where flicker and thermal noise-induced jitters are equal) is 7 KHz with SPICE process parameters [169] and L_N = 1 μm, W_N/L_N = 1, β_C = 1, N = 9, β_l = 3, $W_{N,CS}/L_{N,CS}$ = 5, so at lower frequencies we expect the 1/f noise of Equation 7.22 to be dominant.

We have also observed that this model underestimates the jitter if we take the thermal noise coefficients, γ_N and γ_P, to be 2/3 (assuming the MOSFETs are in saturation). This underestimation of thermal noise has been reported in the literature and results from short channel effects in the MOSFET [170, 171] as well as excess noise effects [170, 172]. We additionally attribute this underestimation to a couple of factors. First, we assumed that the conducting MOSFET in the previous stage remains in saturation but it would be in the linear region for some part of the transient. The thermal noise coefficient for the linear region is 1 instead of 2/3. Second, we have considered the GCSINVs as simple inverters in the derivation while we have also accounted for noise contributions from *M1* and *M4*. However, *M1*, *M4*, and *M5* contribute resistance to the sources of *M2* and *M3*, and previous work has suggested that the source/drain resistance has a significant effect on thermal noise and can enhance the thermal noise by a factor of 3 [170]. In addition, we overestimate the gate capacitance by assuming the MOSFETs in the next stage remain in the linear or cut-off region and neglect *M1*, *M4*, and *M5* in the GCSINVs. If *M1* and *M4* contribute capacitance to the sources of *M2* and *M3* respectively, this reduces the effective capacitance. This leads to an overestimation of β_C and, in turn, an underestimation of thermal noise.

7.4.4 Power

Measured power consumption for selected modes as a function of output frequency is shown in Figure 7.8. For most cases the power increases as the output frequency increases. This is due to increasing internal frequency causing more power consumption for most components. However, when the internal frequency is so low

that the switching is slow, the inverters conduct short-circuit current during switching and consume a large amount of power.

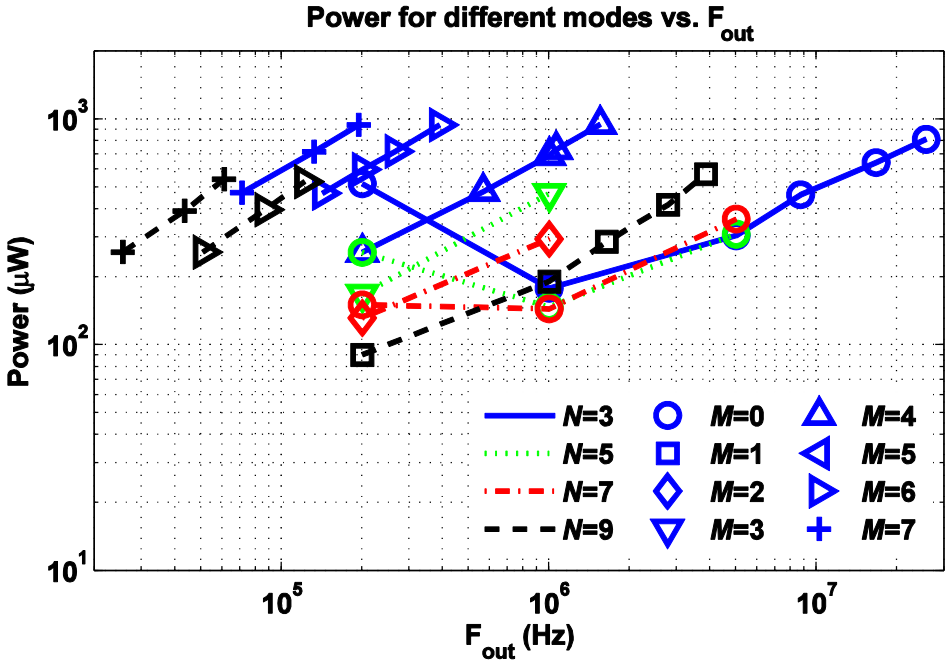


Figure 7.8. Power for selected modes as a function of output frequency. The mode is represented by a combination of line style and markers.

7.4.5 Jitter, Power, and FOM

A. Measurement Results for Fixed Output Frequencies

We measured power and jitter for all modes at three fixed output frequencies: 200 KHz, 1 MHz, and 5 MHz (see Figure 7.9). At each mode we have adjusted the input voltage so the output frequency matched the desired value. There are no data points for a few modes because under those modes the circuit cannot reach the desired output frequency. Data points on the same curve corresponds to modes with the same N (the same number of inverter stages) and, from left to right, adjacent points represent different values of M starting from zero. Therefore, the internal

oscillator frequency increases by a factor of two from the left to the next point starting from the set frequency. For most cases the power increases as the number of DB2 stages increases. However, when the internal frequency is low, the inverters conduct short-circuit current. This phenomenon can be seen for the lowest number of DB2 stages at 200 KHz; the power doesn't increase monotonically. These results clearly demonstrate that the jitter can be reduced by using the frequency-boost technique at the cost of power.

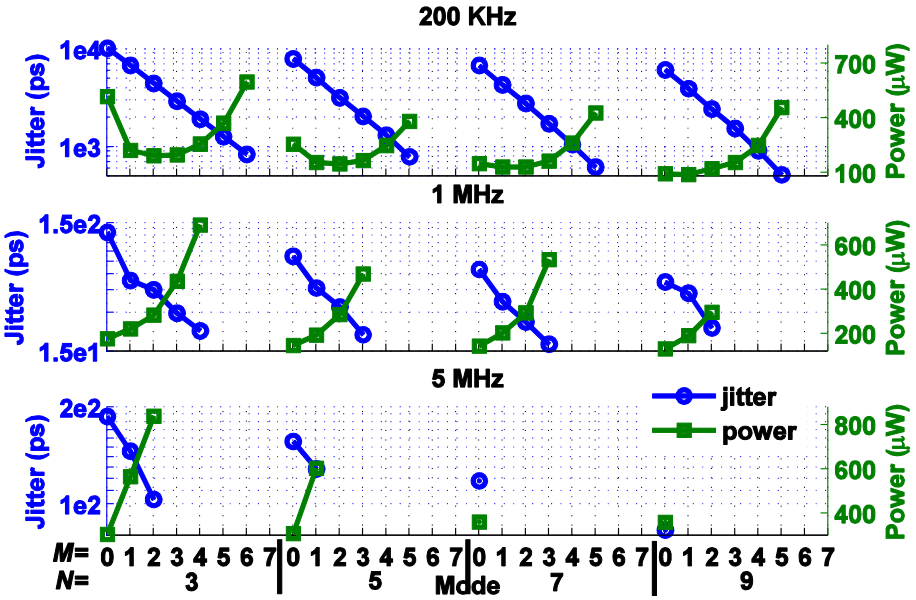


Figure 7.9. Jitter and power for three fixed output frequencies.

To further compare the overall performance of the oscillator under different conditions, we used a FOM that is defined as [65, 173-176]

$$\text{FOM} = \frac{\text{s}^2}{J^2} \cdot \frac{\text{Hz}}{\text{F}_{\text{out}}} \cdot \frac{\text{W}}{P} \tag{7.31}$$

where J^2 is the RMS jitter at the output and P is the total power. FOM is usually in units of dB and higher FOM represents better performance. The FOM is shown in

Figure 7.10. In all cases, higher N and M are favored with a preference for higher internal oscillation frequency over larger rings, i.e. higher M is better than higher N. As seen in Table 7.7 below, FOM is usually higher for (N,M+1) than (N+2,M) in the reported implementation. This further demonstrates that the frequency-boost technique not only reduces jitter but improves overall performance of the RO.

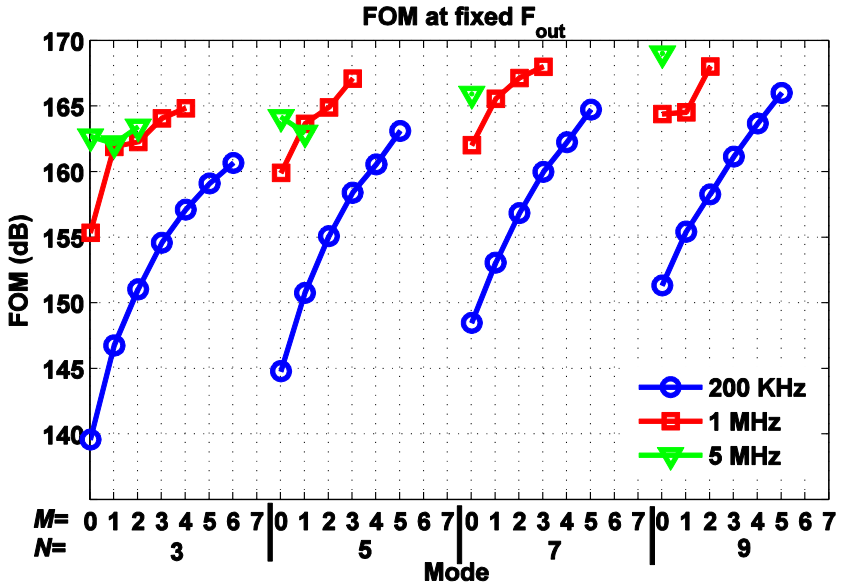


Figure 7.10. FOM for three fixed output frequencies.

B. Measurement Results for Fixed V_{in}

We measured the jitter and power for N = 3 and M = 0-7 with V_{in} fixed. The results are shown in Figure 7.11. This experiment investigates how the FD affects the overall performance because the RO signal should remain the same with N and V_{in} fixed. From the left to the next point, M increases by 1 and the output frequency decreases by a factor of two. As the number of DB2 stages increases, the jitter increases exponentially because the accumulation period increases as discussed in Section 7.2 ($J_{out}^2 = A \cdot J_{osc}^2$ while $A = 2^M$). The period jitter theoretically increases by a

factor of $2^{0.5M}$ (1.41^M) which agrees with the factor of 1.54^M found from the data of Figure 7.11 and one other similar set of data with N = 9 (we discarded jitter measurements that might have experience flooring).

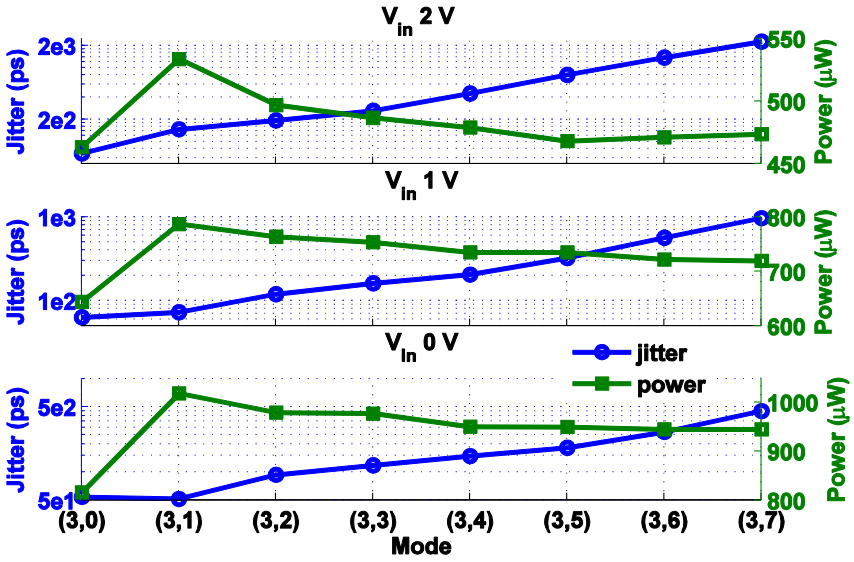


Figure 7.11. Jitter and power for fixed input voltage.

The resampling DFF consumes the most power in the FD section and power is proportional to operating frequency. This results in a huge increase of power at the second mode because the DFF is disabled in the first mode. From modes M = 1 to M = 7 the power decreases because the operating frequency of the DFF decreases as the number of DB2 stages increases.

C. FOM Measurement Results

A total of 99 data points across 28 modes were collected. FOMs for selected modes as a function of output frequencies is shown in Figure 7.12. The same trend was found as for the FOM at fixed output frequency: higher N and M are favored with a priority for M. Modes with nine inverters outperform other modes until they reach

the maximum output frequency. From low to high frequency the best mode changes from $M = 7$ to $M = 0$.

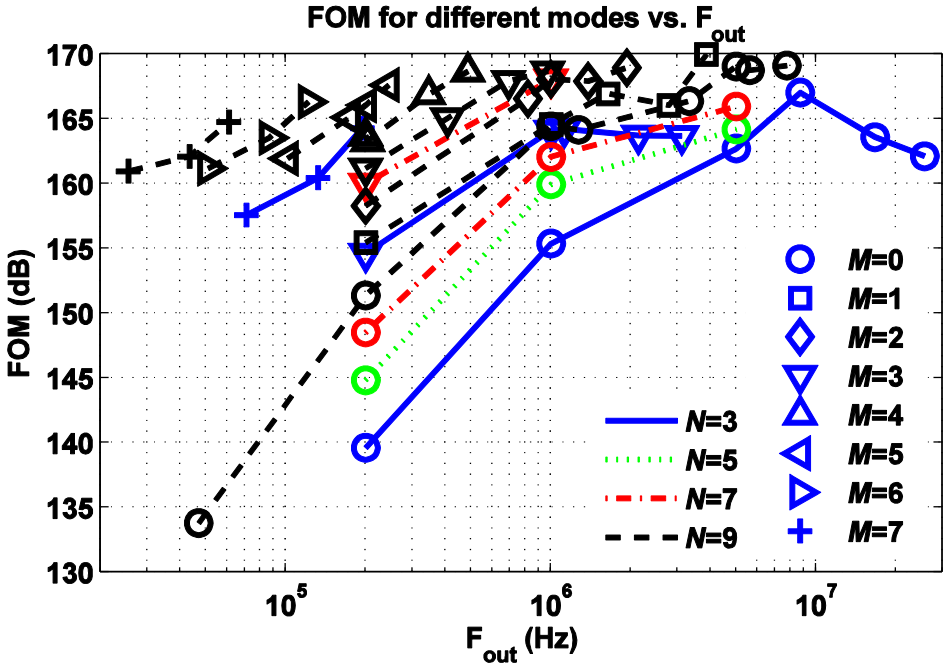


Figure 7.12. FOM for selected modes as a function of output frequency. The mode is represented by a combination of line style and markers.

7.4.6 Phase Noise

We have measured phase noise under two modes: (3,0) at 25 MHz and (9,2) at 1 MHz. The measurements are shown in Figure 7.13. Mode (3,0) has a phase noise of -108.1 dBc/Hz at 1 MHz offset; this corresponds to 31.5 ps RMS jitter assuming that the phase noise is dominated by thermal noise [72]. The measured RMS jitter was 54 ps but this measurement experienced equipment flooring and the predicted real jitter as in Appendix D was 30 ps which agrees with the phase noise converted jitter. Mode (9,2) has a phase noise of -112.6 dBc/Hz at 100 KHz offset. Its converted jitter is 235.4 ps which agrees with the jitter measurement of 229 ps.

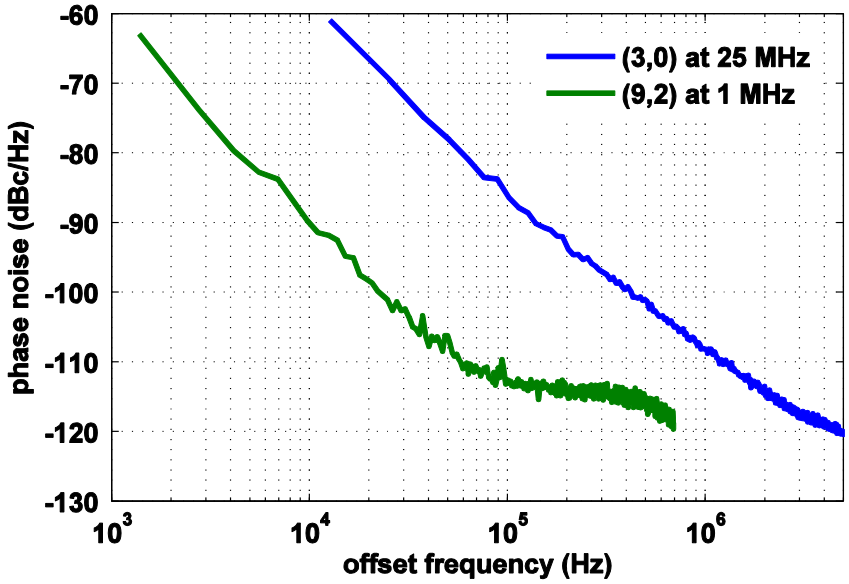


Figure 7.13. Phase noise measured at two conditions. Traces are average of 16 measurements.

7.4.7 Performance Summary and Comparison

Performance of this design (at two operating conditions) is compared with other reported works for timing signal generation in Table 7.3. The two reported performances are under modes (3,0) (25 MHz) and (9,2) (1 MHz), respectively. FOM defined in Equation 7.31 is used for comparison. Among these designs, the frequency-boost design achieves one of the lowest jitter in UI as well as highest FOM for the second condition. This is a substantial improvement knowing that these designs have been well studied and optimized for decades. To further understand how lowering supply voltage affects the performance, we measured performance of the frequency-boost design under a 3 V supply. One operating condition ((*N*,*M*) = (3,0) and V_{in} = 0 V) yielded an output frequency of 19 MHz. The power consumption and jitter were 186 μW (62 μA) and 66 ps, respectively, which corresponds to FOM of 168.1 dB.

Table 7.3 Comparison of Reported Works

	Yu [177]	Chen [178]	Höppner [176]	Cheng [64]	Sebastiano [179]	William [155]	Lasanen [152]	This work
Technology	90 nm	0.18 μm	65 nm	90 nm	65 nm	0.18 μm	0.35 μm	0.5 μm
Approach	Digital				Mixed-signal			
Design	DCO PVT	ADDB	ADPLL	ADPLL	MRO PVT	RO PLL	RO VT	RO
Supply (V)	1.0	1.8	1.2	0.5	1.2	1.8	1.5	5
Area (mm^2)	0.04	0.09	0.0078	0.02	0.11	0.16	0.077	0.15
Frequency (Hz)	1M–20M	250M-625M	83M–4000M	96M–720M	100K	100M–500M	200K–150M	<0.1–25.6M
Power (μW)	650 @5MHz	8400 @250MHz[b]	2700 @2GHz	570 @720MHz	41 @100KHz	24300	105 @5MHz[d]	817 @25MHz 297 @1MHz[f]
RMS Jitter (ps) / (UI)	240 / 1.2e-3 @5MHz	5.19 / 1.3e-3 @250MHz	5.4 / 1.1e-2 @2GHz	13.3 / 9.6e-3 @720MHz	52000 / 5.2e-3 @100KHz	4 / 1.08e-3 @270MHz[c]	1000 / 5e-3 @5MHz[e]	54 / 1.4e-3 @25MHz 229 / 2.3e-4 @1MHz
Voltage accuracy (%/V)	<2 @5MHz, 0.9–1.1V[a]	NA	NA	NA	1.5 @100KHz, 1.05–1.39V	NA	<0.28 @5MHz, 1.2–3V	3.5 @25MHz, 4.5–5.5V
FOM (dB)	157.3 @5MHz	162.5 @250MHz	158.0 @2GHz	161.4 @720MHz	139.6 @100KHz	159.8 @270MHz	152.8 @5MHz	162.3 @25MHz 168.1 @1MHz

DCO: digitally controlled oscillator; PVT: process, voltage, temperature compensation; ADB: all digital deskew buffer; ADPLL: all digital PLL; MRO: mobility-referenced oscillator; VT: voltage and temperature compensation

a: estimated from paper; b: has duty cycle correction function; c: accumulated jitter; d: has start-up circuit and sensing function; e: results from better samples; f: excluding the power of the driver

7.5 Discussion

7.5.1 Improvement of the New Jitter Model

The new jitter model (Equations 7.23 – 7.28) improves upon the prior model (Equations 7.9 and 7.10). In the old model the design parameters cannot be varied arbitrarily; parameters I_{INV}, L_N, V_{DD}, and f_{osc} have a relationship similar to the current-starved RO as in Equation 7.18 which introduces constraints among these parameters. The new model introduces an approximation of effective loading capacitance for the inverters in the RO and imposes this as a constraint to determine the relationship among these parameters. Parameter I_{CSI} is then replaced with the other three parameters. Thus, the design parameters in the new model can be varied arbitrarily unless the corresponding I_{CSI} is physically impossible to achieve. This yields a more intuitive and straightforward model. As a result, the model is a function of f_{osc}, V_{DD}, noise coefficients, design geometry (N, W_N, L_N), and constants. The model predicts jitter accurately over the entire operating range of the circuit (Figure 7.7).

7.5.2 Design Guidance for Current-Starved RO Design

The exploration of design space allows us to draw some conclusions and offer design guidance for current-starved ROs. Figure 7.6 and Figure 7.7 suggest that, to minimize jitter without power constrains, higher values of N and M corresponding to more inverter stages and DB2 stages (higher f_{osc}) are generally better at a fixed output frequency. However, in that case the maximum output frequency is strictly limited. This is consistent with the jitter model (Equations 7.23 – 7.28). The model suggests that in order to minimize jitter at a fixed output frequency, we need to have higher N,

M, V_{DD}, W_N/L_N, and L, but lower $W_{N,CS}/L_{N,CS}$ (this noise source vanishes when $W_{N,CS}/L_{N,CS} = 0$). Equation 7.26 (the dominant jitter contribution term) suggests that, as long as the desired output frequency can be achieved, the circuit should be configured to the highest N and M with a priority for M given the stronger dependence on M than N. Table 7.4 summarizes the N and M values that minimize jitter from measurements at different output frequencies. We also use Equation 7.26 to predict which (N,M) combinations minimize jitter as a function of frequency (labelled "Theory"). The measured optimal values were determined by using linear regression in the log domain to approximate the data for each mode (average slope was used for modes with only one data point) at specific frequencies. The order matches the theoretical model. The model also predicts some modes as optimal which were not found in the measurements; these modes are generally associated with lower N or M values. We attribute this mismatch to not having enough data points for some modes so the regression models used to determine the optimal modes are not precise. For fixed output frequencies (Figure 7.9), we summarize the tradeoffs between N and M in Table 7.6: to minimize jitter, bigger N and M are better; increasing M by 1 is better than N by 2; increasing N by 4 is better than M by 1; increasing M by 2 is better than N by 6.

Table 7.4 Design Scenario for Minimum Jitter

F_{out} (Hz)	Optimal N	Optimal M	Theory
$F_{out} \leq 60K$	9	7	1
$60K < F_{out} \leq 120K$	9	6	4
$120K < F_{out} \leq 190K$	3	7	6
$190K < F_{out} \leq 240K$	9	5	8
$240K < F_{out} \leq 280K$	7	5	9
$280K < F_{out} \leq 380K$	3	6	10
$380K < F_{out} \leq 480K$	9	4	12
$480K < F_{out} \leq 570K$	7	4	13
$570K < F_{out} \leq 770K$	3	5	14
$770K < F_{out} \leq 960K$	9	3	16
$960K < F_{out} \leq 1.1M$	7	3	17
$1.1M < F_{out} \leq 1.6M$	5	3	19
$1.6M < F_{out} \leq 2.3M$	7	2	21
$2.3M < F_{out} \leq 4.5M$	7	1	25
$4.5M < F_{out} \leq 7.7M$	9	0	28
$7.7M < F_{out} \leq 8.9M$	7	0	29
$8.9M < F_{out} \leq 13.0M$	5	0	31
$13.0M < F_{out} \leq 25M$	3	0	32

Missing modes (order (N,M)): 11 (5,5), 15 (5,4), 18 (3,4), 20 (9,2), 22 (3,3), 23 (5,2), 24 (9,1), 26 (3,2), 27 (5,1), 30 (3,1)
Modes without data points (order (N,M)): 2 (7,7), 3 (5,7), 5 (7,6), 7 (5,6)

Table 7.5 Design Guidance for Minimum Jitter at Fixed F_{out}

Guidance	Valid cases at F_{out}		
	200 KHz	1 MHz	5 MHz
$J(N+2,M) < J(N,M)$	18/18	10/11	4/4
$J(N,M+1) < J(N,M)$	21/21	12/12	3/3
$J(N,M+1) < J(N+2,M)$	15/16	9/10	2/3
$J(7,M) < J(3,M+1)$	6/6	3/4	1/1
$J(N,M+2) < J(N+6,M)$	5/5	3/3	0/1

When power is also an important design consideration, the FOM is a more relevant performance metric. To maximize FOM, more inverter stages are generally beneficial but the optimal number of DB2 stages will depend on the output frequency as shown in Figure 7.12. The model suggests that in order to maximize FOM at a fixed output frequency, we need to increase N, M, and L while lowering V_{DD}. This

agrees with the guidance for number of inverter stages and number of DB2 stages suggested by our measurements. The optimal N and M for maximum FOM from measurements at different output frequencies are summarized in Table 7.6. The optimal values were determined by using second degree polynomial regression in the log domain to approximate the data for each mode and estimating the optimal mode at specific frequencies. For FOMs at fixed output frequencies (Figure 7.10), we summarize the tradeoffs between N and M in Table 7.7. In general the design choices follow those suggested by minimizing jitter. However, the priority for M decreases as the output frequency increases. We attribute this to the power of the circuit being dominated by the FD such that adding an additional DB2 increases the power more than it decreases the jitter.

Table 7.6 Design Scenario for Maximum FOM

F_{out} (Hz)	Opt. N	Opt. M
$F_{out} \leq 60K$	9	7
$60K < F_{out} \leq 120K$	9	6
$120K < F_{out} \leq 240K$	9	5
$240K < F_{out} \leq 480K$	9	4
$480K < F_{out} \leq 960K$	9	3
$960K < F_{out} \leq 1.1M$	7	3
$1.1M < F_{out} \leq 2.3M$	7	2
$2.3M < F_{out} \leq 4.5M$	7	1
$4.5M < F_{out} \leq 7.7M$	9	0
$7.7M < F_{out} \leq 8.9M$	7	0
$8.9M < F_{out} \leq 25M$	3	0

Table 7.7 Design Guidance for Maximum FOM at Fixed F_{out}

Guidance	Valid cases at F_{out}		
	200 KHz	1 MHz	5 MHz
FOM(N+2,M) > FOM(N,M)	18/18	10/11	4/4
FOM(N,M+1) > FOM(N,M)	21/21	12/12	2/3
FOM(N,M+1) > FOM(N+2,M)	12/16	4/10	1/3
FOM(N+4,M) > FOM(N,M+1)	10/11	6/7	2/2

FOM(N,M) is the FOM for mode (N,M)

7.5.3 Technology Dependence

To investigate how technology affects the overall performance of the oscillator, we derive the theoretical FOM as defined in Equation 7.31. Here we assumed that total power consumption is dominated by the RO. RO power is taken to be $2 \cdot I_{CSI} \cdot V_{DD}$ (where the 2 accounts for the current bias generation and RO), and, thus,

$$\mathrm{FOM} = \frac{1}{\left(J_w^2 + J_f^2\right) \cdot \mathrm{F_{out}} \cdot I_{CSI} \cdot V_{DD}}. \tag{7.32}$$

Substituting Equations 7.18, 7.19, and 7.23 – 7.28 into Equation 7.32 we have

$$\mathrm{FOM} = \left(\frac{1}{\mathrm{FOM}_{w1}} + \frac{1}{\mathrm{FOM}_{w2}} + \frac{1}{\mathrm{FOM}_{f}} \right)^{-1} \tag{7.33}$$

$$\mathrm{FOM}_{w1} = \beta_4 V_{DD}^{-0.5} \mu_n^{-0.5} \sqrt{\frac{\mu_n}{\mu_p} + 1} L_N (\gamma_N + \gamma_P)^{-1} \tag{7.34}$$

$$\mathrm{FOM}_{w2} = \beta_5 \tag{7.35}$$

$$\mathrm{FOM}_{f} = \beta_6 V_{DD}^{-2} L_N^2 \left(\mu_n K_{fN} + \mu_p K_{fP} \right)^{-1} \tag{7.36}$$

where FOM_{w1} and FOM_{w2} are contributed by thermal noise, FOM_f is contributed by flicker noise, and β_4, β_5, and β_6 are parameters that do not vary with technology and can be expressed as

$$\beta_4 = 2^{-1.5} k_B^{-1} T^{-1} \sqrt{\beta_C} \sqrt{\frac{W_N}{L_N}} \left(\sqrt{\frac{W_N}{L_N}} + \sqrt{\frac{W_{N,CS}}{L_{N,CS}}} \right)^{-1} \sqrt{N} \sqrt{2^M} \sqrt{\mathrm{F_{out}}} \tag{7.37}$$

$$\beta_5 = 2^{-2} k_B^{-1} T^{-1} \tag{7.38}$$

$$\beta_6 = \frac{4}{25} \left(1 + \frac{1}{\beta_L^2} \right)^{-1} N 2^M \mathrm{F_{out}}. \tag{7.39}$$

The power of auxiliary circuits like the PLL or FD was not considered. The three contributions to FOM (Equations 7.34, 7.35, and 7.36) and FOM (Equation 7.33) are listed in Table 7.8. The process parameters used here were from SPICE models released by the foundry [169]. In the calculation we assume that both thermal and flicker noise coefficients are constant across technologies.

To understand more precisely the effects of technology scaling, variations in noise coefficients need to be considered. Studies have demonstrated that thermal noise coefficients, γ_N and γ_P, are relatively insensitive to technology [170, 180] but increase for short channel devices [170]. However, the coefficients increase by less than 2x from channel lengths of 500 nm to 40 nm [170], by less than 1.5x from 240 nm to 100 nm [172], and less than 20% from 2 μm to 200 nm [181]. On the other hand, flicker noise coefficients, K_{fN} and K_{fP}, are insensitive to channel length but depend on technology. However, they vary less than 50% across 0.35 μm, 0.25 μm, and 0.18 μm technologies [180]. Noise scales linearly with these noise coefficients so they don't significantly change the trend in Table 7.8. Although the benefits of more advanced technologies are modest in Table 7.8, there is greater opportunity for frequency boosting and associated FOM improvement due to higher transition frequencies. One additional DB2 stage brings a 1.5 dB and 3 dB increase to FOM_{w1} and FOM_f, respectively. Therefore, we believe that the frequency-boost technique will offer additional benefits when implemented in more advanced technologies. While the experimental validation has been carried out with a relatively long channel technology, the benefits are expected to be even greater when this technique is used in smaller feature size technologies.

Table 7.8 FOM for Different Technologies

Technology (nm)	FOM_{w1}	FOM_{w2}	FOM_f	FOM
500*	167.0	197.8	178.9	166.7
350	166.9	197.8	180.6	166.7
180	166.0	197.8	181.1	165.8
90	163.8	197.8	178.1	163.7
65	163.9	197.8	179.4	163.7

Parameters used in the calculation: L_N = 2 x minimum feature size, W_N/L_N = 1, β_C = 1, f_{osc} = 10e6, N = 3, $\gamma_N = \gamma_P$ = 2/3, $K_{fN} = K_{fP}$ = 1e-24, β_l = 3, $W_{N,CS}/L_{N,CS}$ = 5
* Technology for this work

7.5.4 Applications of Frequency-Boost Technique

The frequency-boost technique generally applies to standalone ROs or ROs incorporated into PLLs. However, when incorporated in a PLL special attention should be paid to the overall PLL stability performance because frequency division is equivalent to scaling the VCO gain; loop characteristics should be adjusted accordingly to optimize stability. For the purposes of lowering thermal and flicker noise-induced jitter in standalone ROs, this technique is valid for ROs meeting these two conditions: 1) Maximum oscillation frequency is higher than the targeted output frequency so there is room to allow frequency division; 2) The dominant jitter variance is inversely proportional to the oscillation frequency raised to a power greater than one. The benefit of this technique is greater if the frequency dependence is stronger. For the purpose of achieving better FOM the first condition still holds while the second condition becomes stricter. The dominant jitter variance should be inversely proportional to the oscillation frequency raised to a power greater than two under the assumption that power scales linearly with oscillation frequency.

This technique also presents an opportunity to tradeoff between power and jitter performance. It can be potentially used for dynamic clock scaling in dynamic

system management to achieve modest jitter with low power and high efficiency, or lower jitter with increased power and reduced efficiency, in analogy to dynamic voltage frequency scaling (DVFS) in VLSI systems [182]. DVFS adjusts supply voltage and operating frequency to tradeoff performance and power while this design can potentially adjusts clock quality to tradeoff performance and power; when the system requirements are relaxed, a lower quality clock can be used to save power and vice versa.

Chapter 8: Conclusion and Future Work

8.1 Conclusion

In this dissertation we developed enabling hardware technologies that represent progress towards achieving autonomous tiny robots (from a few cm^3 in volume to much less than one cm^3). The work focused on the design of electronics and motion mechanisms. State-of-the-art robots were reviewed and analyzed; we came to the conclusion that current approaches for large robots cannot achieve a high level of autonomy at tiny scales. Technical problems of current approaches were identified: excessive use of commercial off-the-shelf electronics modules and mechanisms, centralized processing, and the use of multiple chips or boards for integration of high-voltage devices. These result in technical hurdles at tiny scales for compact integration, fast processing, and the unavailability of COTS motion mechanisms. In order to address these technical problems and difficulties, we proposed a decentralized single application-specific integrated circuit (ASIC). The electronics units offer actuation, control, power supply, and sensing functions in parallel. Locomotion is provided by post-fabricated actuators (and mechanisms) on the AISC. This modular architecture can adapt to different combinations of functions according to the task.

We presented a circuit architecture that is able to generate two phased square signals with tunable frequency for actuation control. Digital logic circuits were used to generate different duty cycles and overlap between two controls. This technique

can be applied to the generation of other types of control signals. Ultralow frequency of the control signal was achieved by the introduction of a frequency divider. Furthermore, we presented a design methodology based on an analytical and model-based framework; the goal was to minimize the weighted product of area, power, and frequency error. A simulation-based model of effective capacitance of the circuit was developed to enable this optimization. This model links these three performance metrics to design parameters and also reduces the dimensionality of the optimization problem. The optimization flow produces a solution that minimizes the weighted product and meets design constraints if any. Design examples were given for a 0.5μm CMOS technology with 3.3 V supply.

Programmability is important for robots to adapt to changes in the environment or tasks. The ability to store information or command sequences is also important for autonomous robots. Otherwise, an initialization is required every time the robots are powered up. We used the actuation signal generator as an example of frequency storage and programming. This was enabled by incorporating the actuation signal generator into a floating-gate phase-locked loop. This programming loop was able to automatically operate with only one frequency reference and power supplies. Furthermore, the information remained stored even after power off. Measurement results showed that the circuit worked as expected. The programming range at different biases was characterized. The equivalent programming resolution was 6 bits compared to 4 bits for state-of-the-art solid state memory. Moreover, this chip was integrated on a legged robot to demonstrate gait control.

Tiny robots need high-voltage devices to drive tiny actuators and to program floating gates as in our case. We have reported implementation results for a family of high-voltage n-type metal-oxide-semiconductor devices. N-well and field oxide buffer regions in the drain were introduced to suppress the avalanche and surface breakdown effects. A total of 63 separate devices in four configurations were fabricated in three different runs of a 0.5 μm standard 5V CMOS technology. This work reports the highest known breakdown voltages that have been achieved in similar technology and provides the first direct comparison between drain-centered and source-centered circular devices fabricated in the same technology. Measurement results showed that a circular structure with a central drain has the highest breakdown voltage as well as the highest transconductance which is comparable to standard transistors in the same run. Other parameters of the devices including threshold voltage and Early voltage were also characterized.

We described the fabrication and optimization of thermal actuators using complementary metal-oxide-semiconductor (CMOS) compatible procedures. The actuators and the ASIC chip have to be designed jointly to realize the integration of these heterogeneous systems. The actuators were optimized to produce forces and bending angles that are necessary to lift the ASIC chip and to produce reasonable displacement. The implementation results showed that the actuators were able to bend properly and lift three times the weight of the CMOS chip. The prototyping actuators fabricated on a dummy wafer could potentially be ported to a CMOS chip in the future. Challenges and design guidance for this effort are discussed in Appendix C.

Many computations are performed for different robotic functionalities where fast and energy efficient processing with minimum delay is required. We used odometry as an example to demonstrate the feasibility of using mixed-signal circuits in this application. Detailed analysis and simulation results of the dynamics simulator suggested that mixed-signal implementation can dramatically reduce power consumption at an acceptable loss in precision. Additionally, detailed analysis and design guidelines for a sine shaping circuit were presented. Transconductance attenuation and resistive source degeneration were applied to achieve arbitrary input voltage range with high accuracy. The design procedure and analysis were verified with two examples in simulation and one example in measurement.

Jitter greatly affects the timing margin and overall performance of a computational system. A frequency-boost technique to reduce jitter was proposed. A reconfigurable oscillator circuit was implemented and can be used for dynamic power and performance management, which is highly desirable in a resource-constrained system. While this technique was developed and demonstrated for a ring oscillator, it can apply to any oscillator whose jitter scales inversely with oscillation frequency. Jitter reduction was verified by measurement results for chip fabricated in a 0.5 μm CMOS process. This design achieved the best phase noise figure of merit at one operating frequency and competitive performance at other frequencies compared with other reported designs. This figure of merit is a commonly used metric to compare across oscillators operating at different settings and fabricated using different technologies. The performance advantages are expected to readily extend to more advanced technologies.

Tiny robots have a lot of potential and are expected to redefine and broaden many applications. Although current tiny robots are still far behind their natural counterparts, ants and bees, these gaps identify the lack of existing platforms and indicate future directions for tiny robot research. This dissertation demonstrated that robotic functions can be implemented under strict size and power constraints. As discussed in Chapter 1, there are four major function components for an autonomous tiny robot. We contributed substantially to actuation and control and also addressed power, sensing, and the integration of these components. These results help to advance a step towards the implementation of autonomous tiny robots and will help to eventually realize the tiny robots that are comparable to biological systems.

8.2 Future Work

There are a few research tracks that can be extended from this work to achieve an autonomous tiny robot: 1) development a system on chip (SoC) ASIC that can fulfill the electronics requirements for the robot to perform simple tasks; 2) integration of MEMS actuators with CMOS chips; 3) wireless power supply with an onboard energy unit.

To complete the SoC, more function components will have to be integrated in the CMOS chip. This is similar to the main processing chip in our smart phones and computers, which shows that current technology is capable of integrating multiple sophisticated functions on chip. Temperature [37, 38] and image [39, 40] sensors can also be integrated on chip while pressure [41, 42] and gas [43, 44] sensors can be integrated on top of the chip. In the latter case the weight of the sensors has to be light

so they do not overload the legs. The position of the sensor has to be carefully designed so that it does not affect the function and the fabrication of the actuators.

A straightforward method to integrate MEMS devices with CMOS dies is to do wafer level processing [112, 113]: ordering a whole CMOS wafer or a large portion of it and fabricating the MEMS devices on this large substrate using the fabrication procedure discussed in Section 5.3. One minor change to the procedures is that the wafer should be diced before releasing as we discussed in Appendix A.8 that releasing should be the last processing step. However, ordering a whole or a larger area of CMOS wafer is costly (more than $10,000 for a wafer fabricated with a relatively old CMOS technology). One other method that Appendix B suggested is to increase the die size and save enough buffer region at the periphery so there is enough space to perform edge bead removal either by hands using Q-tips or by a machine. Another method is to adopt a two-chip solution: one CMOS chip and one MEMS chip. The MEMS devices are still fabricated the same way on a large substrate but, again, diced before releasing. The two chips can be different dimensions. Then the two chips are glued together on both of their back sides as shown in Figure 8.1. There are a few potential ways for creating electrical connections. They can be made of patterned metal layers but it will be challenging to pattern all faces of the stacked chips especially the narrow side faces. Any sorts of conductive materials that can be applied with precise control of amount and positions might also work, for example conductive paste or solder. Through-silicon via (TSV) is a technique to create vertical electrical connections passing through the whole chip and has recently become popular for 3D IC packaging [183]. Instead of creating wraparound connections,

through holes are created on the chips and then filled with metal to create vertical connections. The holes can be created on two chips separately and then align the chip to glue them or the chips can be glued first and then create through holes once. It will require more investigations to determine which order produces a higher yield. One concern for the two-chip solution is that the weight increases. Fortunately, this can be solved by thinning one or both chips down to decrease the weight.

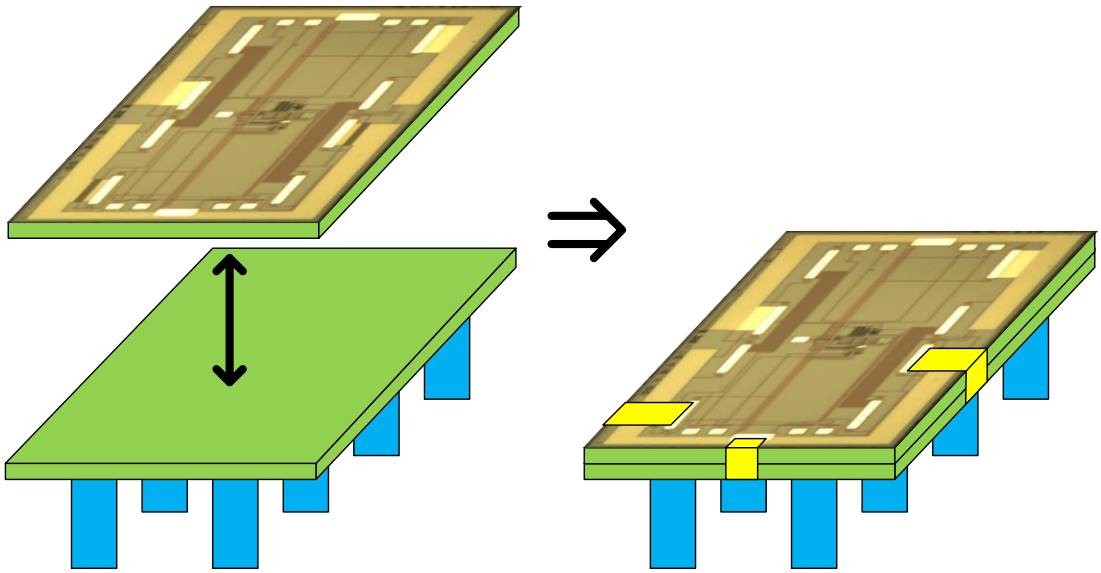


Figure 8.1. One implementation of the two-chip solution for the legged chip. The CMOS chip and the MEMS chip are aligned and glued together. Control signal is applied through wraparound electrical connections (yellow) from the top face of the CMOS chip to the device side of the MEMS chip. Blue legs are for demonstration purpose. Releasing might have to be done after staking and creating connections.

It seems to be a practical and reasonable powering strategy to have an external wireless power source at the station with an onboard energy unit that is large enough to sustain the robots for a few times longer than the mission requires [20]. On board energy units can be batteries, accumulators, supercapacitors, springs, solar cells, or fuel cells [13] that can be integrated on a CMOS chip or are small enough for the robot to carry.

For a huge swarm of robots it is important to have a scalable, efficient, and wide-coverage power source. The wider the coverage of the power sources the more flexibility and functionality the robots are able to gain. We investigated several techniques that can be used to power the robots wirelessly including photovoltaics, radio frequency (RF), inductive coupling, and resonant magnetic coupling. Photovoltaics require a high intensity light source and line-of-sight. A light source is not easily scalable and is inefficient due to an additional conversion from electrical energy into light energy, assuming most accessible energy is electricity. The necessity of line-of-sight makes this technique geometrically limited. The transmitting power for RF is strictly limited by the Federal Communications Commission regulations so it is difficult for the robots to receive enough power using this technique. The design tradeoff for antennas between directivity and coverage range is also hard to balance. Researchers have reported 222 μW [184] and 427 μW [185] with a directive RF source beaming at the receiving antennas (lengths are 15 cm and 20 cm respectively) which are 1 m away from the source. Inductive coupling as discussed in MiCRoN robot can only apply to a short distance of mm scale and is difficult to scale. WiTricity [186, 187] demonstrating transfer of 60 watts over a distance of 2 meters. Other follow-up research efforts have shown that resonant magnetic coupling (shown in Figure 8.2) is a promising technique for mid-range highly efficient wireless power transfer. The decline in efficiency due to misalignment of the transmitting and receiving coils is not significant [186]. High efficiency, relatively large coverage, and low geometric dependence make resonant magnetic coupling a promising candidate for this application.

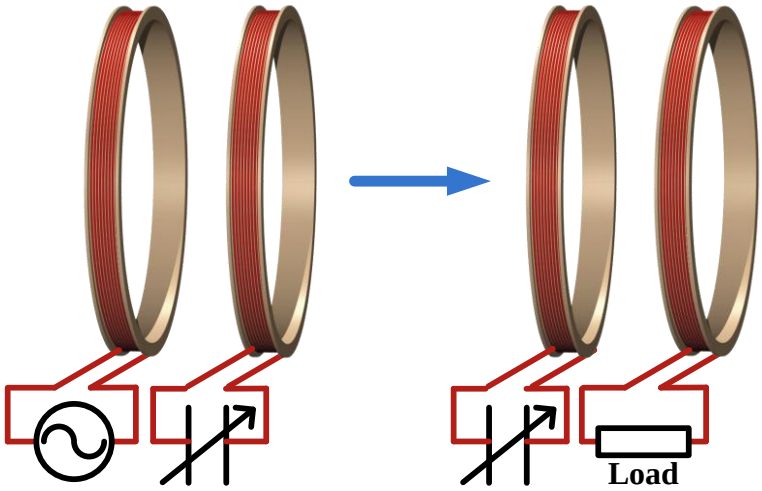


Figure 8.2. Resonant magnetic system. Transmitter and receiver both have a pair of coils to enhance the efficiency. Four coils resonate at the same frequency.

Appendix A. Discussion of Fabrication Procedures

A.1 Material Selection for Sacrificial Layer and Electrode Layer

The materials we tested as candidates for the sacrificial layer and the electrode layer were summarized in Table 5.5. In this section we describe why the materials for Process 1, 2, and 3 did not work and why changes needed to be made.

Process 1

The usage of SiO_2 as the sacrificial layer proved to be difficult. The main problem was that it was hard to determine when etching was done using a substrate with silicon oxide on top or even a bare silicon wafer. These two types of wafers are most popular and are used as substrates for microfabrication. However, silicon wafers naturally oxidize and grow a thin layer of silicon oxide [188, 189].

For process 1 we deposited silicon oxide on the samples using an Oxford Plasmalab System 100 which is a plasma enhanced chemical vapor deposition system. The deposition temperature was set to 300 ℃. Photoresist 1813 was then deposited and patterned on top. The buffered oxide etchant (BOE) etches both the sacrificial layer and the substrate as shown in Figure A.1. Although the etchant removed the sacrificial layer faster than the wafer because the sacrificial layer has lower quality, it was still difficult to determine when etching is done. Normally we etch more than the time required but this method does not work here.

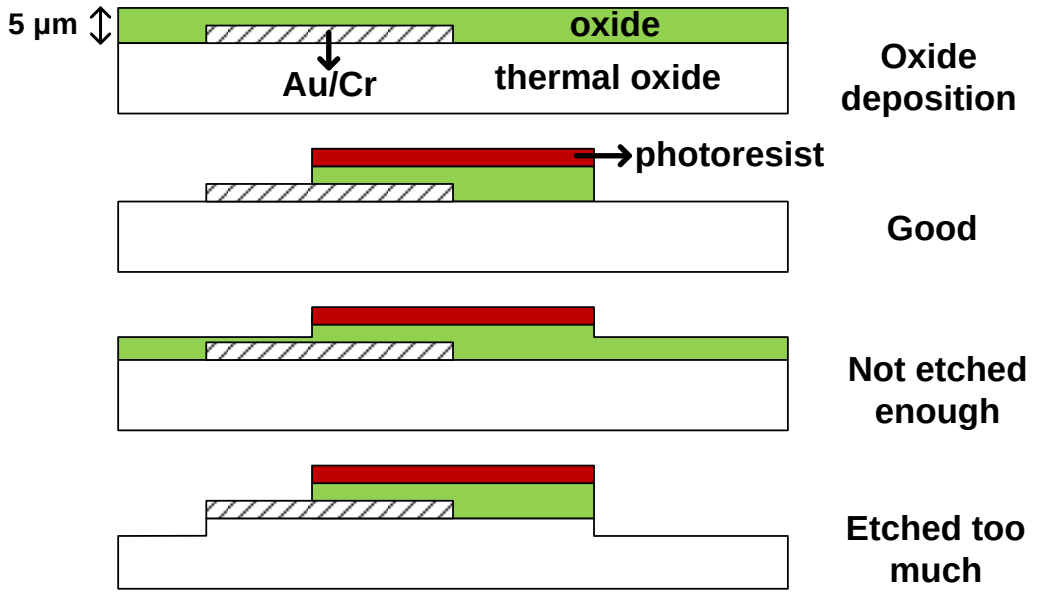


Figure A.1. The oxide etchant etches both the sacrificial layer and the oxide substrate so that it becomes difficult to determine when etching is done.

Process 2

The unsuccessful experience using silicon oxide as the sacrificial layer led us to consider other materials. We chose aluminum because our colleagues previously found it easy to pattern. The aluminum tested in Process 2 was deposited using a Metra Thermal Evaporator. Using aluminum as the sacrificial layer worked well but its combination with copper as the electrode became problematic. We found that the commercial etchant we used to etch aluminum (Transene Aluminum Etchant – Type A) etched copper faster than aluminum [190]. Given that the copper layer (electrode) is thinner than the aluminum layer (sacrificial), it is inevitable that the copper layer would be removed completely before the aluminum layer. Even if the copper layer is protected by photoresist, this result still does not change because part of the aluminum layer is also covered by SU-8 and requires undercut etching. Therefore, we sought other aluminum etchants that do not etch copper. We experimented with

phosphoric acid (H_3PO_4) and potassium hydroxide (KOH) [190]. A commercial 99% phosphoric acid and KOH mixed with water 3:7 by weight were found to etch aluminum too fast and SU-8 legs came off during etching. Nevertheless, the etchants did not react with or etch SU-8. KOH mixed with water 1:19 by weight did not etch too fast and could keep SU-8 attached to the substrate. However, both etchants were found to react with copper because we observed that part of the copper surface turned color (but not to etch copper). Although the copper was still conductive after being exposed to the etchants, we were not confident that the copper could remain functional after placing it in the etchant for the time required to release the devices. As a result, these two etchants were deemed unsuitable for our applications.

Process 3

Although we could not use aluminum and copper as the sacrificial layer and the electrode layer, respectively, we could possibly use them in reversed roles, aluminum for electrode layer and copper for sacrificial layer. The motivation for this was that we did not care if the copper layer (sacrificial) was etched a little during patterning of the aluminum layer (electrode). The copper layer tested in Process 3 was deposited using a Metra Thermal Evaporator. However, we found that SU-8 has poor adhesion to copper as shown in Figure A.2. Some legs peeled during developing.

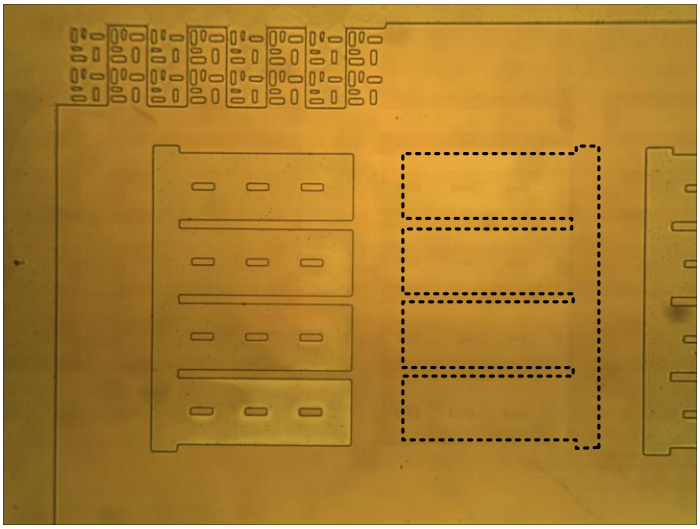


Figure A.2. Pattern SU-8 on a copper substrate. One row of legs did not adhere to the substrate and completely came off during developing. Approximated location is outlined with dotted lines.

Process 4

The difficulties of finding compatible materials and etchants for the three different layers, metal protection, sacrificial, and electrode, motivated us to reuse some materials for multiple layers. We decided to reuse chromium/gold because they are easy to work with. However, using gold for the sacrificial layer is too costly because this layer is thick. A better option was to use chromium/gold for the protection and electrode layer. We also chose aluminum as the sacrificial layer since silicon oxide and copper did not work well. In order to reuse chromium/gold, the mask had to be redesigned so that the pad protection layer was protected during patterning of the electrode layer. The new mask design will be discussed in the next section. The procedures for Process 4 were described in Section 5.3.

A.2 New Mask Design

The new mask had four major changes compared to the first one. First, we added wet etch vias on the leg layer 1 as shown in Figure A.3 to decrease the time required to release the legs. This method was used in other work to release their structures [111-113]. We found from our previous experiments that releasing the actuators took more than 48 hours because a large portion of the sacrificial layer is covered by the legs and the etchant has only limited access to the targeted material (height of the sacrificial layer is 1 μm). The vias were placed in positions to equalize the required time for etching from all directions. Second, the new mask had a wider anchor for the legs as shown in Figure A.3. It was increased from 50 μm to 70 μm and the leg length was decreased by 20 μm accordingly. The reason for this increase was because, from our experiences using the old mask, we had experienced poor adhesion for the SU-8 during both developing the SU-8 and releasing the legs. We hypothesized that due to a small anchor even a weak liquid flow could destroy the structure. Third, we extended the sacrificial layer beyond the legs in the y-direction so that the error tolerance for alignment during exposure was increased. Previously, the top-most and bottom-most edges of the legs aligned with the edges of the sacrificial layer. Fourth, we combined the mask for pad protection with the mask for the leg layer 2 (electrode) as shown in Figure A.4. The reason was to protect the gold deposited in the pad protection layer because we reused chromium/gold for these two layers as discussed in the previous section. The old masks were used for Processes 1, 2, and 3 while the new masks were only used for Process 4.

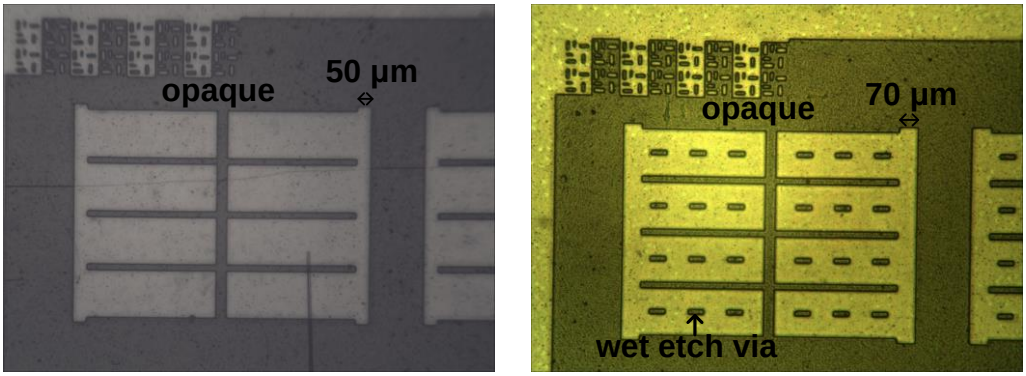


Figure A.3. Pictures of old (left) and new (right) mask designs for leg layer 1. The new mask has wider anchor and wet etch vias. The masks were used for a negative photoresist (SU-8) so the opaque area would have no photoresist left after patterning. Due to different light conditions and white balance, the pictures show different colors.

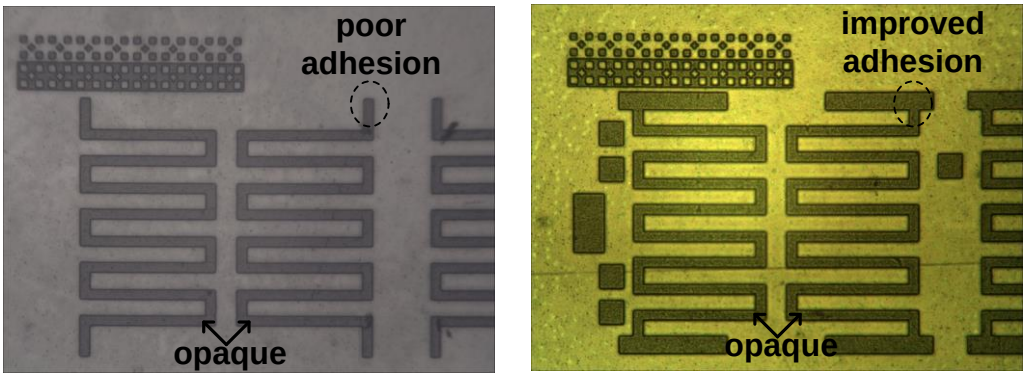


Figure A.4. Pictures of old (left) and new (right) mask designs for leg layer 2. The new mask has a duplicate mask for pad protection to, again, protect the gold previously deposited on the pads. The masks were used for a positive photoresist (AZ 9260) so the opaque area would be the photoresist left after patterning. Due to different light conditions and white balance, the pictures show different colors.

A.3 Electrode Continuity from Leg Surface to Pad

From step 4 of Figure 5.4 and Figure A.4, we observe that the electrode has to continue from the surface of the leg anchor to the sidewall of the leg, down to the substrate, and to the pad. The left picture of Figure A.4 shows that there is only a thin line (labeled “poor adhesion”) all the way to the pad to make an electrical connection.

This thin pattern has poor adhesion as indicated in Figure A.5. The electrode layer after patterning was supposed to be identical to its etching mask, AZ 9260. Unfortunately, the electrode layer down to the substrate level (thin pads are assumed to be the same level as the substrate) shrank, which indicated that the AZ 9260 had poor adhesion at that area and the etchant came underneath to etch the electrode. Sometimes the shrinkage only happened at the pad area. We hypothesized that this thin pattern might not adhere strongly to the substrate and the pad, becoming sustained above the surface like a cantilever. This problem was mitigated using the new mask. When the electrode layer extends to the pads, it covers the entire pad and has stronger adhesion than a thin pattern as shown in the right picture of Figure A.4. However, the yield only improved to close to 50%. Figure 5.5, Figure 5.10, and Figure 5.11 show continuity of the electrode; we applied signal to the pads and the signal was correctly transferred to the electrode and actuated the legs.

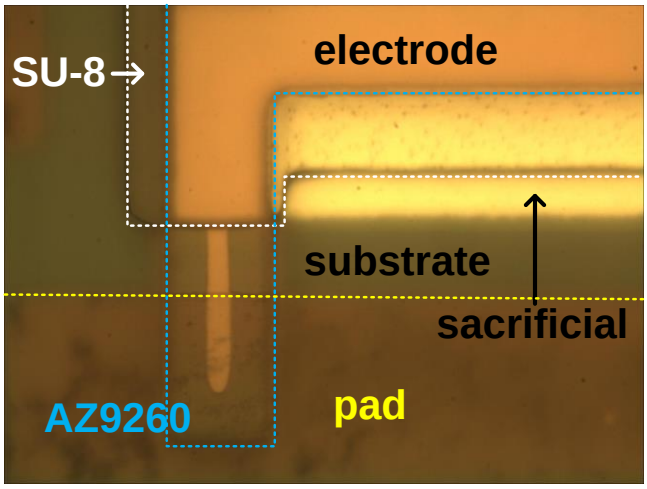


Figure A.5. Picture of the area around the tip of leg anchor and the pad taken after patterning the electrode layer in Process 2. The electrode and sacrificial layers are copper and aluminum, respectively. The SU-8, the AZ 9260, and the pad are labeled and outlined with white, blue, and yellow, respectively. The electrode layer is supposed to be identical to the AZ 9260 etching mask. Layers from bottom to top are substrate, pad and sacrificial, SU-8, electrode, and AZ 9260.

A.4 Post Bake 1813 to Improve Adhesion

Photoresist 1813 was used in Step 1.2 and Step 2.2. We have found that post baking the photoresist at 115 ℃ for at least one minute before placing the chip in the etchant at Step 2.3 helped promote adhesion of the photoresist in Process 1 (see Table 5.5). This poor adhesion was not observed at step 1.3 because both chromium and gold are thin and the required etching time is short. Without post baking, 1813 sometimes lifted and broke at the edges during etching as shown in Figure A.6 (a). This non-flat and rough-edged sacrificial layer caused problems in subsequent steps. One major issue was the adhesion of the SU-8. Minor or local lifting of the SU-8 could not be easily seen under microscope after developing. However, it became evident after depositing the electrode layer as shown in Figure A.6 (b). The legs that did not adhere to substrate well enough would usually completely come off during the release step.

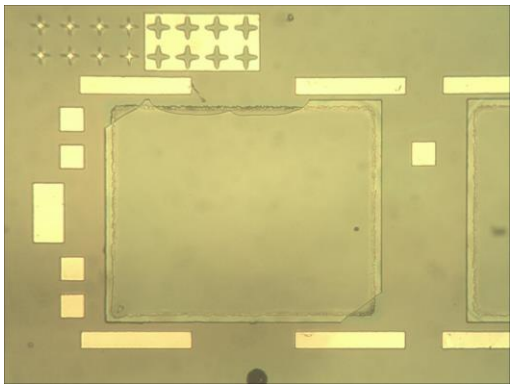
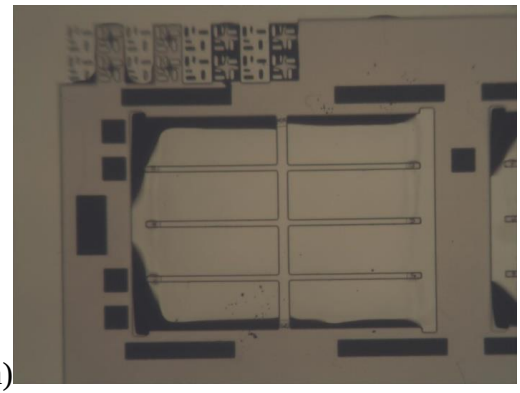

(a)

(b)

Figure A.6. (a) One sample after etching the sacrificial layer (Step 2.3). The sacrificial layer (the rectangle in the middle with rough edges) used here was SiO_2 as Process 1 in Table 5.5 and the etchant was buffered oxide etchant. Photoresist 1813 (the incomplete rectangle on sacrificial layer) lifted and broke at the edges. (b) Another sample after depositing electrode layer (Step 4.1). The electrode layer used here was Cu (Process 1). The dark color around the legs caused by missing direct light reflection indicate that the legs are not flat due to poor adhesion. The pads were also dark because there were residues left from previous steps making pads non-flat.

A.5 Al Deposition at Step 2.1

The deposition rate of aluminum cannot be too high. We have observed that speckles might form on the aluminum surface at a deposition rate of 30 Å/sec (thermal evaporation). The speckles are shown in Figure A.7. The aluminum was flat without speckles at deposition rates below 20 Å/sec. Although we did not observe the speckles to affect the process, it is still advised to avoid this abnormal situation.

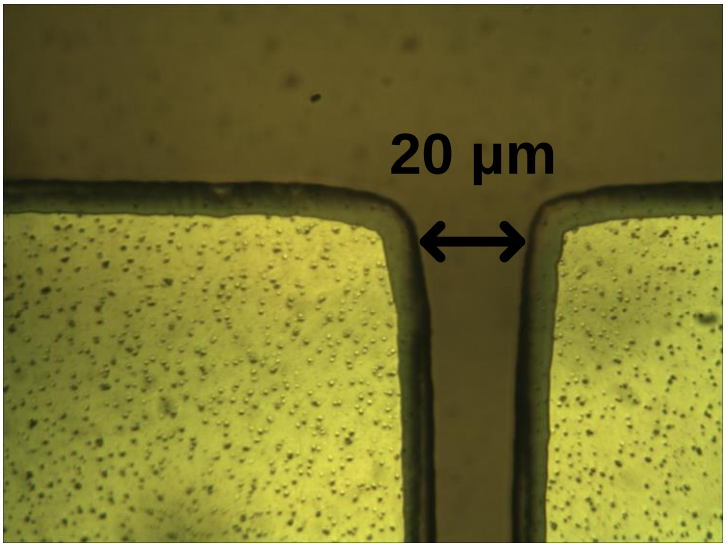


Figure A.7. Speckles on the aluminum surface (bottom left and bottom right rectangles). This photo shows aluminum after being patterned but speckles existed after deposition. Photomicrograph taken at 50x magnification after patterning.

A.6 Patterning of AZ 9260 photoresist at Step 4.2

One interesting observation was that AZ 9260 has orange peel like surface after spin coating as shown in Figure A.8. The height difference can be as high as 2 µm. This was proved to be a normal feature according to the manufacturer and did not adversely affect the successful patterning of the photoresist.

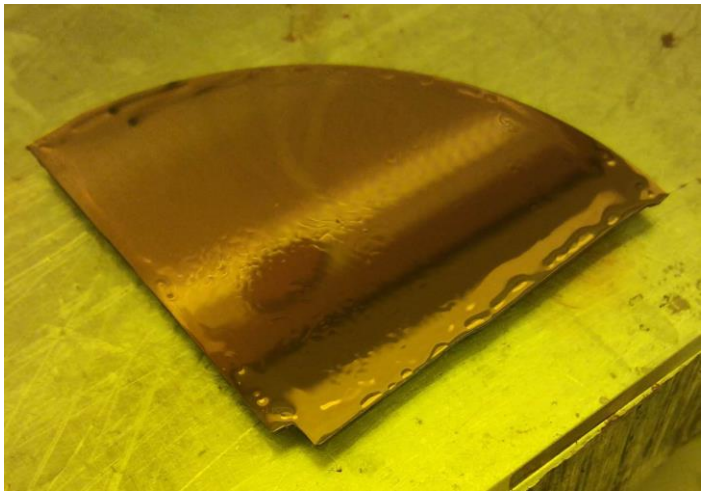


Figure A.8. Orange peel like surface of AZ 9260 on a quarter wafer. There are edge beads forming at the edges of the substrate. More discussion of edge bead will be given in Appendix B.

Descum is required for AZ 9260 because a layer of residue still remains after patterning. This residue could not be seen using optical examination so we hypothesized its thickness was thin. However, it prevented the metal underneath being etched (see Figure A.9). Descum was done using the March Jupiter III O2 plasma system with radio frequency power of 20 W and oxygen flow of 5 mTorr for 10 seconds.

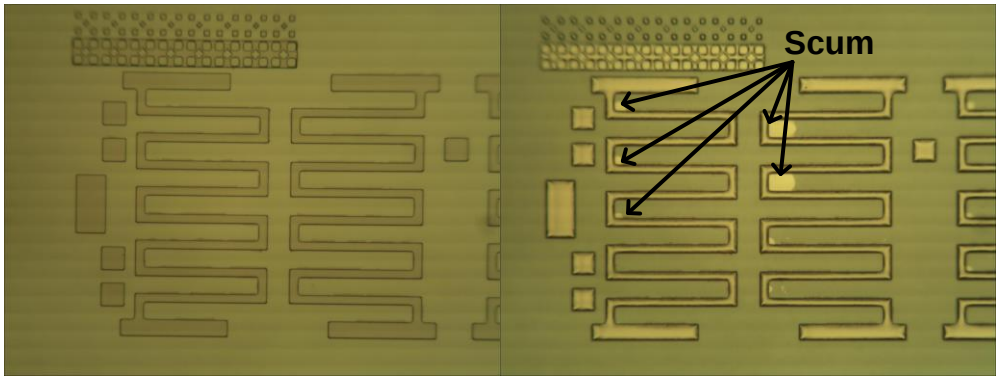


Figure A.9. (Left) Photomicrograph taken after developing. Darker color represents the photoresist; brighter color is gold substrate. The picture shows a perfect pattern with no sign of any residual photoresist. (Right) After etching some metal was not etched due to the residual scum. Both pictures were taken at 5x magnification.

Another interesting observation was that AZ 9260 was sensitive to the baking temperature. Current soft baking condition is 3 minutes at 110 ℃. We found that if we baked at 115 ℃ for 3 minutes, the photoresist became hard to develop. Even after developing for more than 20 minutes, a thin layer of photoresist which was supposed to be gone still remained on the substrate and was visible, unlike scum.

A.7 SU-8 Optimization

SU-8 and SU-8 2000 series were known to have poor adhesion so a SU-8 3000 series with improved adhesion was tested and successfully released. Moreover, adhesion promoter HMDS only works for the SU-8 3000 series. We have experienced adhesion problems with SU-8 and so its recipe optimization has been difficult. From our experience the most critical step is the post exposure baking time. In general there is a tradeoff between adhesion and development; the longer the baking time, the better the adhesion but the more difficult it becomes to develop the SU-8, and vice versa. A step temperature profile (subsequently 60 ℃, 95 ℃, and 60 ℃) was adopted to avoid sudden change of temperature for SU-8 to improve adhesion. We found that SU-8 residues did not develop (see the left image of Figure A.10) when the baking time was too long, especially for the first 60 ℃ bake. If the 95 ℃ bake was too long, cracks were found on the SU-8 as shown in the right image of Figure A.10. Overall there is only a small window on the baking conditions for the SU-8 to work properly.

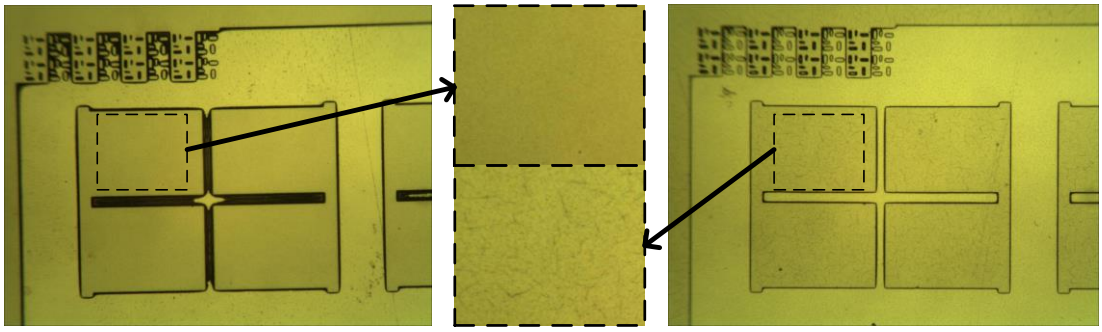


Figure A.10. SU-8 2005 on aluminum substrate after development. (left) Some SU-8 between legs could not be developed properly. (right) Cracks on SU-8. (middle) Two zoomed-in views. Dashed lines in the left and right photos identify the zoomed-in areas. The bottom view shows the cracks.

A.8 No Processing after Releasing

It is normally recommended that the releasing should be the last step in the process because the released structures are usually fragile. We have confirmed this in one experiment for Process 3. We were unsure if the copper etchant used to release the structure would affect the aluminum electrode after a long time (there was no reaction for short term). Therefore, we did not remove AZ 9260 and used it as protection for the electrode. After the releasing is done, we tried to remove AZ 9260 by placing the samples in acetone, but the legs were not robust enough to survive and most of them came off.

A.9 Heat Mass of the Legs

We observed in our actuation experiments that the actuators responded quickly; actuation occurred simultaneously with the application of the driving signal. Applying a lower current to the actuators would not produce actuation even after waiting for a long time. These factors indicated that the heat mass of the legs is small

and heat dissipation is relatively large. Consequently, heat did not accumulate on the legs. This observation also indicated that our initial thought to keep the actuation duty cycle low (one fourth and one eighth in our design) so that the actuators have enough time to passively cool down is incorrect. Therefore, the duty cycle should be designed based on the mechanical response time of the legs instead of heat response.

A.10 Robustness of the Actuators

In our experiments, we have tested the actuators multiple times during an eight-month period while they have been stored in a standard cleanroom environment. We have observed that the actuators are robust and function properly. As long as they are actuated with an appropriate amount of current, they can last for a long time. We believe that using gold as the top surface of the electrodes helps to increase the lifetime of the actuators. Gold is stable and does not get oxidized easily like other metals that are commonly used in MEMS processes, such as aluminum and copper. Furthermore, gold is ductile and malleable which prevents the electrode from getting damaged during bending.

Appendix B. Device Fabrication on a CMOS Chip

We have demonstrated that the fabrication procedures of the actuators (shown in Figure 5.4) can be carried out on a large substrate (a quarter wafer). However, the ultimate goal is to fabricate the microelectromechanical systems (MEMS) actuators directly on top of the tiny CMOS dies and this imposes unique challenges that are not encountered in wafer level processing. Die level processing is necessary because wafer level processing is often costly and not accessible. Semiconductor facilities at research institutions are usually not capable of fabricating complicated CMOS chips that meet modern requirements. Therefore, CMOS chip fabrication typically has to be arranged through commercial foundries. As the feature sizes shrink and wafer size increase, it is usually not affordable for research labs or small companies to acquire a whole wafer or even a large of portion of a wafer; they often share space on a multi-project wafer. The price to order a whole 8" wafer on a relatively old technology (0.35 μm, 1-poly, 4-metal) is more than $10,000. After fabrication the foundry dices the wafer into small dies and delivers them to different customers. Therefore, researchers usually do not have the option of performing processing on a whole wafer or even a large portion of a wafer other than working with tiny dies if they want to integrate CMOS chips with MEMS processing.

One problem of processing tiny dies is how to handle these dies. They are not easily transportable with tweezers and are too small for standard equipment. One common way to deal with this problem is to attach the small dies to a large substrate like glass, wafer [191, 192], or PDMS [193, 194]. Another more serious problem is

that the edge bead of photoresist might cover a large percentage of the die surface that would potentially interfere with the photolithography process [193, 195, 196].

During spin coating of the photoresist, the fluid flows gradually outward to the edges by centrifugal force. The photoresist does not fly off the substrate once it reaches the edge but, instead, gathers at the edge and form a thick bump due to surface tension [197-199]. Moreover, high air flow at the edge of the substrate dries the photoresist and makes it more viscous and easier to accumulate [200]. This thick photoresist at the edges is called an edge bead. The edge bead introduces nonuniformity and reduces the useful area of the wafer or the chip. Photolithography recipes strongly depend on the thickness of the photoresist. This thick bead cannot be patterned properly with the same recipe as the thin and uniform photoresist area. If the photoresist is used as an etching mask for example, we have to make sure that the undeveloped photoresist and improperly etched material do not cause severe problems, like accidently shorting components or covering structures that are supposed to be exposed.

Another problem for the edge bead is that the exposure tool has a limited depth of focus, so we have to bring the samples in contact to the mask. The edge bead, if not removed, would stick to the mask and/or prevent the uniform area (usually the center of the sample) of the photoresist to contact the mask. This additional space between mask and the photoresist introduces scattering of the UV light and distorts the pattern as observed in our experiment. One other problem is particulate contamination due to cracking of the bead during handling [201]. In addition, the edge bead might wrap around the substrate edge and contaminate the back side. The

photoresist on the back side of the substrate will contaminate the subsequent equipment and disturb leveling during exposure [200].

What makes the edge bead problem more severe is that our process requires usage of a thick photoresist to pattern the electrode layer. In order for the actuators to have enough bending force, the SU-8 has to be thick enough as per the actuator optimization described in section 5.2.3. The bending force is approximately linearly proportional to the thickness of SU-8 given fixed electrode thickness as in Table 5.3. The heights of the individual layers of the actuators are 1.0 μm, 5.0 μm, and 0.5 μm for the sacrificial layer, SU-8, and electrode, respectively. The total height of the actuator before patterning the electrode is 6.5 μm. In order to pattern the electrode layer, the photoresist has to cover the tall structures. Coating photoresist on tall features has been discussed by other researchers. Cooper et al. reported that photoresist film tends to tear at the topography edges and proposed a closed chamber coating system [202]. Fischer and Süss suggested the use of spray coating on steep topography instead of spin coating [203] while Pham et al. also suggested that spray coating brings controllability and produce better results [204]. Previously our group found that thick photoresist has better adhesion than thin photoresist on tall features given that we do not have access to spray coating equipment. Therefore, thick photoresist AZ 9260 was selected as the etching mask for the electrode layer. However, the high viscosity of the thick photoresist makes the edge bead problem even more severe.

The formation of the edge bead is complicated and several mechanisms are involved including viscous, capillary, gravitational, centrifugal, Coriolis, and finite-

contact-angle effects [205]. There is no compact equation to predict the width and height of the edge bead; numerical simulations of the fluids have to be performed [198, 205]. In general our observation has shown that the more viscous the photoresist, the more severe the edge bead (higher and wider). We can get some insight from the thickness equation for spin coating at the center uniform area of the photoresist to understand this observation. The photoresist thickness T is proportional to the viscosity of the photoresist [206]

$$T = \frac{\alpha C^{\beta} \eta^{\gamma}}{\omega^{\kappa}}. \tag{8.1}$$

where C is polymer concentration, η is dynamic viscosity, ω is the spin speed, and α, β, γ, and κ are fitting parameters. Kinematic viscosities of Shipley 1813, SU-8 2005, SU-8 2010, and AZ 9260 (a few photoresists that we use most often) are 25, 45, 380, and 486 cSt, respectively; the number for AZ 9260 was calculated from its dynamic viscosity (520 cps) and density (1.07 g/cm^3), others are directly from the manufacturers' datasheet. It is worth noting that SU-8 2010 yields about twice the thickness of SU-8 2005 with the same spinning program but its viscosity is more than eight times of SU-8 2005.

Without the aid of automatic processing machine, fabrication procedures usually involve manual processing that introduces unreliability and nonrepeatability into the process. The purpose of Appendix B is to describe our efforts to develop a stable die-level process, rather than rely on a low-yield process due to the random nature of spin coating on tiny chips, and also to review the methods that researchers have proposed to deal with the edge bead problem.

B.1 Edge Bead Reduction Methods

Researchers have shown that they are able to properly do photolithography on tiny dies. We have divided these edge bead reduction methods into four categories: alternatives to spin coating, spin coating recipe modifications, die surface extension, and edge bead removal. These methods generally apply better to situations with the following characteristics: 1) features are away from the edges of the dies; 2) the process does not require a thick photoresist; 3) exact feature size is not critical so allowance for feature distortion is higher. In practice some of the methods can be combined to better mitigate the edge bead problem. We have experimented with most of these methods and will discuss them below.

Dummy dies used in the experiments were prepared as follows: 1) deposit Cr/Au on a 4" silicon oxide wafer; 2) coat Shipley 1813 on the wafer for protection and bake at 90 ℃ for 2 minutes; 3) attach standard dicing tape to the back of the wafer and trim the tape; 4) dice the wafer using Micro Automation Industries (MAI) Dicing Saw Model 1006; 5) detach dies from tape; 6) put dies in acetone, methanol, and IPA bath; 7) rinse the dies with DI water; 8) collect the dies on a napkin and put it on the hotplate to dry. The purpose of coating the wafer at step 2 is to protect the wafer surface from being scratched by the debris during dicing. The dicing tape at step 3 is to keep the dies intact for easier cleaning. At step 4 the saw stop needs to be carefully set to just cut through the wafer but not to cut through the tape so the tape is complete to hold the dies. Five random dies were measured using a caliper; the average (standard deviation) of length, width, and height are 3.050 mm (0.025 mm), 1.522 mm (0.011 mm), and 538 μm (16.4 μm).

The actuation controller chips have a layout 3.0 mm × 1.5 mm in size. However, the actual sizes of the chips varies depending on how the foundry decided to dice the chips for their convenience. We have sent out six versions of this design with the same pad frame for fabrication and received 95 chips back in total; the width varies from 3.02 mm to 3.70 mm and the length varies from 1.51 mm to 1.67 mm. The thickness remained relatively constant varying from 276 μm to 286 μm.

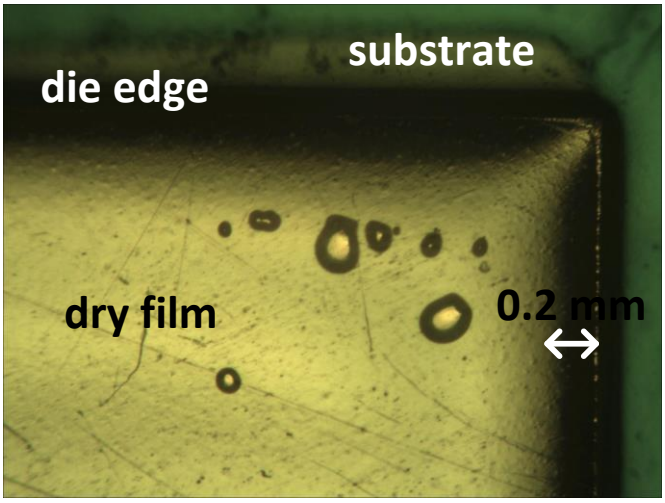
A. Alternative Coating

1) Dry Film

Dry film is a photoresist film commercially fabricated as an adhesive-backed tape, and was originally designed for printed-circuit board (PCB) fabrication [207]. Instead of spin-coating the photoresist, the dry film is laminated onto a substrate for photolithography. One particular advantage of the dry film that is appealing for this application is no formation of edge bead [207, 208]. If the film can be perfectly transferred and adhered to the substrate, edge bead formation can be avoided.

We experimented with MG Chemicals 416DFR-5 dry film photoresist. It is a negative tone and 40 μm thick photoresist. The procedures are: 1) Attach a die to a carrier wafer using photoresist Shipley 1813; 2) Cut a piece of dry film that is large enough to cover the die; 3) Gently attach the dry film to the die surface; 4) Laminate the whole carrier wafer at 110 ℃ using a normal office laminator; 5) Align mask and expose for 15 seconds at 8 mW/cm^2; 6) Develop for 2 minutes in developer (57 g potassium carbonate mixed with 1 gallon DI water); 7) Rinse with DI water and dry.

One problem of using the dry film on tiny dies is that the photoresist is not uniform after lamination as shown in Figure B.1 (a). The rollers of the laminator tend to apply more pressure on the edges so that the photoresist assumes a dome shape. Bubbles also appeared in most samples. This nonuniformity might be the cause of improper patterning of the dry film on tiny dies as shown in Figure B.1 (b). The adhesion of the photoresist should be improved observing that while the pattern still exist, some of them shifted position.



(a)

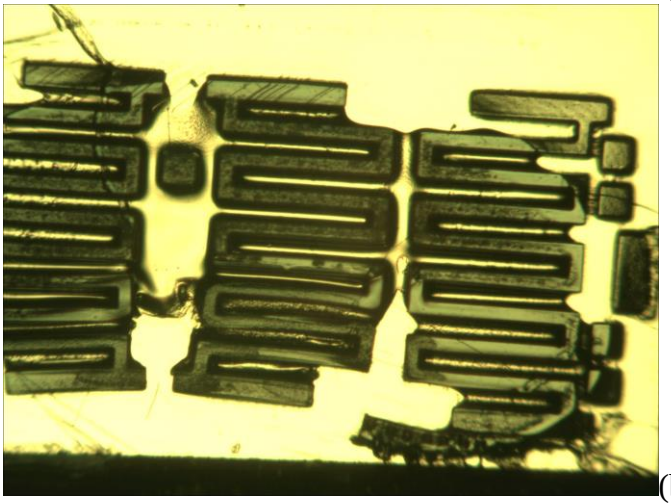

(b)

Figure B.1. (a) Dry film photoresist after lamination. The photoresist formed a dome shaped profile and contained bubbles. The width of the shaded area at the periphery is 0.2 mm. (b) Dry film photoresist after developing. Patterns were not properly developed.

2) Stamp

The idea of this method is to transfer a thin and uniform layer of photoresist on a flat substrate to a tiny chip. If the photoresist can be transferred completely, there is potentially no edge bead on the chip. First, photoresist is spin coated on a flat and wide substrate, for example a glass substrate or a wafer. Then, the tiny chip is evenly and gently pressed down onto the photoresist and slowly pulled up. However, we have found that if the photoresist is thin, there is still edge bead on the edges of the chip. We believe the formation of edge bead was due to the surface tension during the release of the chip from the coated substrate. On the other hand, if the photoresist is thick, the photoresist transferred to the chip has a dome shape.

One stamp sample was prepared as follows: 1) Spin coat Shipley 1813 on a quarter wafer at 1200 RPM for 5 seconds; 2) Apply the stamp method mentioned above on a 3 × 1.5 mm^2 chip; 3) Soft bake the chip on a hot plate at 90 ℃ for 1.5 minutes. 4) Rehydrate for 3 minutes; 5) Align mask and expose with Karl Suss MJB-3 mask aligner set at 8 mJ/cm^2 for 10 seconds; 6) Develop for 7 minutes; 7) Rinse and dry. The profile of the sample after developing is shown in Figure B.2. It demonstrates that edge bead cannot be avoided and cannot develop properly. The thickness of the edge bead is 8 μm which is even worse than simple spin coating. The width of the edge bead is around 300 μm.

Figure B.2. (Left) Blue line is the scan direction on a developed sample using stamp method. (Right) The profile of the sample showing an edge bead that cannot be developed.

In conclusion, it is difficult to develop a useful process for the stamp method. The amount of photoresist transferred to the chip is not well controlled. It depends on the thickness of the photoresist on the substrate and the poorly controlled manual technique of pressing down and pulling up the die.

3) Constant Volume Injection of Photoresist

The constant volume injection proposed by Lin et al. applies an exact volume of photoresist equal to the surface area of the die multiplied by the required photoresist height [209]. The volume of photoresist has to account for the solvent content of the photoresist because the solvent will evaporate during drying. After application of the photoresist the die is baked for several hours to allow the photoresist to reflow and self-planarize. Lin et al. used a commercial plastic 1 mL syringe with 0.01 mL resolution to control the volume of MicroChem SU-8 50 [209]. They achieved ultra-thick thicknesses of 500, 1000, and 1500 µm on dies sized 10 × 10 and 4 × 4 cm^2. Due to the ultra-thick nature of the photoresist, it forms a dome

shape with thickness at the edges being thinner than the center area. It was reported that the thickness reached the settling value within 3 mm from the edges. Pattern distortion occurred at the edges because of the incomplete contact.

This method does not apply well to small dies and non-ultra-thick photoresist. The first problem was the control precision of the injection volume: 10 μm of photoresist on a 3 × 1.5 mm^2 chip only yields a volume of 45 nL. Volume control at this resolution cannot be easily achieved. The second problem was that such a small amount of photoresist does not spread out to cover the whole die surface because of the surface tension of the photoresist; it tends to form a droplet. When the amount of photoresist is enough to cover the whole die surface, it is already too thick and forms a dome shape.

4) Coating using Special Equipment

These methods include spray coating and roller coating, but they all require special equipment and may not be widely available. As a result we have not tested the these methods.

B. Recipe Modification

1) Increase Spin Radius

Increasing spin radius increases the centrifugal force at the outer edge the sample, which could overcome the surface tension of the photoresist and reduce the edge bead. The centrifugal force *F* can be expressed as [210]

$$F = m \cdot \omega^2 \cdot r \,. \tag{8.2}$$

where m is the mass of the photoresist being considered, ω is the angular speed, and r is the distance to rotating center.

Li et al. placed dies on a carrier wafer to increase spin radius and found that this method could effectively increase the uniform area of the photoresist [196, 211]. Their experimental conditions were using Shipley 1813 spinning at 3000 RPM, dies sized 1.5 × 1.5 mm^2 and 3 × 3 mm^2, and a few different spin radii. It was reported that at the same spin radius larger dies provided better uniformity and less edge bead effect. Moreover, the edge bead was the worst on the trailing edge relative to the direction of rotation and slightly better on the interior edge rather than the exterior edge with respect to the center [211]. It was also reported that the best result was obtained from the 3 × 3 mm^2 die spinning at 34 mm radius; all patterns more than 130 μm away from the edge were properly developed, and on two edges patterns could be formed up to the edge of the die, resulting in reliable patterning of 87% of the surface area [211]. Liu et al. also adopted this method to reduce edge bead on tiny dies [192]. They used Shipley 1813 and placed a die (1.5 × 1.5 mm^2) close to the edge of a 4" wafer [192].

2) Increase Spin Speed

Increasing spin speed has a similar effect to increasing spin radius as shown in Equation 5.2. Our experiments and characterizations of the combination of increasing spin radius and speed will be discussed in more detail later.

3) Photoresist with Optimal Viscosity

The rationale for this method is to find a photoresist that is just thick enough to cover the tall structures (lowest viscosity possible) so that the edge bead is less

severe according to Equation 5.1. This can be achieved by finding a different photoresist or diluting an existing thick photoresist. As we mentioned previously, the photoresist has to cover 6.5 μm tall features in order to pattern the electrode layer. The height of AZ 9260 following our recipe is about twice the height of the features (12 μm). Therefore, we diluted the AZ 9260 photoresist with propylene glycol monomethyl ether acetate (PGMEA) solution to make it less viscous. We used MicroChem SU-8 developer as the added solution because its main ingredient (> 99.5%) is PGMEA. The mixing ratio was 2 cc photoresist to 1 cc PGMEA. The mixed photoresist was then applied to a quarter wafer followed by spinning at 2000 RPM for one minute with acceleration of 100 RPM/sec. After coating the sample was baked at 110 ℃ for 3 minutes and patterned. The photoresist thickness was reduced from 12 μm to 8.8 μm. The edge bead was not significantly improved and still unacceptable for direct photolithography.

4) Multiple Coating

MicroChemicals suggested that a multiple coating of photoresist with an elevated spin speed for each coating cycle gives better results than single-coating of thick resist films [212]. We believe the overall effect of this method is similar to dilution of the photoresist.

C. Die Surface Extension

The rationale for the methods in this category is to create an extdeded surface that is level with the die surface. If there is minimal discontinuity between the

extended surface and the die surface, the edge bead can be transferred from the die surface to the extended surface during spinning.

1) Attaching Side Pieces

This method involves placing 4 rectangular side pieces around the CMOS die as shown in Figure B.3. The side pieces can be any size as long as they can attach closely to the CMOS chip without any gap but they should be the same height as the CMOS chip to minimize discontinuity of the entire surface. The side pieces can be prepared by dicing a wafer into pieces. However, it is difficult to find wafers exactly the same height as the CMOS chips received from the foundry. Therefore, either the CMOS chips or the side pieces have to be polished to match the heights. It is preferred to thin the side pieces to avoid damaging the CMOS chips during grinding.

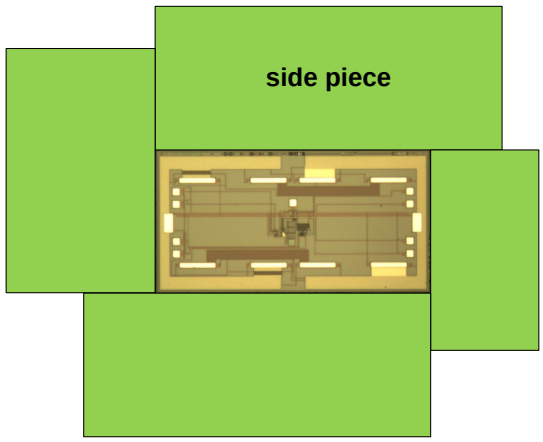


Figure B.3. Top view of the CMOS chip and four dummy chips. The center one is a real chip photo. The green ones are dummy chips.

Agarwal and Chen et al. used this method to pattern Shipley 1813 into an etching mask to define open windows on a 2×2 mm^2 CMOS chip [193, 194]. Blanco Carballo patterned MicroChem SU-8 50 on cm scale chips using this method [213]. As reported by Blanco Carballo [213], this method reduces edge beads but does not

get rid of them completely. Therefore, the mask patterns cannot be close to the edges as these examples showed.

The difficulties of using this method is explained as follows. It might seem straightforward to align multiple chips tightly together and glue them on a larger substrate. However, any discontinuity or gap between these chips could make the edge bead even worse than without using this method. From our experience we found that this method had low yield of less than 10 % due to excessive manual processing. In order to glue the chips on the wafer we have to first apply adhesive on the carrier substrate. Low-viscosity adhesive is preferred because it is easier make a thin and uniform film from this adhesive. Otherwise, nonuniformity of the adhesive would directly transfer to the interfaces between chips and even sometimes be amplified. Therefore, wax, a common adhesive used in the cleanroom, is too thick and cannot be used.

We have experimented with using Shipley 1813 and AZ 4620 as the adhesive. Different means to apply the adhesive were tested. First, we applied adhesive on the carrier wafer to cover an area that the five chips would occupy. We then spun the carrier substrate on the photoresist spinner to make a uniform layer. If the spinning is too fast or too long, the adhesive was already hardened and not sticky. On the other hand, we spun the carrier substrate at a lower speed or for a shorter period, and found problems as shown in Figure B.4. Because the adhesive was still thick, we had to try to evenly press down the chips via a glass slide during placement or they would float on the adhesive and become uneven later. However, the chips are also likely to be tilted if pressed down unevenly. One other problem was that the adhesive was

squeezed and might overflow to the top of the chip surface when we tried to align the chips together. In addition, if the adhesive was too wet, it did not hold the chip that we placed first and made alignment impossible because all chips slid when we pushed the chips toward each other. Therefore, if the adhesive was too thin or too thick, it did not work. We found that a spin program for Shipley 1813, 1500 RPM for 3 seconds, yielded an appropriate thickness. It was still sticky but not too thick. Unfortunately, the time window between the end of spinning and curing of the adhesive was less than 20 seconds which was too short for placing and aligning five chips. In addition, the chips were sometimes still tilted.

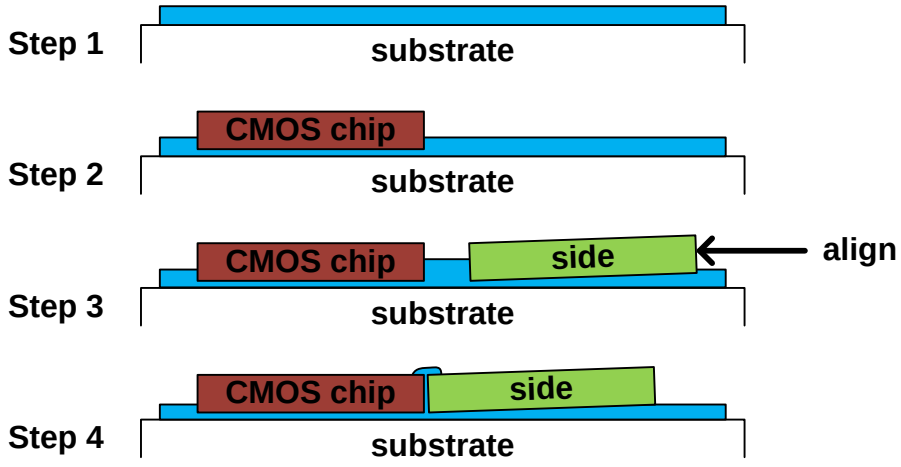


Figure B.4. Procedures (observing from the side) to attach a CMOS chip and side pieces closely on a carrier substrate. Step 1 shows that the adhesive is uniform but still thick. The chip is pressed down into the adhesive in step 2. The chip can be tilted as shown by the side piece in step 3. The adhesive might overflow to the top of the chip surface during aligning of the chips.

Since it is difficult to apply adhesive once for all five chips, the other way is to apply the right amount of adhesive once for each chip. In such case we do not have to worry about the allowed time window of the adhesive to place all chips. Before placing one chip we applied a tiny amount of adhesive on the carrier substrate or on

the back of the chip. Then the chip was evenly pressed down via a glass slide. Two problems were encountered at this step. The first problem was the adhesive overflowed out of the coverage of the chip as shown in the left of Figure B.5. This extra adhesive prevents another chip that will be placed on the left to attach closely to the chip already placed. The other problem was the tiny amount of adhesive tended to form a droplet as shown in the right of Figure B.5. This nonuniform adhesive tilted the chips placed on top of it. Both problems were rooted in the difficulty of applying the exactly tiny amount of adhesive as required and making it into a uniform film. As discussed in constant volume injection, the required volume is nL level. Moreover, such a tiny amount of liquid is always dominated by surface tension.

Figure B.5. (Left) Top view. Adhesive overflows outside the coverage of the chip. (Right) Side view. The chips are tilted because the adhesive forms a droplet due to surface tension.

2) Dry Film

The purpose of dry film discussed here is different from in the previous section. Temiz first thinned chips down to 50 μm, coated a negative dry film on a substrate, patterned the dry film using a chip as mask, and embedded the chip into the hole in the patterned dry film [183]. Dry film was also used for other photolithography steps. Square 3.5 mm dummy chips and square 4 mm CMOS chips were tested in this process [183].

3) Chip Carrier

Methods in this category are used to create a carrier with a cavity that is slightly larger than a chip so that the chip can fit in tightly and have an extensive surface for photolithography processing.

3.1) 3D Printing Carrier

3D printing technology is convenient and cheap. We designed carriers in Autodesk Inventor (a 3D CAD tool). The carriers had at least three cavities corresponding to each chip: one cavity size is the same as the size measurement of the chip, another cavity is the chip plus a few μm, and the other cavity is the chip minus a few μm. The latter two were designed to account for the fabrication variation. The printing was carried out through Terrapin Works service at the University of Maryland. A close view at a corner of a cavity of two printed parts is shown in Figure B.6. These two carriers were printed by a Stratasys Object30 Pro printer using Polyjet RGD 515 material and by a Stratasys Object500 Connex3 printer using Polyjet RGD 525 material, respectively. Both printers have an X-Y resolution of 42 μm (600 dpi) while Stratasys Object500 Connex3 has a better accuracy of 5 μm over 100 μm and a thinner layer thickness of 16 μm over 28 μm. Both carriers showed curved corners and uneven surface on the plateau areas and cavity areas. These non-ideal structures prevented the chips from fitting in tightly. Given that the Stratasys Object500 Connex3 printer is a high-end state-of-the-art printer, we conclude that current 3D printing technology is still not able to provide the accuracy required for microfabrication.

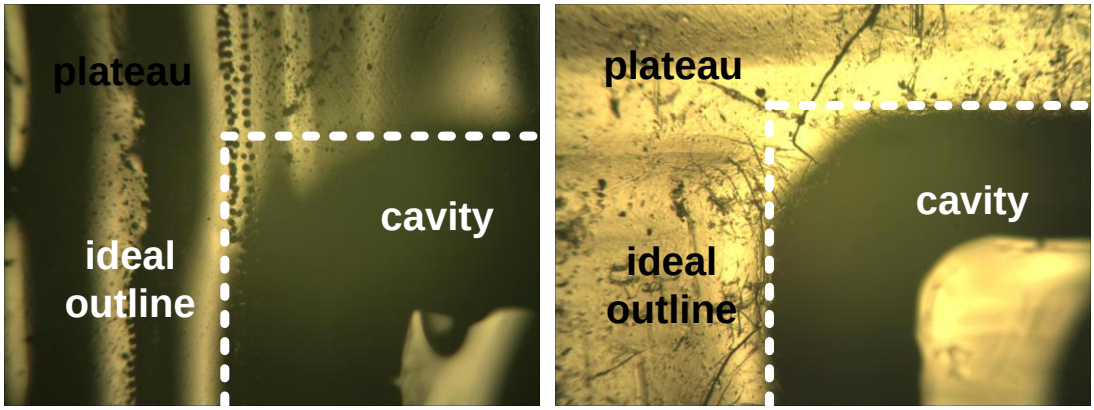


Figure B.6. (Left) First carrier printed by Stratasys Object30 Pro printer using Polyjet RGD 515 material. (Right) Second carrier printed by Stratasys Object500 Connex3 printer using Polyjet RGD 525 material. The height of both pictures is 1.5 mm. Dashed white lines represent ideal outlines of the cavity. Most of the scratches were caused when removing the supporting material from the printed parts.

3.3) Polymer Packaging

This method uses a polymer to seamlessly package the chips to create an extended surface while making sure that the package surface levels with the chip surface. Datta, Abshire, and Smela embedded CMOS chips in an epoxy handle wafer [214]. However, the epoxy used by the authors was a permanent one. As a result, the embedded chip cannot be removed from the handle wafer easily.

We therefore picked another polymer, polydimethylsiloxane (PDMS), to package the chips. PDMS is not rigid and easy to take out the chips from it. It is also a popular material for microfluidics in microfabrication. The generalized procedures are shown in Figure B.7. We first applied a uniform adhesive layer on a flat substrate. Sometimes the substrate is sticky and can serve as the adhesive. Step 2 was to attach the chip top surface to adhesive. Step 3 was to cure PDMS surrounding the chip. The last step was to detach the chip packaged with PDMS from the adhesive. The adhesive has to be chosen carefully since its stickiness affects the yield of the process

significantly. If the adhesive is not sticky enough to hold the chip, the chip will float in the uncured PDMS during pouring or while curing PDMS as shown in Figure B.8. Once PDMS cures on chip surface, it is very difficult to clean the PDMS. If the adhesive is too sticky, it is hard to detach the package and keep the chip intact with the PDMS package because PDMS does not stick to the chip. It is possible to put the chip back to the hole on PDMS but it is difficult to level their surfaces because PDMS is flexible. After a series of failures, we determined that it is the best to have a dissolvable adhesive so that detachment involves less manual operation. Details of our experiment using PDMS will be discussed later.

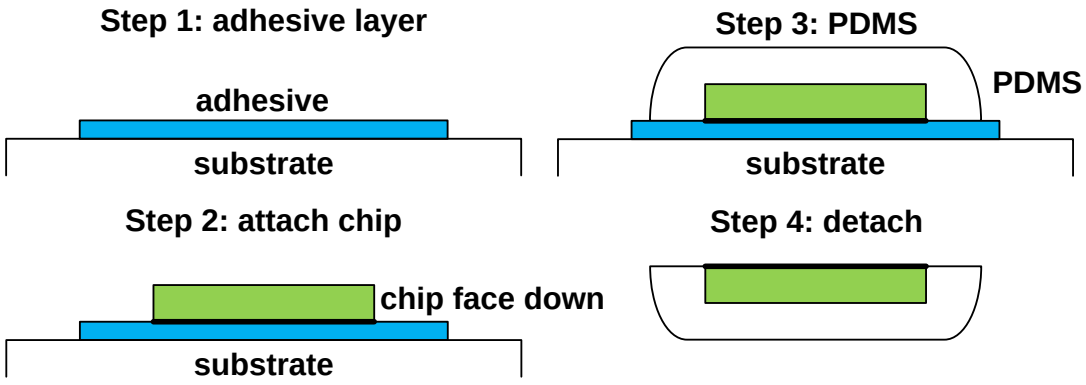


Figure B.7. Procedures to package the chip using PDMS. Detailed descriptions of the steps are in the text. A thicker line is used to represent the face (active side) of the chip.

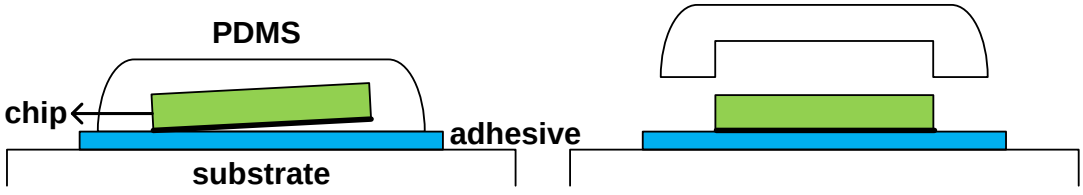


Figure B.8. (Left) Adhesive does not hold the chip well. The chip floats up in the uncured PDMS. (Right) Adhesive holds the chip tightly. PDMS detach from the chip.

3.5) Carrier Made of Other Solids

Ersen et al. designed an aluminum jig to embed a 5×7 mm^2 chip for the purpose of planarizing the chip [215]. They deposited 3 μm polyimide and 7.5 μm Shipley SP- 9019044 photoresist and then etched down to obtain a flat surface. They reported that this method did not get rid of the edge bead since the periphery within 1 mm of each edge was not flat [215]. The photoresist was not patterned but instead used as a buffer material so we hypothesize that the tolerance for edge bead is higher for their situation. They also tried to embed the chip in wax but were not successful [215]. Huang et al. made a chip holder by first dicing a silicon wafer into pieces that are larger than the CMOS chips, depositing SiO_2, patterning photoresist using the chip as mask, and etching SiO_2 and Si until the hole depth is the same as the height of the chip. The CMOS chips were sized 5 × 6.5 mm^2. The photolithography on the chip and chip holder was two layers of AZ 4620 both spun at 3000 RPM for 30 sec. The edge bead was improved from 1 mm wide and 20 μm thick without the chip holder to 0.4 mm wide and 15 μm thick with the holder [216]. In this work the photolithography pattern is regular (most shapes are circular) and inherently tolerates the distortion.

D. Edge Bead Removal

1) Manual Edge Bead Removal

The edge bead can be removed by soaking a Q-tip with solution depending on the targeted photoresist and using it to wipe the edge bead on the substrate to remove the extra photoresist. However, this manual wiping requires some space to operate. In

addition, the wiped area can no longer be patterned and will have to be sacrificed. From our experience and our colleagues' experience, with careful operation and selection of Q-tip, the manual wiping of the edge bead requires at least 1.5 mm of width from edge toward the center of the substrate to operate. To increase the reliability of process it is best to save 2 mm buffer space.

2) RIE

J.-B. Lee et al tried to planarize a 3 × 3 mm^2 chip by depositing thick polyimide/benzocyclobutene/polyimide on chip. This work included a suggestion to remove the edge bead using RIE but no details were reported [191]. We hypothesized that their RIE could only apply to a selective area.

3) Precise Edge Bead Removal Equipment

There are different types of equipment that can perform edge bead removal precisely using much smaller widths than manual operation. Many of them are designed for circular wafers used to fabricate CMOS processes in the foundries. Two main approaches for topside edge bead removal are chemical and optical edge bead removal [217, 218]. Chemical edge removal uses a nozzle to dispense edge bead remover to the edge of a wafer. Precise position control of the nozzle and/or the wafer is required. Optical edge bead removal subjects the photoresist at the edge of the wafer to a broadband exposure (also called wafer edge exposure) but this method only applies to positive photoresist. For negative photoresist a mechanical ring is used to prevent the photoresist from getting exposed. However, the edge exclusion width for wafer processing transitioned from 2 mm to 1.5 mm in 2007 as suggested by ITRS [93] and they predicted no further reduction up to year 2020. One new emerging

technology using laser cleaning achieved sub-mm edge exclusion width; UVTech System claimed that their LEC-300 Laser Edge System can operate with a width below 500 μm [219]. Although these numbers are aimed for mass production where the highest robustness and repeatability are required, it still indicates that edge bead removal is not easy even with the aid of state of the art equipment. Some equipment can handle rectangular substrates. They might be suitable for die level processing. MBRAUN Edge Bead Remover had an edge accuracy specification of < ± 100 μm but no width information was given [220].

B.2 More Effective Edge Bead Reduction Methods

We tested the methods described in the previous section that were available to us. From the preliminary investigations, we abandoned most of the methods and decided to focus on the two most promising methods: combination of increasing spin speed and radius, and PDMS packaging.

A. Increase Spin Speed and Radius

A combination of increasing spin speed and radius is simple and is available to most researchers who have access to micro-fabrication facilities. The only additional item is a larger carrier wafer which can be reused. The spinning speed cannot be increased without limit. As the spin speed reaches some level, the photoresist starts to become nonuniform. For example, AZ 9260 spun at 3000 RPM starts to sometimes have large ripples on the surface that can have a height difference larger than 10 μm (normally < 2 μm). Therefore, we only increased the spin speed up to 3000 RPM in our experiment.

In order to characterize how this method reduces the edge bead, we experimented with different spin radii and spin speeds as well as different chip sizes. The dummy chips were placed close to the edge of carrier wafers with two different sizes, 4” and 6”. We then spun the wafer for one minute, starting from stationary and accelerated by 100 RPM/sec until the targeted spin speed is reached. Normally edge removal is performed immediately after spinning so we are interested in the photoresist profile while it is still wet. Nevertheless, in order to use a profilometer to profile the edge bead, the samples were dried to harden the photoresist. This drying causes the bead to change shape as suggested by Shiratori and Kubokawa [201]. They also suggested that the shape change can be minimized by first drying the samples at room temperature before baking them. Hence we first dried the samples at room temperature for 10 minutes and then placed on a hot plate. If facilities permit, optical ways of measuring the thickness by interferometer [201] or light absorption [221] are better options because the thickness profile can be obtained immediately after spinning and the distortion introduced by drying is minimized.

The data shown in Table B.1 is the measured edge bead width and height (defined as the height difference between the peak of the edge bead and the flat center area of the photoresist) at the outer edge of the chip where the edge bead is the worst. We observed that increasing the spin radius and increasing the spin speed could both effectively reduce the edge bead width and height. The size of the chips did not significantly affect the edge bead.

Table B.1 Edge Bead (EB) Characteristics for Different Experimental Conditions

wafer (inch)	size (mm)	RPM	sample #	EB width (μm)	EB thickness (μm)
4	3	1000	1	1233	35
4	3	1000	2	1125	35
4	5	1000	1	1366	30
4	5	1000	2	1319	28
4	7	1000	1	1371	23
4	7	1000	2	1371	14
4	9	1000	1	1500	26
4	3	2000	1	829	18
4	3	2000	2	825	19
4	5	2000	1	1018	22
4	5	2000	2	803	17
4	7	2000	1	971	21
4	9	2000	1	721	11
4	3	3000	1	450	12
4	3	3000	1	714	7
4	5	3000	1	592	10
4	5	3000	2	651	9
4	7	3000	1	700	10
4	7	3000	2	480	10
4	9	3000	1	644	13
6	3	2000	1	440	11
6	5	2000	1	340	12
6	7	2000	1	644	16
6	9	2000	1	720	12

B. PDMS Packaging

PDMS packaging was described earlier in polymer packaging. We have experimented with many materials as the adhesive. The results are summarized in Table B.2. Detailed descriptions of each experiment for the adhesives were written by Ms. Deepa Sritharan and are in Appendix C. She did all of polydimethylsiloxane (PDMS) curing for packaging.

Table B.2 Summary of Results Experimenting Different Adhesive

Adhesive	Release	Results
Ecoflex 30	Peeling	PDMS glued to Ecoflex
Ecoflex 30 swab	Peeling	PDMS glued to Ecoflex
Ecoflex 30 + mold release	Peeling	PDMS glued to Ecoflex
PDMS + mold release	Peeling	PDMS went underneath
Vinyl tape	N.A.	PDMS did not cure
Adhesive spray	N.A.	Bumpy surface. PDMS went underneath
Screen protector	N.A.	PDMS went underneath
Scotch painters tape	Peeling	Chip stuck to adhesive
Scotch packaging tape	Peeling	30 μm step
Scotch packaging tape	Heating	Residue
3M masking tape	Peeling	PDMS went underneath
Scotch magic tape	peeling	Chip stuck to adhesive
Scotch magic tape	Heating	Residue
Hot glue	Peeling	Chip stuck to adhesive
2 way glue spin coated on glass	Peeling	Chip stuck to adhesive
2 way glue on paper	Peeling	Chip stuck to adhesive
Glue stick on glass	Peeling	Chip stuck to adhesive
Glue stick on paper	Peeling	Chip stuck to adhesive
Firm wax	Heating	Residue
Reusable scotch tape squares	Heating	Residue
Sugar	Dissolving	PDMS went underneath
Dashboard phone holder	Peeling	PDMS went underneath

Once the dies are packaged and the surface of the chip is level with the package. Some issues may remain. First, PDMS is hydrophobic. This property encourages some photoresist to spread out on the surface, for example Shipley 1813. However, other photoresist like AZ 9260 prefer a hydrophilic surface. Chen et al. suggested to use O_2 plasma treatment to temporarily change the PDMS surface from hydrophobic to hydrophilic to improve the adhesion of AZ 9260 [222]. The adhesion promoter PDMS, on the other hand, changes the surfaces from hydrophilic to hydrophobic or enhances their hydrophobicity, and this effect is more permanent. Therefore, researchers should proceed with caution when using HMDS on PDMS because it might not work well with some photoresist. Second, PDMS sticks to the masks. This becomes an issue when aligning the mask and the sample during contact exposure. Once they are brought into contact, PDMS sticks to the mask which prevents us fine alignment. Under such a small scale and low error tolerance, alignment usually has to be done iteratively. This sticking problem requires one step alignment without adjustment, which is an extraordinarily difficult task.

B.3 Other Anticipated Issues

We also anticipate that the non-flat surface of the CMOS chips might cause issues in the process. This nonuniformity is due to the CMOS structures and metal connections underneath [191] as well as the open windows on the top passivation which are created for access to the top metal layer. Although some works successfully performed photolithography on CMOS chips without encountering these issues [183, 193, 194, 213-216], this nonuniformity should be taken into account in the process sequences. The chips received from the foundry have open windows 10 μm in depth

and have hills 10 μm in height surrounding the open windows as shown in Figure B.9. Photoresist might get stuck in the cavity of the pad and not be developed properly. The hills create a gap between the mask and the sample distorting the features.

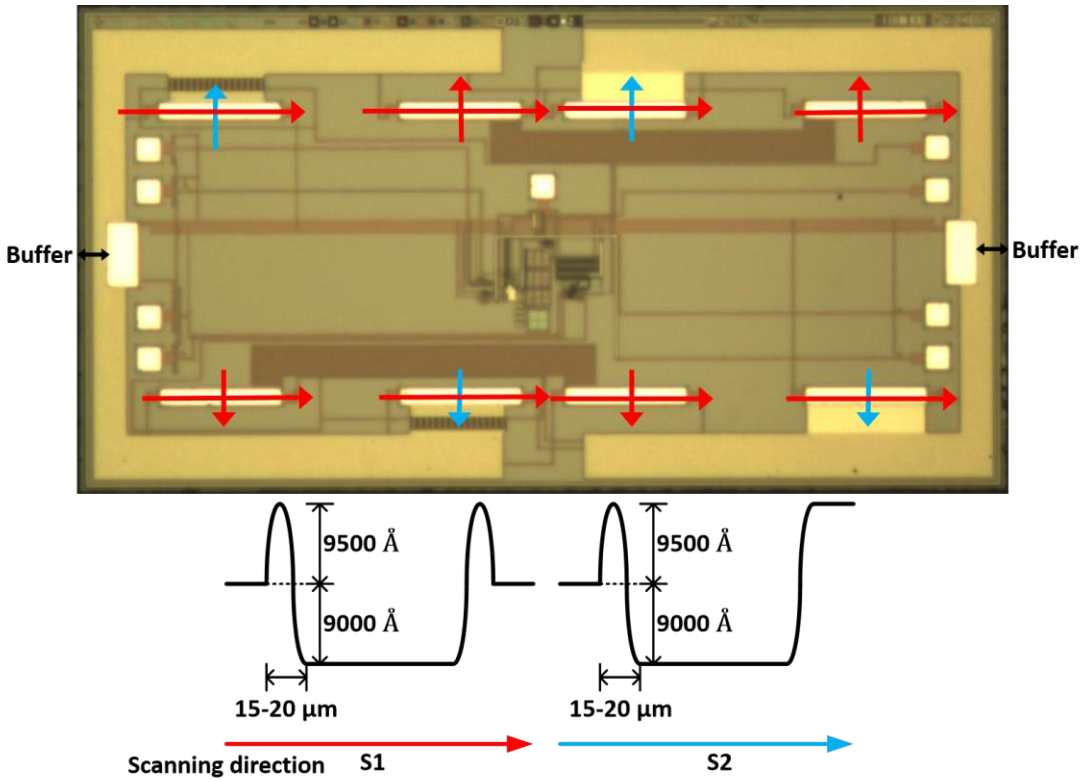


Figure B.9. Chip profile around the open pads. There are two types of profile (indicated with different colors) depending on the layers underneath. The buffer area will be discussed later.

In order to perform MEMS processing on non-flat CMOS chips, chip planarization might be required. A common idea is to deposit thick layers of materials on top to smooth the height difference and optionally etch down the materials. One example is to coat the chip surface with polyimide and photoresist followed by plasma etching [215]. A similar method uses a combination of polyimide, benzocyclobutene, and SiO_2 to smooth the height differences [191]. However, open

windows still need to exist so that electrical connections can be formed from the CMOS circuits to external devices.

B.4 CMOS Chip Design Strategy

A. Leave Buffer Area at the Periphery

An optimal scenario to deal with the edge bead problem is to first use one of the edge bead reduction methods described previously to reduce the size of the edge bead, particularly its width, so that it can be later removed by manual operation or by a specialized machine without sacrificing too much active area of the surface. Note that the area after edge bead removal cannot be used for patterning. A buffer area at the periphery of the chip is required to provide clearance for edge bead removal. Open pads cannot be placed in the buffer area (shown in Figure B.9) because the metal connections cannot be patterned properly in that area. Layout features other than open pads are not affected and can be used freely. The masks should be designed accordingly. The buffer width depends on the edge bead removal method. With careful operation and selection of Q-tip, manual removal of the edge bead would require 1.5 mm of width with limited yield. It is preferable to save greater than 2 mm of width for manual operation. Specialized equipment can reliably operate within 1.5 mm. Some equipment can even achieve below 0.5 mm as discussed earlier. However, they are designed for large circular wafers and their application on tiny dies has not been evalutaed in this work.

B. Fill in Metals to Planarize Chip Surface

The layout of the CMOS chip affects the uniformity of the top surface as reported by J.-B. Lee et al. [191]. Therefore, we suggest, without affecting the

functionality of the chip, to fill blank space with all different metal layers in the layout design so that the chip surface can be as flat as possible.

B.5 Discussion

There is no easy way to get around the edge bead problem if die level processing with of thick photoresist is required. However, edge bead effects can be mitigated if the MEMS actuators are redesigned with different materials or structures to reduce their height as well as avoid using sharp edges while maintaining the force output. A particularly effective strategy will be to eliminate the requirement for thick photoresist and use thinner photoresist. The edge bead can be improved with this much lower viscosity material (note the nonlinear relationship between the photoresist viscosity and its thickness).

Appendix C. Descriptions of Different Adhesives for PDMS Packaging*

Screen Protector

A plastic cell phone screen protector was tested. It was likely made of PET and meant for dry mounting on the cell phone screen. The hypothesis was that the electrostatic adhesive force that could attach the film onto the screen may be strong enough to hold our chip in place while forming the polymer handle. This was investigated due to the potential advantage of being a clean and dry method that can enable chip encapsulation in a single step.

The chip was placed, metal side down on the screen protector film. When the film was inverted, the chip did not fall down, implying that the electrostatic adhesion was able to counter the force of gravity on the chip. A drop of 10:1 PDMS was then applied carefully on the chip using an applicator stick. Care was taken to avoid any displacement of the chip during application. The setup (the screen protector with the encapsulated chip) was placed in an oven at 65 ^{o}C for 30 minutes to cure the PDMS. After the PDMS drop encasing the chip was solidified, more PDMS was cast on the chip and cured to create a handle that was approximately 1 inch in diameter and 3 mm thick. The cured structure was peeled off.

It was observed that PDMS had crept under the chip and completely lifted it off the screen protector film. This likely because the electro-adhesive force was weaker than the force transmitted by the viscous liquid PDMS.

* The text in this appendix was written by Ms. Deepa Sritharan. The experiments mentioned here were performed by her.

Dashboard Phone Pad

In addition to the screen protector film, we tested a sticky pad that is manufactured to secure a phone against the dashboard of a car by electroadhesion. The product is marketed for its ability to function in moving car. Our chip being much lighter, and the force of the PDMS liquid exerted on the chip being much smaller than those present in a car, we anticipated that this may be a better dry mount than the screen protector. However, we observed the same result that we faced with the screen protector. We deduced that although the electroadhesive films did hold the chip up against gravity, contact with the chip surface was likely to be incomplete near edges of the chip, where the PDMS was able to creep underneath through micro-gaps.

Smooth-On Ecoflex

Ecoflex 0030 and Ecoflex 0050 were tested as adhesive substrates to hold the chip while molding the polymer handle around it. Both thick and thin films of Ecoflex were tested. Thick films of Ecoflex (mixed 1:1 according to manufacturer's instructions) were cast to a thickness of 3 mm in a petri dish and cured in an oven at 65 ^{o}C for 20 minutes. The thick Ecoflex films were used to dry mount the chip and encapsulate it in PDMS, in a manner similar that described using the screen protector. Thin films were fabricated by spin coating a glass slide with Ecoflex at 2500 rpm for 360 sec at 500rpm/sec ramp. The chip was then wet mounted on the coated slide, with the metal coating facing the adhesive. The mounted slide was placed in an oven

at 65 ^{o}C for 20 minutes to cure the Ecoflex. The chip was encapsulated with PDMS as previously described.

In both cases, the PDMS adhered strongly to the Ecoflex layers (in both Ecoflex 0030 and Ecoflex 0050 samples); in some places the two layers could not be separated without damaging the films. When the PDMS handle was eventually released, the peeling force required for detachment caused the chip to dislodge from the PDMS.

Hot Glue

Hot glue was tested as an adhesive layer due its ability to reflow at higher temperatures and solidify upon cooling. In the first test, a thick layer of hot glue was applied to a glass slide using a hot glue gun. The slide was then placed on a hot plate set to 100 ^{o}C and melted to create a flat surface. The slide was then cooled at room temperature for 2 minutes, so that it partially solidified. The chip was gently placed on the surface of the glue. The glue was brought to room temperature and the chip was then encapsulated in PDMS.

In the second test, a small amount of hot glue was applied on a glass slide heated from beneath using a hot plate set to 100 ^{o}C. The molten hot glue was then wiped off, while the slide was still on the hot plate, leaving a thin film of glue on the glass surface. The chip was placed on the molten thin film and then allowed to cool to room temperature.

In both cases, it was observed upon release that the PDMS had crept under the chip. Upon further investigation, we observed that the cause for this was that upon

solidification the hot glue lost its adhesive grip on the chip, allowing PDMS to seep into the gap between the chip and the hot glue during casting. In addition, it is also possible that a small amount of deformation of the the hot glue layer may have occurred while baking the PDMS in the oven at 65 ^{o}C.

Zig 2 Way Acrylic Glue

This commercially available adhesive is an acrylic emulsion adhesive that cures to form a tacky surface. A glass slide was spin coated with this liquid adhesive for 2500 rpm for 60 seconds at 1000 rpm/sec ramp. The chip was mounted on the dried tacky adhesive and encapsulated with PDMS as previously described. After curing, it was observed that several bubbles had formed at the interface of the adhesive and the PDMS, and a smooth PDMS surface was not formed around the chip. This bubble formation during baking is likely due to evolution of the solvent used in the glue formulation. The PDMS handle was peeled off eaily, the chip remained adhere to the adhesive layer.

Elmer's Glue Stick

A layer of adhesive was applied using Elmers glue stick on a glass or paper substrate. The chip was stuck to to the glue and the glue was allowed to dry. The chip was encapsulated in PDMS. Bubble formation similar to that described using the acrylic glue was observed. The PDMS handle was peeled off easily, but the chip remained adhere to the adhesive layer. To check if if the adhesive could be softened

or dissolved, the set up was placed in hot water, but the adhesion was still too strong to release the chip before dislodging it from the PDMS.

Pressure-Sensitive Tapes

The following commercially available pressure sensitive tapes were tested for use as the adhesive layer: 3M Blue Painters Tape, ScotchBlue Painters Tape, Scotch Reusable Tabs, Scotch Magic Tape, and transparent Scotch Packaging Tape, listed in increasing order of bond strength. Pressure sensitive tape has an adhesive coating on a plastic or paper-based backing. The tape was attached to a glass slide with the adhesive facing up, by fixing its edges to the glass base with more tape. The chip was wet mounted on the adhesive. The chip was then encapsulated with PDMS as previously described. Upon release, it was observed that the best chip molding was achieved using the Reusable Tabs, Magic Tape, and the Packaging Tape. Some amount of PDMS did creep under the chip in these samples, but about 80% of the samples appeared to have good chip molding. Painter's tapes were observed to be unsuitable for this application. This is due to the uneven adhesive layer on the tape, which creates gaps under the chip, causing PDMS to creep underneath. The main issue with using these pressure sensitive tapes was that, their adhesion to the chip was much stronger than their adhesion to the cured PDMS handle. Thus during peeling, the chip remained attached to the tape and became dislodged from the PDMS handle despite taking care to peel the tape off horizontally to minimize the lifting force on the chip.

In order to peel the tape off more easily, we tested methods to soften the adhesive. The samples were soaked in acetone overnight to swell/dissolve the adhesive. Soaking in acetone caused the PDMS to swell significantly and deform (curving), pushing the chip out of its mold. The tape was then slowly peeled off. The adhesive had softened to a gel, but was still too strongly bonded to the chip for clean release. An ultrasonication step in acetone was then added to the overnight soak, but this did not appear to improve the adhesive dissolution. The other method we tried for softening the adhesive was to apply heat. It was observed that the tape was easier to peel when the sample was warmed to a temperature of about 65 ^{o}C. The warm tape was peeled off horizontally, or slid off horizontally. Both release method resulted in the chip remaining securely in place in PDMS handle. However a layer of tacky adhesive remained on the surface of the chip. We tried to dissolve this residue using acetone and gently wiping the surface using a swab. However, it did not appear to create an adequately clean surface. It may be possibly that a stronger solvent such as hexane may be successful for the cleaning. However, hexane rapidly swells PDMS, and it is likely that there still may be residue that is undissolved in the crevices around the edges of the chip.

Wafer-Mounting Film Wax

It was established, based on our findings using the pressure sensitive tape, that a release method that involves sliding off the encapsulated chip instead of peeling would ensure that the chip does not dislodge during the release process. A sheet of wafer-mounting wax was placed on a glass slide, and the chip was placed on the dry

film. The set up was placed on a hot plate at approximately 90 ^{o}C for 15 min to heat-activate the adhesive wax. The glass slide was then allowed to cool down to room temperature. The chip was encapsulated in PDMS as previously described. The set up was placed on the hot plate again to soften the adhesive, the PDMS handle containing the chip was slid off horizontally off the glass slide. Upon inspection, it was observed that a thick layer of wax was still attached to the chip. We attempted to dissolve this using acetone, however we were unable to create a clean chip surface. Although the film wax is designed to mount and release chips during processing, it may be assumed that the surface of the chip adhered to the wax is typically the back of the chip not the surface of interest.

Sugar

Based on our experience with previous approaches we learnt that to achieve clean chip encapsulation and release the chip must wet-mounted on an adhesive layer that is tacky but resistant to reflow for strong grip on the chip during encpasulation, and the adhesive residue must easily dissolved. So, we looked towards employing adhesive films that could dissolved using water. The first approach was to make a sugar-based based film. 10 grams of granulated white sugar was mixed with a drop of glacial acetic acid. DI Water (1 gram) was added to the sugar to form a slurry. The sugar was heated at 200 ^{o}C for 10 minutes to melt the sugar and form a smooth caramel; the acetic acid prevents recrystallization of the sugar during caramelization. The sugar was then applied to coat a glass slide maintained at temperature of 100 ^{o}C using a metal spatula and the sugar was allowed to melt to form a film on the glass.

A sheet of printer paper was affixed to the sugar and allowed to soak up the molten sugar. The chip was placed to the sticky paper. The chip was encapsulated in PDMS as before. The setup was then placed in a beaker of hot water and sonicated to dissolve the sugar; the paper aids in release by absorbing water and allowing better contact for the water with the sugar for faster dissolution. The PDMS handle with chip was slid off the glass slide, and it was observed that the chip molding appeared to secure the chip in place without PDMS covering it. However, the issue with using the sugar film was that it was not possible to create a perfectly flat film. When maintaining the sugar in a molten form, there was also bubble formation due the water content in it boiling. Thus, it was not possible to create a perfectly flat film, and the PDMS handle contained several dents due to these bubbles.

PVA

Our final attempt was to employ polyvinyl alcohol (PVA) films that can be dissolved easily in water. A commercially available pressure sensitive PVA tape was tested. The tape was one-sided and therefore a double-sided adhesive sheet was used to attach it securely to a glass slide, adhesive side facing up. This step was necessary, instead of simply securing the edges of the tape, as we did in the case of the pressure-sensitive tapes, because otherwise the PVA tape experienced severe shrinkage and warping during PDMS baking at 65 ^{o}C. No warping occurred when the double-sided adhesive backing was used to fix the PVA tape to the glass. The chip was mounted on the adhesive and encapsulated in PDMS. The setup was placed in boiling water for 30 minutes. The PVA tape entirely dissolved to form a thin gel and the PDMS

handle with chip was easily slid off the glass. Good encapsulation was observed. In addition to using PVA tape we also spin coated a PVA solution (1 gram PVA powder dissolved in 20 grams of hot water) on a glass slide at 2500 rpm for 5 seconds at 1000 rpm/sec ramp. The chip was mounted immediately on the slide; the PVA film dried within 10 seconds of coating to form a film that was dry to touch. The sample was baked in an oven at 65 ^{o}C for 1 hour to drive off moisture from the PVA film that may affect the PDMS during curing. The chip was encapsulated in PDMS. The set up was placed in boiling water for 30 minutes to dissolve the PVA film. No residue remained, the adhesive was completely removed and the PDMS handle with chip was easily slid off the slide. Good encapsulation was observed.

Appendix D. Analysis of Jitter Floor for the Testing Equipment

Most jitter measurements were done by using the histogram function of Tektronix MSO4034B that finds the time distribution of the threshold crossing points one period after triggering. However, we noticed that this method exhibits a floor for jitter measurement at about 120 ps. We found that an alternate method provided a lower jitter floor by using the automatic measurements functionality for digital channels to measure the standard deviation of the period. The digital channels have a higher sampling rate of 16.5 GHz compared to 2.5 GHz for the analog channels. This method was used for all jitter measurements below 150 ps. We discuss the resolution of this method below. Three major limiting factors include discrete sampling, sampling clock jitter, and uncertainty in threshold voltage.

The mechanism of digitization of the signal is illustrated in Figure D.1. The time axis (x axis) is shifted by one ideal cycle; the voltage axis (y axis) is shifted by the triggering threshold voltage ($V_{DD}/2$ or 2.5 V). Therefore, the ideal crossing point is at time zero and zero volts. The actual crossing time is assumed to be a random variable J which has a normal distribution, $N(0,J_{rms}^2)$. The ideal sampling time sequence t_i is defined as $\Delta t+i \cdot t_{sp}$ where i = -j, -j+1, …, 0, …, j, with j chosen so that the sampling range extends more than ten J_{rms} on either side of the origin in order to cover the entire range of simulated crossing times. Δt is the offset between sampling and the origin, and t_{sp} is the sampling period (60.6 ps as provided by the manufacturer [223]). The real sampling time sequence is subject to sampling clock jitter at each

time step: $t_i' = t_i + N(0,J_{clk}^2)$, where J_{clk} is 3 ps + 0.1 ppm × record duration (60.6 ps × 10K) according to the instrument specification [223]. Likewise, the real quantization of threshold crossing is subject to uncertainty in the threshold voltage, where the threshold voltage V_{ti} at t_i is $N(0,V_{rms}^2)$ where we assumed that V_{rms} is one-fourth of the maximum uncertainty 100 mV + 3% × nominal threshold voltage (2.5 V) according to the instrument specification [223].

In order to understand the effects of sample clock jitter and threshold uncertainty on the measured jitter, we performed Monte Carlo simulation in the stochastic variables J, t_i', and V_{ti}. The ideal jitter-free signal starts at the negative threshold and increases linearly with slew rate SR, with a zero crossing at time zero: $S_{ideal}(0)=0$. The signal incorporating the effects of circuit jitter starts at the negative threshold and increases linearly with slew rate SR, with a zero crossing at time J: $S_{measured}(J) = 0$ – however that timing is impossible to measure due to the effects of discrete sampling times, sample clock jitter, and threshold uncertainty. Each Monte Carlo run has a unique value for J, with unique values for clock jitter t_i' and threshold V_{ti} at each sampling point. The signal at sampling time t_i' $S_{measured}(t_i') = S_{measured}(J) + (t_i'-J) \cdot SR = (t_i'-J) \cdot SR$ is compared to the threshold V_{ti}, and the observed digital signal is one if $(t_i'-J) \cdot SR \geq V_{ti}$. When the first binary one occurs, the period ends and one run is completed. This process is executed repeatedly for different values of J_{rms} and Δt, and the simulation results are shown in Table D.1. We find that this method for measuring jitter has a jitter floor 45 ps under the set up mentioned above. Δt does not affect the jitter measurement much. When the incoming signal has a jitter of 20

and 30 ps, the resulting measured jitter is 47 and 52 ps, respectively. This matches our empirical observations (see Figure 7.8).

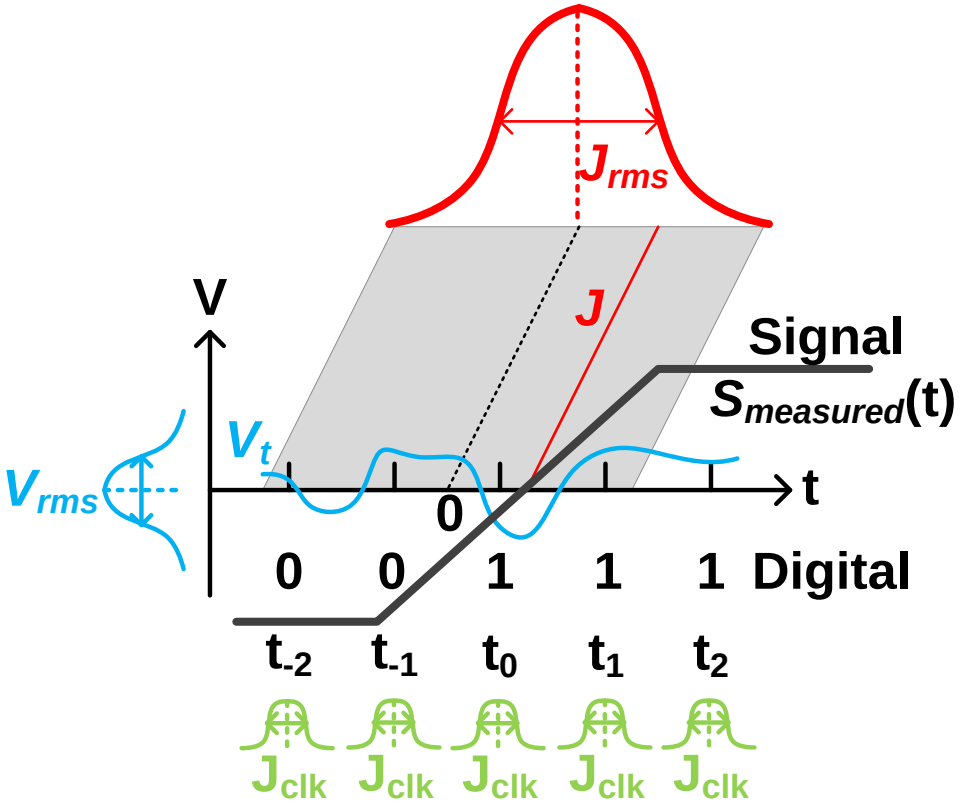


Figure D.1. Diagram for digitization of signal. The observed digital signal should be 00011 if there is no uncertainty of threshold voltage.

Table D.1 Monte Carlo Simulation Results (ps)

J_{rms} (ps) \ Δt (UI)	0	0.2	0.4	0.6	0.8
1	43	42	45	42	44
5	44	44	44	42	44
10	43	46	46	45	44
20	48	48	46	47	47
30	54	52	53	53	50
60	73	74	73	76	73
90	103	101	99	100	100
120	127	128	127	127	127
150	154	155	157	156	156
180	185	186	186	184	185

Bibliography

[1] I. F. o. R. (IFR). [Online]. Available: http://www.ifr.org/.

[2] BostonDynamics. [Online]. Available: http://www.bostondynamics.com/.

[3] E. Nebot, "Surface Mining: Main Research Issues for Autonomous Operations," in *Robotics Research*. vol. 28, ed: Springer Berlin Heidelberg, 2007, pp. 268-280.

[4] WIKIPEDIA. "*Robot*," [Online]. Available: http://en.wikipedia.org/wiki/Robot. [Accessed: Aug. 22, 2013]

[5] J. Pepitone. "*Amazon buys army of robots*," [Online]. Available: http://money.cnn.com/2012/03/20/technology/amazon-kiva-robots/index.htm. [Accessed: Aug. 21, 2013]

[6] WIKIPEDIA. "*da Vinci Surgical System*," [Online]. Available: http://en.wikipedia.org/wiki/Da_Vinci_Surgical_System. [Accessed: Aug. 22, 2013]

[7] L. Fotoohi and A. Graser, "Building a safe care-providing robot," in *Proc. IEEE International Conference on Rehabilitation Robotics (ICORR)*, 2011, pp. 1-6.

[8] M. Rubenstein, C. Ahler, and R. Nagpal, "Kilobot: A low cost scalable robot system for collective behaviors," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2012, pp. 3293-3298.

[9] A. T. Baisch, C. Heimlich, M. Karpelson, and R. J. Wood, "HAMR3: An autonomous 1.7g ambulatory robot," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011, pp. 5073-5079.

[10] A. M. Hoover, E. Steltz, and R. S. Fearing, "RoACH: An autonomous 2.4g crawling hexapod robot," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2008, pp. 26-33.

[11] G. Caprari, P. Balmer, R. Piguet, and R. Siegwart, "The autonomous micro robot "Alice": a platform for scientific and commercial applications," in *Proc. International Symposium on Micromechatronics and Human Science (MHS)*, 1998, pp. 231-235.

[12] G. Caprari, K. O. Arras, and R. Siegwart, "The autonomous miniature robot Alice: from prototypes to applications," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2000, pp. 793-798 vol. 1.

[13] G. Caprari, T. Estier, and R. Siegwart, "Fascination of down scaling-Alice the sugar cube robot," *Journal of micromechatronics*, vol. 1, no. 3, pp. 177-189, 2001.

[14] G. Caprari and R. Siegwart, "Mobile micro-robots ready to use: Alice," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2005, pp. 3295-3300.

[15] J. Brufau, M. Puig-Vidal, J. Lopez-Sanchez, J. Samitier, N. Snis, U. Simu, *et al.*, "MICRON: Small autonomous robot for cell manipulation applications,"

in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2005, pp. 844-849.

[16] R. Casanova, A. Saiz, J. Lacort, J. Brufau, A. Arbat, A. Dieguez, *et al.*, "Towards co-operative autonomous 1cm^3 robots for micro and nanomanipulation applications: MICRON," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2005, pp. 789-794.

[17] R. Casanova, A. Saiz-Vela, A. Arbat, J. Colomer, P. Miribel, A. Dieguez, *et al.*, "Integrated Electronics for a 1cm^3 Robot for Micro and Nanomanipulation Applications: MiCRoN," in *Proc. IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2006, pp. 13-18.

[18] W. A. Churaman, A. P. Gerratt, and S. Bergbreiter, "First leaps toward jumping microrobots," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011, pp. 1680-1686.

[19] W. A. Churaman, L. J. Currano, C. J. Morris, J. E. Rajkowski, and S. Bergbreiter, "The first launch of an autonomous thrust-driven microrobot using nanoporous energetic silicon," *IEEE Journal of Microelectromechanical Systems*, vol. 21, no. 1, pp. 198-205, 2012.

[20] H. Woern, M. Szymanski, and J. Seyfried, "The I-SWARM project," in *Proc. IEEE International Symposium on Robot and Human Interactive Communication (ROMAN)*, 2006, pp. 492-496.

[21] R. Casanova, A. Dieguez, A. Arbat, O. Alonso, A. Sanuy, J. Canals, *et al.*, "Integration of the control electronics for a mm^3-sized autonomous microrobot into a single chip," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2009, pp. 3007-3012.

[22] R. Casanova, A. Dieguez, A. Sanuy, A. Arbat, O. Alonso, J. Canals, *et al.*, "Enabling swarm behavior in mm^3-sized robots with specific designed integrated electronics," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2007, pp. 3797-3802.

[23] R. Casanova, A. Arbat, O. Alonso, A. Sanuy, J. Canals, and A. Dieguez, "An optically programmable SoC for an autonomous mobile mm^3-sized microrobot," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 11, pp. 2673-2685, 2011.

[24] T. Ebefors, J. U. Mattsson, E. Kalvesten, and G. Stemme, "A robust micro conveyer realized by arrayed polyimide joint actuators," in *Proc. IEEE International Conference on Micro Electro Mechanical Systems (MEMS)*, 1999, pp. 576-581.

[25] T. Ebefors, J. U. Mattsson, E. Kalvesten, and G. Stemme, "A robust micro conveyer realized by arrayed polyimide joint actuators," *Journal of Micromechanics and Microengineering*, vol. 10, no. 3, p. 337, 2000.

[26] T. Ebefors, J. U. Mattsson, E. Kälvesten, and G. Stemme, "A walking silicon micro-robot," in *Proc. International Conference on Solid-State Sensors and Actuators (Transducers)*, 1999, pp. 1202-1205.

[27] S. Hollar, A. Flynn, C. Bellew, and K. S. J. Pister, "Solar powered 10 mg silicon robot," in *Proc. IEEE International Conference on Micro Electro Mechanical Systems (MEMS)*, 2003, pp. 706-711.

[28] E. Erik, S. Niklas, M. Raimon Casanova, S. Oliver, C. Paolo, G. Jianbo, *et al.*, "Evaluation of building technology for mass producible millimetre-sized robots using flexible printed circuit boards," *Journal of Micromechanics and Microengineering*, vol. 19, no. 7, p. 075011, 2009.

[29] O. Alonso, J. Canals, L. Freixas, J. Samitier, A. Dieguez, M. Vatteroni, *et al.*, "Enabling multiple robotic functions in an endoscopic capsule for the entire gastrointestinal tract exploration," in *Proc. ESSCIRC*, 2010, pp. 386-389.

[30] T. Fukuda, H. Ishihara, and F. Arai, "Microrobotics, current of art and future," in *Proc. INRIA/IEEE Symposium on Emerging Technologies and Factory Automation (ETFA)*, 1995, pp. 29-39 vol. 3.

[31] M. Hirose and T. Takenaka, "Development of the humanoid robot ASIMO," *Honda R&D Technical Review*, vol. 13, no. 1, pp. 1-6, 2001.

[32] K. Kaneko, K. Harada, F. Kanehiro, G. Miyamori, and K. Akachi, "Humanoid robot HRP-3," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2008, pp. 2471-2478.

[33] M. Rubenstein, A. Cornejo, and R. Nagpal, "Programmable self-assembly in a thousand-robot swarm," *science*, vol. 345, no. 6198, pp. 795-799, 2014.

[34] S. H. University. "*Micron Public Final Report*," [Online]. Available: https://vision.eng.shu.ac.uk/mmvl/viewfinder/micron.html. [Accessed: Aug. 29, 2013]

[35] C. Perkins, L. Lei, M. Kuhlman, T.-H. Lee, G. Gateau, S. Bergbreiter, *et al.*, "Distance sensing for mini-robots: RSSI vs. TDOA," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2011, pp. 1984-1987.

[36] OpenCores. [Online]. Available: http://opencores.org/.

[37] M. A. P. Pertijs, K. A. A. Makinwa, and J. H. Huijsing, "A CMOS smart temperature sensor with a 3σ inaccuracy of ±0.1°C from -55°C to 125°C," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2805-2815, 2005.

[38] K. Souri, C. Youngcheol, and K. A. A. Makinwa, "A CMOS Temperature Sensor With a Voltage-Calibrated Inaccuracy of ±0.15°C (3σ) From 55°C to 125°C," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 1, pp. 292-301, 2013.

[39] A. Berkovich, M. Lecca, L. Gasparini, P. A. Abshire, and M. Gottardi, "A 30 µW 30 fps 110 × 110 Pixels Vision Sensor Embedding Local Binary Patterns," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 9, pp. 2138-2148, 2015.

[40] Y. L. Wong and P. A. Abshire, "A 144 × 144 Current-Mode Image Sensor With Self-Adapting Mismatch Reduction," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 8, pp. 1687-1697, 2007.

[41] K. Stangel, S. Kolnsberg, D. Hammerschmidt, B. J. Hosticka, H. K. Trieu, and W. Mokwa, "A programmable intraocular CMOS pressure sensor system implant," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 7, pp. 1094-1100, 2001.

[42] A. V. Chavan and K. D. Wise, "A monolithic fully-integrated vacuum-sealed CMOS pressure sensor," *IEEE Transactions on Electron Devices*, vol. 49, no. 1, pp. 164-169, 2002.

[43] J. W. Gardner, P. K. Guha, F. Udrea, and J. A. Covington, "CMOS Interfacing for Integrated Gas Sensors: A Review," *IEEE Sensors Journal*, vol. 10, no. 12, pp. 1833-1848, 2010.
[44] M. Y. Afridi, J. S. Suehle, M. E. Zaghloul, D. W. Berning, A. R. Hefner, R. E. Cavicchi, *et al.*, "A monolithic CMOS microhotplate-based gas sensor system," *IEEE Sensors Journal*, vol. 2, no. 6, pp. 644-655, 2002.
[45] F. Horiguchi, "Integration of Series-Connected On-Chip Solar Battery in a Triple-Well CMOS LSI," *IEEE Transactions on Electron Devices*, vol. 59, no. 6, pp. 1580-1584, 2012.
[46] W. Liu, Y. Wang, W. Liu, Y. Ma, Y. Xie, and H. Yang, "On-chip hybrid power supply system for wireless sensor nodes," in *Proc. Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2011, pp. 43-48.
[47] H. Ballan and M. Declercq, *High voltage devices and circuits in standard CMOS technologies*: Springer, 1999.
[48] M. J. Kuhlman, T.-H. Lee, and P. A. Abshire, "Mixed-signal odometry for mobile robotics," in *Proc. SPIE8725, Micro- and Nanotechnology Sensors, Systems, and Applications V*, 2013.
[49] T.-H. Lee and P. A. Abshire, "Design methodology for a low-frequency current-starved voltage-controlled oscillator with a frequency divider," in *Proc. IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2012, pp. 646-649.
[50] T.-H. Lee and P. A. Abshire, "An ultra-low frequency ring oscillator with programmable tracking using a phase-locked loop," in *Proc. IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2012, pp. 17-20.
[51] T.-H. Lee and P. A. Abshire, "Design and characterization of high-voltage NMOS structures in a 0.5 μm standard CMOS process," *IEEE Sensors Journal*, vol. 13, no. 8, pp. 2906-2913, 2013.
[52] T.-H. Lee and P. A. Abshire, "40 volt NMOS in a 0.5 μm standard CMOS process," in *Proc. IEEE Sensors Conference*, 2012, pp. 1-4.
[53] T.-H. Lee and P. A. Abshire, "Accurate, wide range analog sine shaping circuits," in *Proc. IEEE International New Circuits and Systems Conference (NEWCAS)*, 2014, pp. 285-288.
[54] T.-H. Lee and P. A. Abshire, "Frequency-boost jitter reduction for voltage-controlled ring oscillators," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2016.
[55] M. A. Lewis, R. Etienne-Cummings, A. H. Cohen, and M. Hartmann, "Toward biomorphic control using custom aVLSI CPG chips," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2000, pp. 494-500 vol.1.
[56] K. Nakada, T. Asai, and Y. Amemiya, "Analog CMOS implementation of a CNN-based locomotion controller with floating-gate devices," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 6, pp. 1095-1103, 2005.
[57] S. Still, K. Hepp, and R. J. Douglas, "Neuromorphic walking gait control," *IEEE Transactions on Neural Networks*, vol. 17, no. 2, pp. 496-508, 2006.

[58] J.-M. Breguet, S. Johansson, W. Driesen, and U. Simu, "A review on actuation principles for few cubic millimeter sized mobile micro-robots," in *Proc. International Conference on New Actuators (Actuator)*, 2006, pp. 374-381.

[59] M. Gad-el-Hak, *The MEMS Handbook*, 2 ed.: CRC Taylor & Francis, Boca Raton, 2005.

[60] B. Kim, J. Ryu, Y. Jeong, Y. Tak, B. Kim, and J.-O. Park, "A ciliary based 8-legged walking micro robot using cast IPMC actuators," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2003, pp. 2940-2945 vol.3.

[61] C. Hwang, S. Bibyk, M. Ismail, and B. Lohiser, "A very low frequency, micropower, low voltage CMOS oscillator for noncardiac pacemakers," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 42, no. 11, pp. 962-966, 1995.

[62] A. S. Elwakil and S. Ozoguz, "A low frequency oscillator structure," in *Proc. European Conference on Circuit Theory and Design (ECCTD)*, 2009, pp. 588-590.

[63] G. Smiley, "Ultra-Low-Frequency, Three-Phase Oscillator," *Proceedings of the IRE*, vol. 42, no. 4, pp. 677-680, 1954.

[64] K.-H. Cheng, J.-C. Liu, and H.-Y. Huang, "A 0.6-V 800-MHz All-Digital Phase-Locked Loop With a Digital Supply Regulator," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 59, no. 12, pp. 888-892, 2012.

[65] H. Q. Liu, W. L. Goh, L. Siek, W. M. Lim, and Y. P. Zhang, "A Low-Noise Multi-GHz CMOS Multiloop Ring Oscillator With Coarse and Fine Frequency Tuning," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 17, no. 4, pp. 571-577, 2009.

[66] M.-T. Hsieh, J. Welch, and G. Sobelman, "PLL performance comparison with application to spread spectrum clock generator design," *Analog Integrated Circuits and Signal Processing*, vol. 63, no. 2, pp. 197-216, 2010.

[67] T. Miyazaki, M. Hashimoto, and H. Onodera, "A performance comparison of PLLs for clock generation using ring oscillator VCO and LC oscillator in a digital CMOS process," in *Proc. Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2004, pp. 545-546.

[68] M. M. Mansour, M. M. Mansour, and A. Mehrotra, "Analysis of MOS cross-coupled LC-tank oscillators using short-channel device equations," in *Proc. Asia and South Pacific Design Automation Conference (ASP-DAC)* 2004, pp. 181-185.

[69] M. M. Mansour, A. Mehrotra, W. W. Walker, and A. Narayan, "Analysis techniques for obtaining the steady-state solution of MOS LC oscillators," in *Proc. International Symposium on Circuits and Systems (ISCAS)*, 2004, pp. V-512-V-515 Vol.5.

[70] G. Szczepkowski and R. Farrell, "350 mV, 0.5 mW, 5 GHz, 130 nm CMOS class-C VCO design using open loop analysis," in *Proc. IET Irish Signals and Systems Conference (ISSC)*, 2012, pp. 1-6.

[71] B. Leung, "A switching-based phase noise model for CMOS ring oscillators based on multiple thresholds crossing," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 11, pp. 2858-2869, 2010.

[72] A. A. Abidi, "Phase noise and jitter in CMOS ring oscillators," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 8, pp. 1803-1816, 2006.

[73] R. J. Baker, *CMOS: Circuit Design, Layout, and Simulation*. Hoboken, NJ: Wiley-IEEE Press, 2008.

[74] J. J. Chen, S. I. Liu, and Y. S. Hwang, "Low-voltage single power supply four-quadrant multiplier using floating-gate MOSFETs," *IEE Proceedings - Circuits, Devices and Systems*, vol. 145, no. 1, pp. 40-43, 1998.

[75] J. J. Chen, S. I. Liu, and Y. S. Hwang, "Low-voltage single power supply four-quadrant multiplier using floating-gate MOSFETs," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 1997, pp. 237-240 vol.1.

[76] D. Kahng and S. M. Sze, "A Floating gate and its application to memory devices," *Bell System Technical Journal*, vol. 46, no. 6, pp. 1288-1295, 1967.

[77] K. Kanda, N. Shibata, T. Hisada, K. Isobe, M. Sato, Y. Shimizu, *et al.*, "A 19 nm 112.8 mm^2 64 Gb multi-level flash memory with 400 Mbit/sec/pin 1.8 V toggle mode interface," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 1, pp. 159-167, 2013.

[78] T.-S. Jung, Y.-J. Choi, K.-D. Suh, B.-H. Suh, J.-K. Kim, Y.-H. Lim, *et al.*, "A 3.3 V 128 Mb multi-level NAND flash memory for mass storage applications," in *Proc. IEEE International Solid-State Circuits Conference (ISSCC)*, 1996, pp. 32-33.

[79] K. Rahimi, C. Diorio, C. Hernandez, and M. D. Brockhausen, "A simulation model for floating-gate MOS synapse transistors," in *Proc. IEEE International Symposium on Circuits and Systems*, 2002, pp. II-532-II-535 vol.2.

[80] P. Hasler and J. Dugger, "Correlation learning rule in floating-gate pFET synapses," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 48, no. 1, pp. 65-73, 2001.

[81] R. R. Harrison, J. A. Bragg, P. Hasler, B. A. Minch, and S. P. DeWeerth, "A CMOS programmable analog memory-cell array using floating-gate circuits," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 48, no. 1, pp. 4-11, 2001.

[82] C. Bassin, H. Ballan, and M. Declercq, "High-voltage devices for 0.5-µm standard CMOS technology," *IEEE Electron Device Letters*, vol. 21, no. 1, pp. 40-42, 2000.

[83] A. Mason, A. V. Chavan, and K. D. Wise, "A mixed-voltage sensor readout circuit with on-chip calibration and built-in self-test," *IEEE Sensors Journal*, vol. 7, no. 9, pp. 1225-1232, 2007.

[84] X. Li, L. Gu, Y. Wang, and H. Yang, "Single-wafer-processed self-testable high-g accelerometers with both sensing and actuating elements integrated on trench-sidewall," *IEEE Sensors Journal*, vol. 8, no. 12, pp. 1992-1999, 2008.

[85] W. Wang and J. Fang, "Variable Focusing Microlens Chip for Potential Sensing Applications," *IEEE Sensors Journal*, vol. 7, no. 1, pp. 11-17, 2007.

[86] S. Tudisco, F. Musumeci, L. Lanzano, A. Scordino, S. Privitera, A. Campisi, *et al.*, "A New Generation of SPAD—Single-Photon Avalanche Diodes," *IEEE Sensors Journal*, vol. 8, no. 7, pp. 1324-1329, 2008.

[87] D. Palubiak, M. M. El-Desouki, O. Marinov, M. J. Deen, and F. Qiyin, "High-Speed, Single-Photon Avalanche-Photodiode Imager for Biomedical Applications," *IEEE Sensors Journal*, vol. 11, no. 10, pp. 2401-2412, 2011.

[88] V. Savuskan, I. Brouk, M. Javitt, and Y. Nemirovsky, "An Estimation of Single Photon Avalanche Diode (SPAD) Photon Detection Efficiency (PDE) Nonuniformity," *IEEE Sensors Journal*, vol. 13, no. 5, pp. 1637-1640, 2013.

[89] B. S. Kang, S. Kim, F. Ren, B. P. Gila, C. R. Abernathy, and S. J. Pearton, "AlGaN/GaN-based diodes and gateless HEMTs for gas and chemical sensing," *IEEE Sensors Journal*, vol. 5, no. 4, pp. 677-680, 2005.

[90] B. Thompson and H.-S. Yoon, "A Prestress Measurement Circuit for Piezoceramic Stack Transducers," *IEEE Sensors Journal*, vol. 11, no. 10, pp. 2349-2355, 2011.

[91] P. Ruther, J. Bartholomeyczik, A. Buhmann, A. Trautmann, K. Steffen, and O. Paul, "Microelectromechanical HF resonators fabricated using a novel SOI-based low-temperature process," *IEEE Sensors Journal*, vol. 5, no. 5, pp. 1112-1119, 2005.

[92] M. Dandin, A. Akturk, B. Nouri, N. Goldsman, and P. Abshire, "Characterization of single-photon avalanche diodes in a 0.5 μm standard CMOS process—part 1: perimeter breakdown suppression," *IEEE Sensors Journal*, vol. 10, no. 11, pp. 1682-1690, 2010.

[93] ITRS. "*International Technology Roadmap for Semiconductors*," [Online]. Available: http://www.itrs.net.

[94] F. Chen, K. Wang, Y. Fang, N. Allec, G. Belev, S. O. Kasap, *et al.*, "Direct-conversion X-ray detector using lateral amorphous selenium structure," *IEEE Sensors Journal*, vol. 11, no. 2, pp. 505-509, 2011.

[95] C.-Y. Chen, C.-J. Lin, and Y.-C. King, "A new sensing scheme for sensitivity enhancement of low-temperature polycrystalline silicon photodetecors," *IEEE Sensors Journal*, vol. 11, no. 6, pp. 1478-1483, 2011.

[96] M. Declercq, F. Clement, M. Schubert, A. Harb, and M. Dutoit, "Design and optimization of high-voltage CMOS devices compatible with a standard 5 V CMOS technology," in *Proc. IEEE Custom Integrated Circuits Conference*, 1993, pp. 24.6.1-24.6.4.

[97] K. M. Buck, H. Li, S. Subramanian, H. Hess, and M. Mojarradi, "Development and testing of high-voltage devices fabricated in standard CMOS and SOI technologies," in *Proc. Annual NASA Symposium on VLSI Design*, 2003, pp. 157-163.

[98] P. M. Santos, A. P. Casimiro, M. Lanca, and M. I. C. Simas, "High-voltage solutions in standard CMOS," in *Proc. IEEE Annual Power Electronics Specialists Conference (PESC)*, 2001, pp. 371-377 vol. 1.

[99] P. M. Santos, M. I. C. Simas, M. Lanca, S. Finco, and F. H. Behrens, "Breakdown voltage improvement of standard MOS technologies targeted at smart power," in *Proc. IEEE IAS Industry Applications Conference*, 1995, pp. 937-945 vol.2.

[100] A. Nezar and C. A. T. Salama, "Optimization of the breakdown voltage in LDMOS transistors using internal field rings," in *Proc. International Symposium on Power Semiconductor Devices and ICs (ISPSD)*, 1991, pp. 149-153.
[101] P. M. Santos, A. P. Casimiro, M. Lanca, and M. I. C. Simas, "CMOS compatible HV gate-shifted LDD-NMOS," *IEEE Transactions on Electron Devices*, vol. 48, no. 5, pp. 1013-1015, 2001.
[102] S. Finco, P. Tavares, A. C. Fiore De Mattos, and M. I. C. Simas, "Power integrated circuit drives based on HV NMOS," in *Proc. IEEE Power Electronics Specialists Conference (PESC)*, 2002, pp. 1737-1740.
[103] X. Ouyang, "Design and Characterization Results of Integrating CMOS-Compatible High-Voltage MOSFETs in a 0.5μm CMOS Process," M.S. thesis, Dept. Elect. Eng., Washington State Univ., Pullman, WA, USA, 1998.
[104] F. Conti and M. Conti, "Surface breakdown in silicon planar diodes equipped with field plate," *Solid-State Electronics*, vol. 15, no. 1, pp. 93-105, Jan. 1972.
[105] Z. Parpia, C. A. T. Salama, and R. A. Hadaway, "Modeling and characterization of CMOS-compatible high-voltage device structures," *IEEE Transactions on Electron Devices*, vol. 34, no. 11, pp. 2335-2343, 1987.
[106] A. Bazigos, F. Krummenacher, J. M. Sallese, M. Bucher, E. Seebacher, W. Posch, *et al.*, "A physics-based analytical compact model for the drift region of the HV-MOSFET," *IEEE Transactions on Electron Devices*, vol. 58, no. 6, pp. 1710-1721, 2011.
[107] S. A. Parke, J. E. Moon, H. J. C. Wann, P. K. Ko, and H. Chenming, "Design for suppression of gate-induced drain leakage in LDD MOSFETs using a quasi-two-dimensional analytical model," *IEEE Transactions on Electron Devices (TED)*, vol. 39, no. 7, pp. 1694-1703, 1992.
[108] K.-F. You and C.-Y. Wu, "A new quasi-2-D model for hot-carrier band-to-band tunneling current," *IEEE Transactions on Electron Devices*, vol. 46, no. 6, pp. 1174-1179, 1999.
[109] I. Paprotny and S. Bergbreiter, *Small-Scale Robotics From Nano-to-Millimeter-Sized Robotic Systems and Applications*: Springer Berlin Heidelberg, 2014.
[110] E. T. Enikov and K. Lazarov, "PCB-integrated metallic thermal micro-actuators," *Sensors and Actuators A: Physical*, vol. 105, no. 1, pp. 76-82, 2003.
[111] E. Y. Erdem, Y.-M. Chen, M. Mohebbi, J. W. Suh, G. T. A. Kovacs, B. B. Darling, *et al.*, "Thermally Actuated Omnidirectional Walking Microrobot," *Journal of Microelectromechanical Systems*, vol. 19, no. 3, pp. 433-442, 2010.
[112] J. W. Suh, S. F. Glander, R. B. Darling, C. W. Storment, and G. T. A. Kovacs, "Organic thermal and electrostatic ciliary microactuator array for object manipulation," *Sensors and Actuators A: Physical*, vol. 58, no. 1, pp. 51-60, 1997.
[113] J. W. Suh, R. B. Darling, K. F. Böhringer, B. R. Donald, H. Baltes, and G. T. A. Kovacs, "CMOS integrated ciliary actuator array as a general-purpose micromanipulation tool for small objects," *Journal of Microelectromechanical Systems*, vol. 8, no. 4, pp. 483-496, 1999.

[114] M. J. Sinclair, "A high force low area MEMS thermal actuator," in *Proc. Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITHERM)*, 2000, pp. 127-132.
[115] Y. Lai, M. James, K. Marek, and H. Ted, "Force, deflection and power measurements of toggled microthermal actuators," *Journal of Micromechanics and Microengineering*, vol. 14, no. 1, pp. 49-56, 2004.
[116] M. Ataka, A. Omodaka, N. Takeshima, and H. Fujita, "Fabrication and operation of polyimide bimorph actuators for a ciliary motion system," *Journal of Microelectromechanical Systems*, vol. 2, no. 4, pp. 146-150, 1993.
[117] C. Edwards. (2008, July 22) Temperatur control. *IET Engineering and Technology Magazine* [Online]. Available: http://eandt.theiet.org/magazine/2008/13/temp-control0813.cfm
[118] B. Balakrisnan, A. Nacev, and E. Smela, "Design of bending multi-layer electroactive polymer actuators," *Smart Materials and Structures*, vol. 24, no. 4, p. 045032, 2015.
[119] S. Alben, B. Balakrisnan, and E. Smela, "Edge Effects Determine the Direction of Bilayer Bending," *Nano Letters*, vol. 11, no. 6, pp. 2280-2285, 2011/06/08 2011.
[120] W. Kristof, G. Pieter, and P. Robert, "Comparison of methods for the mechanical characterization of polymers for MEMS applications," *Journal of Micromechanics and Microengineering*, vol. 21, no. 11, p. 115027, 2011.
[121] M. Hopcroft, T. Kramer, G. Kim, K. Takashima, Y. Higo, D. Moore, *et al.*, "Micromechanical testing of SU-8 cantilevers," *Fatigue & Fracture of Engineering Materials & Structures*, vol. 28, no. 8, pp. 735-742, 2005.
[122] C. J. Robin, A. Vishnoi, and K. N. Jonnalagadda, "Mechanical Behavior and Anisotropy of Spin-Coated SU-8 Thin Films for MEMS," *Journal of Microelectromechanical Systems*, vol. 23, no. 1, pp. 168-180, 2014.
[123] C. Luo, T. W. Schneider, R. C. White, J. Currie, and M. Paranjape, "A simple deflection-testing method to determine Poisson's ratio for MEMS applications," *Journal of Micromechanics and Microengineering*, vol. 13, no. 1, p. 129, 2003.
[124] E. Philofsky, "Intermetallic formation in gold-aluminum systems," *Solid-State Electronics*, vol. 13, no. 10, pp. 1391-1394, 1970.
[125] L. R. Barber, K. M. Ghantasala, R. Divan, D. K. Vora, C. E. Harvey, and C. D. Mancini, "Optimisation of SU-8 processing parameters for deep X-ray lithography," *Microsystem Technologies*, vol. 11, no. 4, pp. 303-310, 2005.
[126] A. d. Campo and C. Greiner, "SU-8: a photoresist for high-aspect-ratio and 3D submicron lithography," *Journal of Micromechanics and Microengineering*, vol. 17, no. 6, pp. R81-R95, 2007.
[127] WIKIPEDIA. "*Artificial intelligence*," [Online]. Available: http://en.wikipedia.org/wiki/Artificial_intelligence.
[128] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789-814, June 2000.

[129] S. Holkar and L. Waghmare, "An overview of model predictive control," *International Journal of Control and Automation,* vol. 3, no. 4, pp. 47-63, 2010.
[130] J. Richalet, "Industrial applications of model based predictive control," *Automatica,* vol. 29, no. 5, pp. 1251-1274, 9 1993.
[131] M. J. Kuhlman, E. Arvelo, L. Shuoxin, P. A. Abshire, and N. C. Martins, "Mixed-signal architecture of randomized receding horizon control for miniature robotics," in *Proc. IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2012, pp. 570-573.
[132] A. Alessio and A. Bemporad, "A survey on explicit model predictive control," in *Nonlinear model predictive control*, ed: Springer, 2009, pp. 345-369.
[133] M. N. Zeilinger, C. N. Jones, and M. Morari, "Real-time suboptimal model predictive control using a combination of explicit MPC and online optimization," *IEEE Transactions on Automatic Control,* vol. 56, no. 7, pp. 1524-1534, 2011.
[134] H. G. Tanner and J. L. Piovesan, "Randomized Receding Horizon Navigation," *IEEE Transactions on Automatic Control,* vol. 55, no. 11, pp. 2640-2644, 2010.
[135] P. O. M. Scokaert, D. Q. Mayne, and J. B. Rawlings, "Suboptimal model predictive control (feasibility implies stability)," *IEEE Transactions on Automatic Control,* vol. 44, no. 3, pp. 648-654, 1999.
[136] B. Gilbert, "Circuits for the precise synthesis of the sine function," *Electronics Letters*, vol. 13, no. 17, pp. 506-508, 1977.
[137] R. Fried and C. C. Enz, "MOST implementation of Gilbert sin(x) shaper," *Electronics Letters*, vol. 32, no. 22, pp. 2073-2074, 1996.
[138] M. Valle and F. Diotalevi, "An analog CMOS four quadrant current-mode Multiplier for low power artificial neural networks implementation," in *Proc. European Conference on Circuit Theory and Design (ECCTD)*, 2001, pp. 325-328.
[139] B. Gilbert, "A monolithic microsystem for analog synthesis of trigonometric functions and their inverses," *IEEE Journal of Solid-State Circuits,* vol. 17, no. 6, pp. 1179-1191, 1982.
[140] T.-H. Lee, J.-S. Jhuang, and T.-D. Chiueh, "A high-speed baseband receiver for MIMO OFDM based WLAN," in *Proc. International Symposium on VLSI Design, Automation and Test (VLSI-DAT)*, 2006, pp. 1-4.
[141] J. W. Fattaruso and R. G. Meyer, "Triangle-to-sine wave conversion with MOS technology," *IEEE Journal of Solid-State Circuits,* vol. 20, no. 2, pp. 623-631, 1985.
[142] C.-Y. Yang, J.-H. Weng, and H.-Y. Chang, "A 5-GHz Direct Digital Frequency Synthesizer Using an Analog-Sine-Mapping Technique in 0.35-μm SiGe BiCMOS," *IEEE Journal of Solid-State Circuits,* vol. 46, no. 9, pp. 2064-2072, 2011.
[143] D. N. Loizos, P. P. Sotiriadis, and G. Cauwenberghs, "High-speed adaptive RF phased array," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2008, pp. 1084-1087.

[144] D. Rairigh, X. Liu, C. Yang, and A. J. Mason, "Sinusoid signal generator for on-chip impedance spectroscopy," in *Proc. IEEE International Symposium on Circuits and Systems*, 2009, pp. 1961-1964.

[145] M. Bohrn, L. Fujcik, and R. Vrba, "Novel on-chip sine wave generator," in *Proc. International Conference on Telecommunications and Signal Processing (TSP)*, 2011, pp. 505-508.

[146] J. Mulder, A. C. van Der Woerd, W. A. Serdijn, and A. H. M. Van Roermund, "Translinear sin(x)-circuit in MOS technology using the back gate," in *Proc. European Solid-State Circuits Conference (ESSCIRC)*, 1995, pp. 82-85.

[147] R. G. Meyer, W. M. C. Sansen, and S. Peeters, "The differential pair as a triangle-sine wave converter," *IEEE Journal of Solid-State Circuits*, vol. 11, no. 3, pp. 418-420, 1976.

[148] A. Leuciuc and Y. Zhang, "A highly linear low-voltage MOS transconductor," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2002, pp. 735-738 vol. 3.

[149] J. Gomez-Quinones, H. Moncada-Hernandez, O. Rossetto, R. Martinez-Duarte, B. H. Lapizco-Encinas, M. Madou, *et al.*, "An application specific multi-channel stimulator for electrokinetically-driven microfluidic devices," in *Proc. IEEE International New Circuits and Systems Conference (NEWCAS)*, 2011, pp. 350-353.

[150] K. Pengwon and E. Leelarasmee, "A Quadrature Generator Based on CMOS Triangular-to-Sine/Cosine Converter with 1/4 Frequency Output," in *Proc. IEEE International Conference on Circuits and Systems for Communications (ICCSC)*, 2008, pp. 319-322.

[151] J. McNeill, R. Croughwell, L. DeVito, and A. Gasinov, "A 150 mW, 155 MHz phase locked loop with low jitter VCO," in *Proc. IEEE International Symposium on Circuits and Systems*, 1994, pp. 49-52 vol.3.

[152] K. Lasanen and J. Kostamovaara, "A 1.2-V CMOS RC Oscillator for Capacitive and Resistive Sensor Applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 12, pp. 2792-2800, 2008.

[153] A. Sai, T. Yamaji, and T. Itakura, "A low-jitter clock generator based on ring oscillator with 1/f noise reduction technique for next-generation mobile wireless terminals," in *Proc. IEEE Asian Solid-State Circuits Conference (ASSCC)*, 2008, pp. 425-428.

[154] M. Lont, D. Milosevic, A. H. M. van Roermund, and G. Dolmans, "Requirement driven low-power LC and ring oscillator design," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2011, pp. 1129-1132.

[155] S. Williams, H. Thompson, M. Hufford, and E. Naviasky, "An improved CMOS ring oscillator PLL with less than 4ps RMS accumulated jitter," in *Proc. IEEE Custom Integrated Circuits Conference*, 2004, pp. 151-154.

[156] A. Arakali, S. Gondi, and P. K. Hanumolu, "Low-Power Supply-Regulation Techniques for Ring Oscillators in Phase-Locked Loops Using a Split-Tuned Architecture," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 8, pp. 2169-2181, 2009.

[157] J. Borremans, J. Ryckaert, C. Desset, M. Kuijk, P. Wambacq, and J. Craninckx, "A Low-Complexity, Low-Phase-Noise, Low-Voltage Phase-Aligned Ring Oscillator in 90 nm Digital CMOS," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 7, pp. 1942-1949, 2009.

[158] W. S. T. Yan and H. C. Luong, "A 900-MHz CMOS low-phase-noise voltage-controlled ring oscillator," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 48, no. 2, pp. 216-221, 2001.

[159] L. Dai and R. Harjani, "Design of low-phase-noise CMOS ring oscillators," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 49, no. 5, pp. 328-338, 2002.

[160] T. Kawamoto, M. Suzuki, and T. Noto, "1.9-ps Jitter, 10.0-dBm-EMI Reduction Spread-Spectrum Clock Generator With Autocalibration VCO Technique for Serial-ATA Application," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, no. 5, pp. 1118-1126, 2014.

[161] T. Pialis, E. W. Hu, and P. Khoman, "A 1.8V low-jitter CMOS ring oscillator with supply regulation," in *Proc. IEEE International Symposium on Circuits and Systems*, 2005, pp. 2791-2794 Vol. 3.

[162] S.-Y. Kao and S.-I. Liu, "A Digitally-Calibrated Phase-Locked Loop With Supply Sensitivity Suppression," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 19, no. 4, pp. 592-602, 2011.

[163] M. Behbahani and G. E. R. Cowan, "Phase-noise tuneable ring voltage-controlled oscillator in 90 nm CMOS," in *Proc. IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2013, pp. 1031-1034.

[164] Ü. Güler and G. Dündar, "Modeling CMOS Ring Oscillator Performance as a Randomness Source," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 3, pp. 712-724, 2014.

[165] C. Liu and J. A. McNeill, "Jitter in oscillators with 1/f noise sources," in *Proc. International Symposium on Circuits and Systems (ISCAS)*, 2004, pp. I-773-6 Vol.1.

[166] A. Hajimiri, S. Limotyrakis, and T. H. Lee, "Jitter and phase noise in ring oscillators," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 6, pp. 790-804, 1999.

[167] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits*: Prentice Hall Press, 2008.

[168] D. Tasca, M. Zanuso, S. Levantino, C. Samori, and A. L. Lacaita, "Low-Power Divider Retiming in a 3–4 GHz Fractional-N PLL," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 58, no. 4, pp. 200-204, 2011.

[169] MOSIS. "*Wafer Electrical Test Data and SPICE Model Parameters*," [Online]. Available: www.mosis.com/requests/test-data. [Accessed: Jun. 9, 2015]

[170] G. D. J. Smit, A. J. Scholten, R. M. T. Pijper, L. F. Tiemeijer, R. van der Toorn, and D. B. M. Klaassen, "RF-Noise Modeling in Advanced CMOS Technologies," *IEEE Transactions on Electron Devices*, vol. 61, no. 2, pp. 245-254, 2014.

[171] A. A. Abidi, "High-frequency noise measurements on FET's with small dimensions," *IEEE Transactions on Electron Devices,* vol. 33, no. 11, pp. 1801-1805, 1986.

[172] A. Antonopoulos, M. Bucher, K. Papathanasiou, N. Mavredakis, N. Makris, R. K. Sharma, *et al.*, "CMOS Small-Signal and Thermal Noise Modeling at High Frequencies," *IEEE Transactions on Electron Devices,* vol. 60, no. 11, pp. 3726-3733, 2013.

[173] S. L. J. Gierkink and E. van Tuij, "A coupled sawtooth oscillator combining low jitter with high control linearity," *IEEE Journal of Solid-State Circuits,* vol. 37, no. 6, pp. 702-710, 2002.

[174] Y. Tokunaga, S. Sakiyama, A. Matsumoto, and S. Dosho, "An On-Chip CMOS Relaxation Oscillator With Voltage Averaging Feedback," *IEEE Journal of Solid-State Circuits,* vol. 45, no. 6, pp. 1150-1158, 2010.

[175] P. F. J. Geraedts, E. van Tuijl, E. A. M. Klumperink, G. J. M. Wienk, and B. Nauta, "A 90μW 12MHz Relaxation Oscillator with a -162dB FOM," in *Proc. IEEE International Solid-State Circuits Conference,* 2008, pp. 348-618.

[176] S. Höppner, H. Eisenreich, S. Henker, D. Walter, G. Ellguth, and R. Schüffny, "A Compact Clock Generator for Heterogeneous GALS MPSoCs in 65-nm CMOS Technology," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 21, no. 3, pp. 566-570, 2013.

[177] C.-Y. Yu, J.-Y. Yu, and C.-Y. Lee, "A Low Voltage All-Digital On-Chip Oscillator Using Relative Reference Modeling," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 20, no. 9, pp. 1615-1620, 2012.

[178] Y.-G. Chen, H.-W. Tsao, and C.-S. Hwang, "A Fast-Locking All-Digital Deskew Buffer With Duty-Cycle Correction," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 21, no. 2, pp. 270-280, 2013.

[179] F. Sebastiano, L. J. Breems, K. A. A. Makinwa, S. Drago, D. M. W. Leenaerts, and B. Nauta, "A Low-Voltage Mobility-Based Frequency Reference for Crystal-Less ULP Radios," *IEEE Journal of Solid-State Circuits,* vol. 44, no. 7, pp. 2002-2009, 2009.

[180] V. Re, M. Manghisoni, L. Ratti, V. Speziali, and G. Traversi, "Survey of noise performances and scaling effects in deep submicrometer CMOS devices from different foundries," *IEEE Transactions on Nuclear Science,* vol. 52, no. 6, pp. 2733-2740, 2005.

[181] A. J. Scholten, L. F. Tiemeijer, R. van Langevelde, R. J. Havens, A. T. A. Zegers-van Duijnhoven, and V. C. Venezia, "Noise modeling for RF CMOS circuit simulation," *IEEE Transactions on Electron Devices,* vol. 50, no. 3, pp. 618-632, 2003.

[182] S. Herbert, S. Garg, and D. Marculescu, "Exploiting Process Variability in Voltage/Frequency Control," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 20, no. 8, pp. 1392-1404, 2012.

[183] Y. Temiz, C. Guiducci, and Y. Leblebici, "Post-CMOS Processing and 3-D Integration Based on Dry-Film Lithography," *IEEE Transactions on Components, Packaging and Manufacturing Technology,* vol. 3, no. 9, pp. 1458-1466, 2013.

[184] T. Le, K. Mayaram, and T. Fiez, "Efficient far-field radio frequency energy harvesting for passively powered sensor networks," *IEEE Journal of Solid-State Circuits,* vol. 43, no. 5, pp. 1287-1302, 2008.
[185] D. J. Yeager, J. Holleman, R. Prasad, J. R. Smith, and B. P. Otis, "NeuralWISP: A Wirelessly Powered Neural Interface With 1-m Range," *IEEE Transactions on Biomedical Circuits and Systems,* vol. 3, no. 6, pp. 379-387, 2009.
[186] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannopoulos, P. Fisher, and M. Soljačić, "Wireless power transfer via strongly coupled magnetic resonances," *science,* vol. 317, no. 5834, pp. 83-86, 2007.
[187] A. Karalis, J. D. Joannopoulos, and M. Soljačić, "Efficient wireless non-radiative mid-range energy transfer," *Annals of Physics,* vol. 323, no. 1, pp. 34-48, 2008.
[188] S. I. Raider, R. Flitsch, and M. J. Palmer, "Oxide Growth on Etched Silicon in Air at Room Temperature," *Journal of The Electrochemical Society,* vol. 122, no. 3, pp. 413-418, 1975.
[189] G. Mende, J. Finster, D. Flamm, and D. Schulze, "Oxidation of etched silicon in air at room temperature; Measurements with ultrasoft X-ray photoelectron spectroscopy (ESCA) and neutron activation analysis," *Surface Science,* vol. 128, no. 1, pp. 169-175, 1983.
[190] K. R. Williams, K. Gupta, and M. Wasilik, "Etch rates for micromachining processing-Part II," *Journal of Microelectromechanical Systems,* vol. 12, no. 6, pp. 761-778, 2003.
[191] J. B. Lee, J. English, C. H. Ahn, and M. G. Allen, "Planarization techniques for vertically integrated metallic MEMS on silicon foundry circuits," *Journal of Micromechanics and Microengineering,* vol. 7, no. 2, p. 44, 1997.
[192] Y. Liu, "Fabrication and Characterization of Polypyrrole/Gold Bilayer Microactuators for Bio-Mems Applications," Ph.D. dissertation, Department of Mechanical Engineering, University of Maryland, College Park, Maryland, USA, 2005.
[193] V. Agarwal, "Integrated CMOS microsystems for carbon nanotubes based sensing and integrated micro-array for multiplexed analyte detection using electro-chemiluminescence," M.S. thesis, Department of Electrical Engineering, Tufts University, 2008.
[194] C.-L. Chen, V. Agarwal, S. Sonkusale, and M. R. Dokmeci, "Integration of Single-Walled Carbon Nanotubes on to CMOS Circuitry with Parylene-C Encapsulation," in *Proc. IEEE Conference on Nanotechnology*, 2008, pp. 480-483.
[195] P. Abshire, A. Bermak, R. Bemer, G. Cauwenberghs, S. Chen, J. B. Christen, *et al.,* "Confession session: Learning from others mistakes," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2011, pp. 1149-1162.
[196] L. Li and A. J. Mason, "Post-CMOS parylene packaging for on-chip biosensor arrays," in *Proc. IEEE Sensors*, 2010, pp. 1613-1616.
[197] T. Dobroth and L. Erwin, "Causes of edge beads in cast films," *Polymer Engineering & Science,* vol. 26, no. 7, pp. 462-467, 1986.

[198] D. E. Weidner, L. W. Schwartz, and R. R. Eley, "Role of Surface Tension Gradients in Correcting Coating Defects in Corners," *Journal of Colloid and Interface Science*, vol. 179, no. 1, pp. 66-75, 1996.
[199] C. Mack, *Fundamental Principles of Optical Lithography: The Science of Microfabrication*. London, UK: John Wiley & Sons, 2007.
[200] Y. Wei and R. L. Brainard, *Advanced Processes for 193-nm Immersion Lithography*. Bellingham, Washington: SPIE, 2009.
[201] S. Shiratori and T. Kubokawa, "Double-peaked edge-bead in drying film of solvent-resin mixtures," *Physics of Fluids*, vol. 27, no. 10, p. 102105, 2015.
[202] K. A. Cooper, C. Hamel, B. Whitney, K. Weilermann, K. J. Kramer, Y. Zhao, *et al.*, "Conformal Photoresist Coating for High Aspect Ratio Features," in *Proc. International Wafer-Level Packaging Conference*, 2007.
[203] K. Fischer and R. Süss, "Spray coating - a solution for resist film deposition across severe topography," in *Proc. IEEE/CPMT/SEMI International Electronics Manufacturing Technology Symposium*, 2004, pp. 338-341.
[204] N. P. Pham, E. Boellaard, J. N. Burghartz, and P. M. Sarro, "Photoresist coating methods for the integration of novel 3-D RF microstructures," *Journal of Microelectromechanical Systems*, vol. 13, no. 3, pp. 491-499, 2004.
[205] L. W. Schwartz and R. V. Roy, "Theoretical and numerical results for spin coating of viscous liquids," *Physics of Fluids*, vol. 16, no. 3, pp. 569-584, 2004.
[206] N. Arjmandi, *Resist Homogeneity, Updates in Advanced Lithography*: InTech, 2013.
[207] P. Vulto, N. Glade, L. Altomare, J. Bablet, L. D. Tin, G. Medoro, *et al.*, "Microfluidic channel fabrication in dry film resist for production and prototyping of hybrid chips," *Lab on a Chip*, vol. 5, no. 2, pp. 158-162, 2005.
[208] E. Kukharenka, M. M. Farooqui, L. Grigore, M. Kraft, and N. Hollinshead, "Electroplating moulds using dry film thick negative photoresist," *Journal of Micromechanics and Microengineering*, vol. 13, no. 4, p. S67, 2003.
[209] C.-H. Lin, G.-B. Lee, B.-W. Chang, and G.-L. Chang, "A new fabrication process for ultra-thick microfluidic microstructures utilizing SU-8 photoresist," *Journal of Micromechanics and Microengineering*, vol. 12, no. 5, p. 590, 2002.
[210] L. D. Landau and L. M. Lifshitz, *Mechanics*. Oxford: Butterworth-Heinemann, 1976.
[211] L. Li, X. Liu, and A. J. Mason, "Die-level photolithography and etchless parylene packaging processes for on-CMOS electrochemical biosensors," in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2012, pp. 2401-2404.
[212] MicroChemicals. "*Lithography Trouble-Shooter*," [Online]. Available: http://www.microchemicals.com/technical_information/lithography_trouble_shooting.pdf. [Accessed: January 5, 2016]
[213] V. M. Blanco Carballo, "Radiation Imaging Detectors Made by Wafer Post-processing of CMOS Chips," Ph.D. dissertation, Electrical Engineering, Mathematics and Computer Science, University of Twente, 2009.

[214] T. Datta-Chaudhuri, P. Abshire, and E. Smela, "Packaging commercial CMOS chips for lab on a chip integration," *Lab on a Chip*, vol. 14, no. 10, pp. 1753-1766, 2014.
[215] A. Ersen, I. Schnitzer, E. Yablonovitch, and T. Gmitter, "Direct bonding of GaAs films on silicon circuits by epitaxial liftoff," *Solid-State Electronics*, vol. 36, no. 12, pp. 1731-1739, 1993.
[216] Y. Huang, Y. Zhang, Z. Yin, G. Cui, H. C. Liu, L. Bian, *et al.*, "Indium bump array fabrication on small CMOS circuit for flip-chip bonding," *Journal of Semiconductors*, vol. 32, no. 11, p. 115014, 2011.
[217] I. Jekauc, M. Watt, T. Hornsmith, and J. Tiffany, "Necessity of chemical edge bead removal in modern day lithographic processing," in *Proc. SPIE 5376*, 2004, pp. 1255-1263.
[218] A. Carlson, T. Le, A. Pai, J. Hallen, and B. Rioux, "Use of automated EBR metrology inspection to optimize the edge bead process," in *Proc. SPIE 6518*, 2007, pp. 2L-8.
[219] UVTech Systems. "*Edge Film Removal: Data Sheet*," [Online]. Available: http://uvtechsystems.com/151013_service_data_sheet_v1_4.pdf. [Accessed: Jan. 27, 2016]
[220] MBRAUN. "*MB-EBR*," [Online]. Available: http://mbraun.de/products/coating-equipment/edge-bead-remover/#specifications. [Accessed: Jan. 27, 2016]
[221] A. A. Mouza, N. A. Vlachos, S. V. Paras, and A. J. Karabelas, "Measurement of liquid film thickness using a laser light absorption method," *Experiments in Fluids*, vol. 28, no. 4, pp. 355-359, 2000.
[222] W. Chen, R. H. W. Lam, and J. Fu, "Photolithographic surface micromachining of polydimethylsiloxane (PDMS)," *Lab on a Chip*, vol. 12, no. 2, pp. 391-395, 2012.
[223] Tektronix. [Online]. Available: www.tektronix.com. [Accessed: Jun. 15, 2015]