ABSTRACT

Title of Dissertation: THERMAL CHARACTERIZATION and FEEDBACK DRIVER CIRCUIT DESIGN of INTEGRATED 2D NANOPHOTONIC PHASED ARRAYS for VR and LIDAR APPLICATIONS

Po-Chun Huang, Doctor of Philosophy, 2021

Dissertation directed by: Professor Martin Peckerar
Department of Electrical and Computer Engineering

Nanophotonic phased array (NPA) technology has been an active topic of research for many years. This is due to its widespread applicability in the emerging fields of virtual reality (VR) and light detection and ranging (LIDAR). This thesis describes an integrated NPA system consisting of an optical phased array and its electrical driver circuit. These two components are realized in two separate “chips”: an optical chip (OC) for generating and routing light; and, a “motherboard chip” (MB) for creating the current drivers for the individual elements of the array. The driver circuit sources voltage or current to modulate the phase or amplitude (or both) of the optical output beams of each unit in the phased array. The output beams interfere with each other either constructively or destructively in such a way as to replicate the light field of a portrayed scene (forming a VR image) or to form a steerable beam (LIDAR). This

dissertation centers on the realization of a scalable thermo-optic based NPA system. The thermo-optic based system changes the phase of the output beam emerging from an individual emitter by locally heating the optical path through which the beam emerges from the array. To realize a large NPA system with pixel-level independent phase control, the size of a driver circuit unit must be matched to the size of the individual phased array element, or pixel. This must be accomplished, while at the same time avoiding interconnection congestion issues. This poses a critical design challenge for the driver circuit limiting its functionality. Also, a large amount of heat is generated in the thermo-optic system. Unintentional spreading of this heat through the array (known as proximity effect) not only introduces phase errors across the array, but it also causes reliability issues in the densely integrated electronic elements of the array. To overcome these issues, my thesis was divided into the following tasks.

First, I conducted a comprehensive simulation based thermal study of our proposed integrated NPA system using the COMSOL finite element method (FEM) solver. The study includes detailed single pixel simulations characterizing the thermo-electrical properties of the system. This helped guide the driver circuit design. It further enabled small array simulation for quantifying thermal spread blurring (proximity effect) and phase errors. The thesis includes transient simulations to show the response speed of the system. I show that our system requires less than 50 μs to reach a target temperature. I introduce a model simplification method to reduce the computation resource requirement of system-level simulations. These calculations tell us how large the array can be without incurring thermal damage.

This thesis further discloses a unique broken-loop feedback control system to achieve pixel-level temperature regulation for phase error minimization. The control system uses an integrated thin-film thermocouple/heater device to sense the

temperature feedback signal and to source a current for providing the required phase shift. This device requires but a single contact post between each phase array pixel and its control circuit. In this way, the OC and the MB chips can be integrated by most available flip-chip bonding technologies. Two design implementations of the driver circuit sourcing 4.8 mW per pixel from a 2.5 V supply voltage are provided. One design can be realized in an area of 15 µm x 15 µm per pixel with pixel-level independent phase control using the TSMC 65 nm technology node. This exactly matches the size of the NPA pixel. The other design can meet the same area constraint using a more advanced technology node.

This thesis also provides an experimental characterization of the driver circuit designed and fabricated on the TSMC foundry's 65nm product line. Experimental results of characterizing each component of the driver circuit are provided. The broken-loop feedback control method was electrically evaluated independent of the optical system by using a resistor to generate a simulated feedback signal. The circuit achieves a maximum 3.6% (0.07π) and average 1% (0.02 π) introduced by a ±20% variation of the load resistance. I provide a comparison of the performance of both VR image quality and LIDAR steering accuracy using either the direct control method or our broken loop feedback control method. This was done using the structure similarity index (SSIM) method. This method ranks image quality in a range from 0 to 1 (0 the poorest image and 1 the best image.) On average, the images studied improved their SSIM index from 0.45 to 0.9 using the broken-loop method. In beam steering, our feedback control method achieves less than 0.05° angle deviation and constant main beam intensity as compared to a 0.9° angle deviation and more than 90% reduction in main beam intensity using direct control. This demonstrates that our feed-back controlled driver circuit is essential for NPA systems to achieve high performance.

THERMAL CHARACTERIZATION and FEEDBACK DRIVER CIRCUIT DESIGN of INTEGRATED 2D NANOPHOTONIC PHASED ARRAYS in VR and LIDAR APPLICATIONS

by

Po-Chun Huang

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
2021

Advisory Committee:

Professor Martin Peckerar, Chair
Professor Mario Dagenais
Professor Robert Newcomb
Professor Pamela Abshire
Professor Amitabh Varshney

© Copyright by
Po-Chun Huang
2021

Dedication

To my parents, I-Pin Huang and Shu-Yuan Hsu

and

to my wife, Tzu Pei Huang,

for their love, endless support and encouragement.

I love you with all my heart

Acknowledgements

This dissertation would not be possible without the significant support and contributions from my advisor, colleagues, and family. I am the most fortunate person, having worked with a great team of supervisors, mentors and colleagues.

First and foremost, I would like to express my deepest gratitude and appreciation to my advisor and committee chair, Dr. Martin Peckerar, for his great support, invaluable guidance, and patience throughout my Ph.D. journey. His enthusiasm and encouragement motivated me to overcome all the challenges and his extensive knowledge and experiences directed me to the completion of this work.

I would like to thank Dr. Mario Dagenais and Dr. Amitabh Varshney who provided me great support and invaluable feedback. I would have never been able to finish this thesis without their help. I would like to acknowledge Dr. Feng Yi and Dr. Pamela Abshire for their generously help sharing equipment for my research and experiments. I would like to thank my colleagues at University of Maryland, Dr. Chao Wang, Dr. Gaunnan Liu, Dr. Xuetong Sun and Dr. Yang Zhang, for their assistance and beneficial discussion.

Finally, I want to thank my parents for their unconditional love and endless support. I would like to thank my wife, Tzu Pei Huang, for her exceptional support, love, and immeasurable sacrifices through the entire journey. I also like to thank my sons, Anders, Nathan, and Nolan who have motivated me to finish this dissertation. I am lucky to have this extraordinary team helping me accomplish this wonderful journey.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AR	Augment Reality
BEOL	Back-end-of-line
BW	Beam Width
CMFB	Common Mode Feedback
DAC	Digital-to-analog Converter
EO	Electro-optic
ESD	Electrostatic Discharge
FEM	Finite Element Method
FEOL	Front-end-of-line
FOM	Figure of Merit
FOV	Field of View
FWHM	Full Width Half Magnitude
GSA	Gerchberg–Saxton algorithm
IC	Integrated Circuit
LC	Liquid Crystal
LDE	Layout Dependent effect
LIDAR	Light Detection and Ranging
MB	Motherboard
MEMS	Microelectromechanical system
MIMCAP	Metal-insulator-metal Capacitor
MOMCAP	Metal-oxide-metal Capacitor
MOSFET	Metal-oxide-semiconductor Field-effect Transistor
NPA	Nanophotonic Phased Array
OC	Optical Chip
OPA	Optical Phased Array
PCB	Printed Circuit Board
PEC	Proximity Effect Correction
PTAT	Proportional to Absolute Temperature
PWM	Pulse-width modulation

RF	Radio Frequency
SC	Switch Capacitor
SOI	Silicon-on-insulator
SPICE	Simulation Program with Integrated Circuit Emphasis
TF H/TC	Thin film Heater/Thermocouple
TO	Thermo-optic
TOV	Through-oxide Via
VR	Virtual Reality

Chapter 1: Overview

1.1 Nanophotonic Phased Array

A phased array is a collection of individual electromagnetic emitters, whose output sums in the far field to form a desired radiation pattern. This is accomplished by controlling either the phase or amplitude (or both) of individual beams. The output beams interfere with each other either constructively or destructively in such a way as to create the desired pattern.

The basic idea of a phased array can be traced back to the work of the Nobel laureate Ferdinand Braun [1]. Braun's system consisted of three equally spaced antennas and a quarter-wave delay was introduced to any one of the antenna's output, steering the radiation beam with 120° resolution. Numerous research studies [2,3] have been conducted in this field since then and electronic phased arrays using radio frequency (RF) waves are in widespread use in beam forming [4-8] and beam steering [9-13], radar systems for military [14-18] and weather research [19-21] as well as in mobile communication application [22-27].

Optical phased arrays (OPAs) using optical beams (mostly near infrared or visible light) find even wider applications than their electronic counterpart. The applications of these devices include light-detection-and-ranging (LIDAR) devices [28-34] optical communications [35-42], 3D holographic imaging [43-48] and many others. The advance of deep submicron semiconductor technology has led to remarkable developments in the utilizing ultra-fine nanoscale devices. The goal of this work is to achieve compact, large-scale, and high performance nanophotonic phased array (NPA) system. I have focused on thermal phase control. As such, this thesis focuses on a novel electronic current driver circuit for precision control of micro-heaters. I have designed and realized this circuit for scalable NPA systems in holographic virtual reality (VR) imaging and LIDAR applications.

1.2 Characteristics of a Nanophotonic Phased Array System

The performance and physical characteristics of a NPA system can be characterizes by the following criteria [49]:

Wavelength: the wavelength of the light beam used in the system. Most NPA systems are operated in the infrared region using 1550 nm wavelength. Shorter wavelength systems using visible red light with wavelength around 630 nm has been demonstrated [47-48,50-51].

Array size: the number of the radiating antennas in the system. The radiators can be arranged in one or two dimension(s) and the distance between them is not necessarily equal.

Pitch size: the distance between two adjacent antennas of a uniformly-spaced NPA. Together with the array size, this determines the maximum field of view and the minimum beam width.

Control mechanism: how an NPA varies the phase or the amplitude of its output beams. These mechanisms can be categorized as: mechanical, e.g. micro-electromechanical system (MEMS) mirrors [52], and non-mechanical, e.g. electro-optic (EO) [53] or thermo-optic effect (TO) [43], methods.

System integration: the methods of assembling a NPA and its electronic control circuit. The system can be "unintegrated," utilizing external control [54]. It can be heterogeneously integrated where the NPA and an electronic control circuit are located on separate chips [55], or fully integrated in a single-chip by designing the fabrication process of the NPA to be compatible with silicon process technology [56,57].

Phase error: the difference between the desired and actual phase of the output beams. It may result from process variations in both optical and electrical systems, non-ideal effects in electronic driver devices, and thermal crosstalk between the array elements. A feedback control mechanism can be incorporated to minimize the phase error.

Power efficiency: the power required to generate a 2π phase shift for a full range of amplitude variation. This is a particularly important parameter for NPAs adopting the thermo-optic effect and it limits the array size as will be shown in the later section.

To sum up, an ideal NPA system should have large array size and small pitch to minimize the beam width and maximize the field of view for high resolution. An electronic control circuit with feedback control should be tightly integrated with a NPA to reduce the system complexity and increase the output accuracy. The power requirement should also be kept low for achieving a large-scale system.

1.3 Current NPA Technologies

In this section, I briefly review some current available NPA technologies.

1. Liquid crystal NPAs [58-63]: Liquid crystal (LC) NPAs are based on the birefringence property of liquid crystals. A typical LC device consists of a group of elliptical molecules sandwiched between two parallel alignment glass plates. Transparent electrodes are deposited on the external surfaces of the glass plates. Light is incident normally on one plate and emerges from the other.

Without any applied voltage, the LC molecules align with the direction of the alignment plates. An electric field between the two plates rotates the molecules in the direction of the field, and the rotation angle depends on the field strength. Therefore, the effective refractive index along the normal of a LC device can be modulated by the applied voltage This, in turn, alters the optical path length through the liquid crystal medium. A NPA is realized by applying various voltages to different locations of a LC device so the phases of the output light beams passing through different parts of a LC device can be independently controlled.

The major advantage of this approach is that it benefits from the matured LC display technology, so high resolution is realizable. The device is generally low power.

The response time is limited by the relaxation time for the LC molecules returning to idle state and it is on the order of millisecond depending on the parameters such as the device thickness and LC viscosity. Therefore, it is more challenging to achieve agile beam scanning in LIDAR applications. Another drawback is that it is difficult to integrate a LC NPA with other integrated optical devices such as devices fabricated with silicon-on-insulator (SOI) technology without special process steps.

2. EO NPAs: utilize the Pockels or Kerr effects which describe a linear and a quadratic relationship between the variation of the medium's refraction index and the electric field presented in the medium respectively [53,58,64]. Unlike the case of the LC phase shifter, the refraction index is modulated by changing the material's dielectric permittivity and polarization through the applied electric field. The response time of an EO phase modulator is faster than that of an LC device because no macroscopic molecular movement is involved in the modulation process. Devices with operation speed on the order of gigahertz are commercially available.

The major drawback here is that the EO coefficients of most common EO materials are quite low. To achieve the required 2π phase shift, a high input voltage is needed or the length of a device along the light propagation direction must be several orders of magnitude greater than the wavelength of the light [65,66]. High input voltage requirement poses design challenges on electrical driver circuit using advanced submicron technology while long device size limits the field of view of a two-dimension NPA.

3. MEMS NPAs: consists of an array of thin film micro-mirrors mounted on micro cantilevers or springs [67-70]. Mirrors are moved, deformed or tilted by independent addressable electrical controlled actuators to increase the optical path. For example, mirrors controlled by piston actuators are moved vertically in parallel to the direction

of the incident light. A half a wavelength displacement is required to achieve a 2π phase shift.

In general, large phase shift with fast response time are relatively easy to be achieved by a MEMS NPA comparing with the other NPA technologies. It can be combined with other NPA methods to realize wide angle and high resolution beam steering at the same time by having a first stage MEMS NPA for discrete large angle control and a second stage NPA for fine angle adjustment filling the discrete steps in the first stage. Obviously, the main downside is the device reliability issue associated with physically moving parts.

4. Thermo-optic (TO) NPAs: These devices adjust the refractive index of a phase modulator by raising its local temperature above the ambient temperature [43]. Electric current flowing through a thin-film heater near a phase modulator or directly into a phase modulator provides the heat source to establish a desired temperature distribution. Like LC NPAs, the turn on time, determined by the temperature rise time, can be reduced by changing the input current profile. But the turn off time is fixed, depending on various system parameters. These include factors such as the thermal resistance between a phase modulator and a heat sink. However, the thermal response is usually faster than the rotation speed of LC molecules.

We have selected the TO method for use in our NPA system design because the fabrication is simple, and the device is fast and reliable without any mechanically moving elements. The major limitation is the power density. LC or EO phase modulators are mostly capacitor-based devices which can maintain their voltage without refresh over a short period of time. On the other hand, TO phase shifters are memoryless in most cases. To maintain a desired phase shift, a closed loop system to stabilize the pixel temperature and continuous input current are required. The power

dissipation may become prohibitively large when the number of elements in a NPA is large and the size of individual element is small at the same time. When a driver electrical circuit is closely integrated with a TO NPA, the maximum temperature at the electrical circuit must be kept below a certain threshold to ensure long-term reliability. This thesis provides a thorough thermal simulation study whose goal is to determine the maximum achievable array sizes for several different heater and system structures.

Another issue of a TO NPA is the thermal crosstalk. Since pixels are not thermally insulated with each other, heating up a target pixel would unavoidably increase the temperature at the adjacent pixels. A feedback control mechanism is needed to address this issue to prevent thermal runaway.

Table 1.1 shows a summary of the performance of 4 NPA technologies.

Table 1.1 Comparison of performance of current NPA technologies

NPA TYPE	**Speed**	**Power**	**Array Size**	**System Integration**
Liquid Crystal	SLOW	Low	High	Low
Electro-optic	Fast	Medium	Medium	High
MEMS	Medium	Medium	Medium	High
Thermo-optic	Medium	High	Low	High

1.4 NPA System Architecture and Control Circuit Design

Figure 1.1 presents the design of our highly integrated NPA system with its control circuitry. An electrical circuit chip, we referred to as the motherboard, (MB) chip. This chip is flip-chip bonded with an optical chip (OC). The OC contains the waveguides and heater elements, optical "antennas" and other components used for light-wave beam formation.

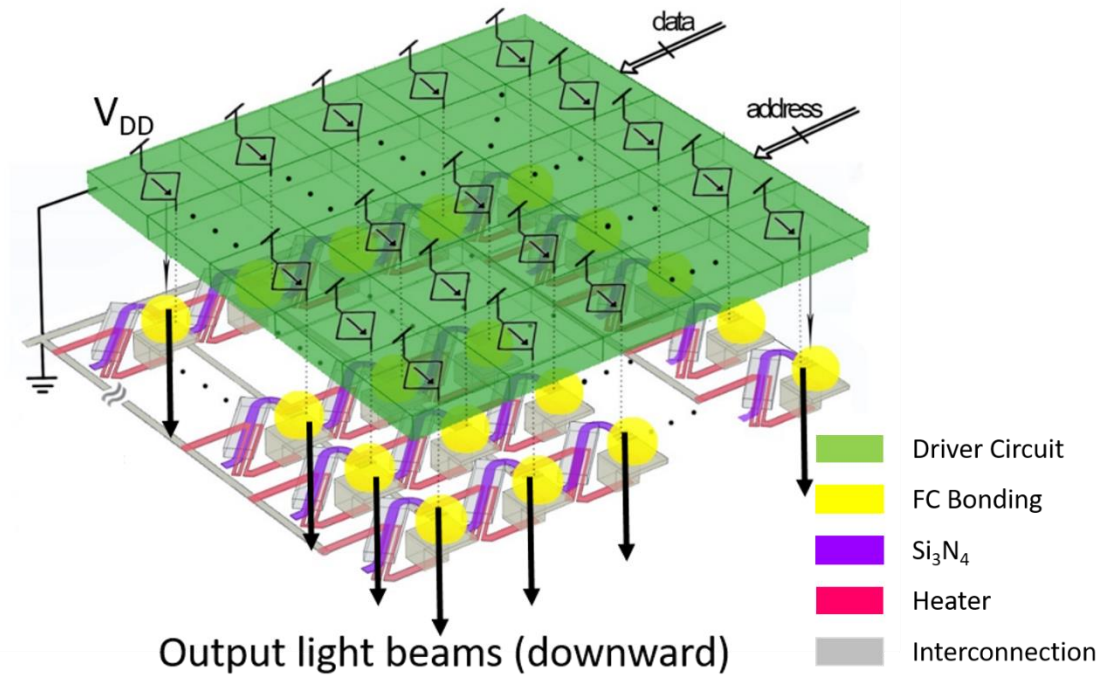


Fig. 1.1 Integrated two-dimensional NPA system

Each unit pixel in the OC consists of an integrated thin film heater/thermocouple (TF H/TC), a silicon nitride integrated waveguide/phase modulator and a radiating antenna. As mentioned in the previous section, the OC utilizes the thermo-optic effect to set the phases of each phase element's output individually. The transparent blocks around the waveguides represent air cladding which helps increase the effective optical refractive index contrast. The waveguide/phase shifter subsystems incorporate Bragg gratings which create "slow light" [71,72]. That is, the gratings effectively increase the optical density and polarizability of the waveguide. This increases the optical effective

refractive index so the maximum required temperature change for a 2π phase shift and the pixel's size are minimized.

On-chip metal thin-film heaters are used to control waveguides/phase modulator temperature. This, in turn, changes the optical path length and creates a phase delay with respect to other non-heated pixels. The zig-zag structure of the heater provides appropriate resistance for generating needed heat under certain voltage and current constraint. The antenna is designed to radiate the output light downward, so the output light won't be blocked or absorbed by all other parts of our system on top of the optical chip. The constructive and destructive interference of each phase element output resulting from the thermo-optic effect will form the desired radiation pattern in the far field, thus creating a virtual reality (VR) scene or a steered pointing beam useful in light detection and ranging (LIDAR) applications.

The MB chip consists of an array of driver units which provide electric current to the resistive TF H/TCs on the OC according to the input signal from a computer. A desired temperature distribution on the OC is achieved by Joule heating. Each driver contains a controllable MOSFET current source and each NPA unit cell or a group of NPA unit cells has its own dedicated current source depending on whether the NPA is used for a VR imaging system or a LIDAR application.

The yellow balls in Fig. 1.1 represent the bump bonds for flip-chip bonding. The MB and the OC communicate with each other through these bump bonds. To reduce the total number of connections between the OC and the MB, only one of the heater terminals is directly connected to the MB chip. This use of multiplexing to minimize interlayer contact vias is a unique contribution of this thesis. All the other heater terminals are tied together inside the optical chip before making contact to the circuit's ground (or rail voltage, V_{DD}). The flip-chip bonds tightly connect phased array unit cells

with their dedicated current driver unit without any area overhead and long-distance signal routing. Therefore, multiple identical bonded optical/circuit chips can be placed side-by-side and it enables the proposed NPA system architecture to be scalable.

The work reported in this thesis focuses on the thermal simulation of the NPA system and design, simulation, and implementation of the control driver unit in the MB.

1.4.1 System Thermal Simulation

To ascertain optimum system design performance, I conducted a set of finite element simulation to characterize the system thermal and electrical properties. First, a single NPA unit simulation was performed to determine the power required for 2π phase shift and the thermal response time. The results are used to create a simplified heater behavioral model and to guide the MB control circuit design.

Next, I studied the thermal crosstalk between adjacent cells using a 5-by-5 array. The results are used to perform proximity effect correction (PEC) [73-75]. Our control feedback mechanism can effectively reduce the phase error caused by thermal crosstalk and PEC can further minimize this error.

The power efficiency and thermal response time of NPA unit pixels with different thermal insulation structures and sizes are also investigated. A simplified NPA model and a simplified power distribution map of a test scene are used together to explore the maximum possible NPA size under different unit pixel structure conditions while keeping the MB maximum temperature below threshold. The results are useful in providing insight for NPA unit pixel design.

1.4.2 Thermocouple Design

The thin-film heater consists of two sections in series using different metals. The junction of the two sections forms a thermocouple. This integrated TF H/TC design is

also a unique contribution developed in this thesis. A voltage appears across the two heater terminals reflecting the temperature gradient along the heater itself when the circuit is open without any current passing through. Thus, we can use this signal to access the local temperature at the OC without additional contact between each pixel unit on the OC and its corresponding driver unit on the MB. The result is a kind of loop feedback control system. We believe that this "broken loop" feedback control system integrated as described below, is also a significant contribution to the NPA field. It is described in greater detail below. We use this on-the-fly temperature reading to stabilize the temperature over a period corresponding to framerate (for the VR application) or the target acquisition time for Lidar.

The first metal combination for thermocouple fabricating we attempted was chromium and nickel, because chromium has a high Seebeck coefficient. However, chromium is highly susceptible to oxidation. It is difficult to form a uniform and low resistance junction contact. We ended up choosing gold and nickel instead. Gold and nickel thin film thermocouples with thermal response ranging from 10 to 20 μV per degree Celsius have been reported [76]. We used finite element simulation to aid the design and verified our sample with an IR camera.

1.4.3 Driver Unit Circuit Design

The main design challenge of the driver unit circuit is that it must match the size of a NPA unit pixel (15 μm x 15 μm) and it must be repeated often enough to provide the desired image resolution. The driver circuit is fabricated using the Taiwan Semiconductor Manufacture Company (TSMC) 65 nm technology. The small "lambda" (i.e., minimum feature size) of this technology will allow us to create a compact array element. A layout using TSMC 28 nm technology is also provided for comparison purpose but due to the limited resources, we are not able to fabricate the design with 28

nm technology.

I provide two design approaches: a mixed-signal design with digital counter and an analog design. The analog design is more compact while the mixed-signal design is easier to apply a correction curve to offset the non-linear relationship between the current source input voltage and the controlled NPA local temperature. Because of latch-up and electrostatic discharge (ESD) design rules related to the deep submicron process, only the analog design meets the area requirement for a NPA in VR imaging system that needs independent control for all array elements. The mixed-signal design, on the other hand, can be applied to a LIDAR NPA which requires less control freedom. A detailed explanation will be provided in later chapter.

Some interface circuit devices between the computer-generated output and the control circuit array on the peripheral of the array are required for a fully operation NPA system. These include a digital-to-analog converter (DAC) to convert the digital computer outputs to control analog signals for our control circuit and an address decoder to direct the control signals to each individual array unit. These peripheral devices constitute “future work” for follow-ons to this thesis.

1.5 Contributions and Accomplishments

This dissertation has achieved the following goals:

1. I have created a system thermal simulation methodology for NPA system design space exploration essential for guiding integrated circuit design.
2. I have provided a complete integrated circuit design flow including the simulation, implementation, verification, and chip measurement of a novel and compact NPA system control driver circuit is shown.
3. I have demonstrated an integrated, small pitch size NPA system with feedback mechanism for accurate phase control in VR and LIDAR

applications.

4. I have developed a finite element thermal simulation model simplification methodology. The complex and high aspect-ratio geometry structures, thin-film and nanoscale devices, in the NPA system results in high density mesh elements in a finite element model. This exhausts computer memory resources causing either unrealistic long simulation time or out-of-memory error. Thus, an abstract system level geometry model which renders a reliable approximated simulation result within a reasonable time frame is needed.
5. I have developed an integrated TF H/TC device for heating NPA TO phase modulators and sensing NPA local temperature for feedback control. This novel device saves space in NPA unit pixel without introducing an extra device and enables a single connection between each pixel and its control circuit. Single contact ensures the flip-chip bonding is not limited by the NPA fine pitch pixel and avoids long routing wire to improve heating power efficiency.
6. I have designed and demonstrate a compact driver unit circuit using deep submicron semiconductor technology for a TO NPA with broken-loop feedback control to minimize phase error results from fabrication device mismatches and thermal crosstalk between neighboring elements. Two control approaches, analog and mixed signal, which are optimized for VR and LIDAR applications respectively, have been shown. The driver unit on average meets the main design constraint 15 μm -by- 15 μm per pixel.

In summary, this work demonstrates a systematic design flow for a cross domains, thermal and electrical, integrated system and a first scalable NPA system architecture

with an integrated fully independent phase control feedback circuit.

1.6 Organization of the Dissertation

This dissertation has been presented in six chapters and is organized as follows.

Chapter 1 (this chapter) provides a brief overview of nanophotonic arrays, their fabrication and applications. Contributions and accomplishments of this work are presented at the end of this chapter.

Chapter 2 gives the background of beam steering and holography applications followed by literature reviews of current NPA technology. Design challenges for optical-electrical system integration and electrical driver circuit for NPAs, and methods of on-chip temperature measurement in semiconductor technology are discussed.

Chapter 3 presents the thermo-electrical simulation for our NPA system for aiding the driver circuit and thermocouple design. The simulation of local temperature distribution at a single pixel using detailed model geometry is provided first followed by a small array of pixels for the thermal crosstalk (proximity effect) study. Various thermal insulation structures were studied to determine the trade-off between the fabrication complexity and power efficiency. A simplified NPA system model is provided for investigating the maximum achievable array size in light of the possibility of thermal damage.

Chapter 4 describes the control driver circuit design, SPICE simulation results and the layout implementation in 65 nm technology for both analog and mixed-signal approaches. Design considerations of each component are discussed in detail. A trial design in 28 nm technology is also provided for comparison purposed.

Chapter 5 shows the experiment results of our thermocouple and NPA driver IC. The performance of each component in the control driver is characterized and the broken-loop control operation is verified by using a resistor divider to simulate the

actual temperature feedback signal from the OC. The control error is measured. The last part of this chapter discusses the performance degradation of image forming and beam steering applications caused by phase errors from various sources. The results demonstrate that our control circuit improves the image quality and steering accuracy effectively by minimizing the phase errors.

Chapter 6 summarizes this research and points out the direction of future work for completing the entire system.

Chapter 2: Background

This chapter provides background information on the concepts forming the basis of this work, including the following points:

- Background on the theory of beam steering and 3D VR image forming using an NPA.
- Brief explanation of the functioning of the thermo-optic phase modulator.
- Detailed literature review of NPA technology.
- Challenges of integration between the optical and electrical devices.
- Discussion of the topology, design considerations and feedback control mechanism by sensing the local temperature at the NPA pixel of the electrical control driver circuit.
- Review of possible temperature sensor devices for the integrated NPA system.

2.1 Optical Phase Array

An OPA steers its output beam by focusing the output optical power on a specific direction. Constructive interferences are formed at the far field along the target direction while destructive interferences occur at the rest of the area. Assuming an equally spaced one-dimensional N-pixel array with a uniform phase difference between each adjacent pixel pair in the beam forming application, the output beam can be steered in one dimension and the steering angle between the output beam and the normal to the array surface can be represented as:

$$\theta = sin^{-1}\frac{\lambda\varphi}{2\pi d} \tag{2.1}$$

where d is the spacing between neighboring antennas, λ is the wavelength, and φ is the adjacent phase difference [28]. According to equation (2.1), the phase difference between the first pixel and the last pixel is (N-1) φ. Obviously, the phase difference

may exceed 2π but since the phase repeats itself every 2π, the phase of all pixels which exceeds 2π can be replaced with its modulo 2π value. Therefore, the maximum phase shift of a modulator's output needed is 2π.

The spacing, d, determines the location of grating lobes. The angle between the main lobe and the first positive/negative grating lobes limits the steering range or field of view (FOV) as:

$$-sin^{-1}\frac{\lambda}{2d} \leq \Delta\theta \leq sin^{-1}\frac{\lambda}{2d} \tag{2.2}$$

An ideal full range ±90° FOV requires the spacing equal to half wavelength. Since the wavelength commonly used in a NPA is around 1550 μm in infrared spectrum and submicron in visible spectrum, wide FOV can only be achieved with the spacing on the order of 1 μm.

The output beam width (BW), defined as the full width at half power of the main lobe (or FWHM), is proportional to:

$$\mathrm{BW} \propto \frac{\lambda}{Nd} \tag{2.3}$$

This parameter defines the spatial resolution. As explained earlier, the spacing should be kept small for large FOV. Therefore, for a given wavelength, the only parameter left for optimizing the BW is the phased array size N.

The above equations clearly show that a pratical OPA must meet that (i) the array size has to be large enough to deliver high resolution radiation pattern at the far field and while (ii) the spacing between each antennas should be close enough to suppress the magnitude of grating lobes for adequate field of view. As shown in Fig. 2.1, based on the relative location of phase shifters and antennas, OPAs can be categorized into two groups: (i) Antennas are tightly placed together. Phase shifters are on the periphery, outside the array of antennas and are routed by waveguide to their corresponding antennas. (ii) Each antenna and its phase modulator are put together to form a unit pixel.

The former case has been realized in one-dimension using OPAs with long parallel grating antennas placed closely together [54]. Small size 2D OPAs have also been realized, as the routing scales quadratically with the array size in 2D arrays. Larger 2D arrays require a larger pixel pitch to accommodate the interconnection congestion between the antennas and the phase shifters as the array size goes up which ultimately prevents the realization of a compact antennas array. It can be shown that the minimum pixel pitch is related to the array size, the width of waveguide and the minimum spacing between two parallel waveguides without interference. A non-uniform, sparse 2D array with 128 antennas selectively placed on a 27-by-27 grid which has same FOV and BW as a 22-by-22 array (but fewer interconnects) has been reported [57]. However, a sparse array may lead to larger phase error for generating random radiation patterns because of less phase control freedom. The result is like truncating the individual terms in a Fourier expansion.

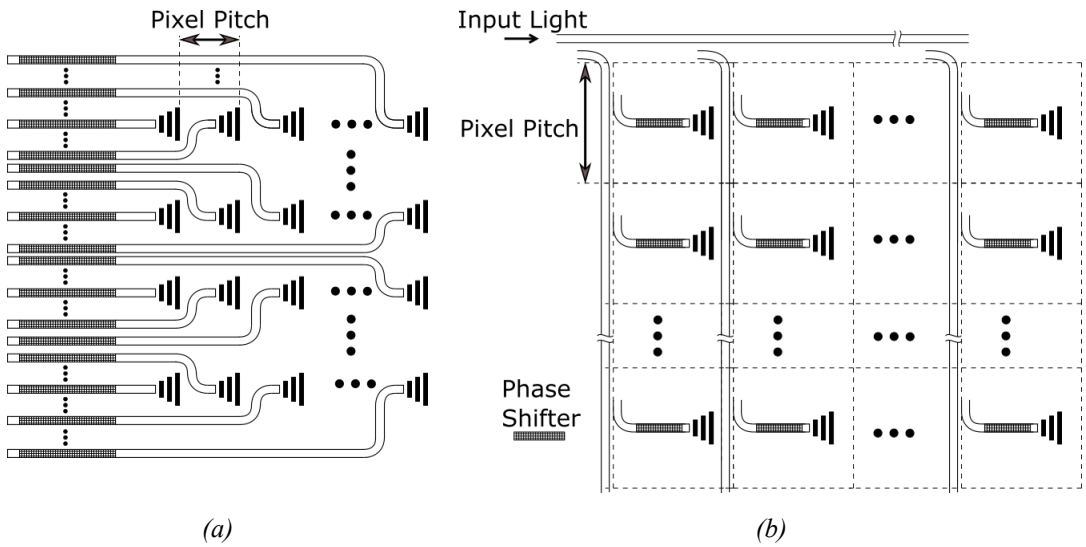


Fig. 2.1 Different NPA architectures (a) Phase shifters outside the pixels (b) Phase shifters inside the pixels

2D beam steering in these 1D NPAs are realized by using a combination of phase

and wavelength tuning. The additional steering dimension is done by changing the wavelength which varies the angle of light coupling out the antenna grating. In contrast, a 2D NPA can perform single wavelength 2D beam steering using pure phase tuning or forming arbitrary radiation patterns. Nevertheless, most current 2D NPA technologies still struggle in the tradeoff between FOV and pixel size due to the size of waveguides, phase shifters, antennas and the power limitation.

2.2 Thermo-optic Phase Modulators

Several phase control mechanisms were briefly introduced in the chapter 1. TO phase modulators offer simple operation and adequate response speed without any moving parts. The detailed discussion of TO phase modulators which is used in our NPA system is presented here. The phase delay of a light beam passing through a waveguide with length L can be expressed as:

$$\varphi = \frac{2\pi L}{\lambda} n \tag{2.4}$$

where n is the effective refractive index of the waveguide. Assumed two identical waveguides, one is uniformly heated above the ambient temperature by ΔT and the other is maintained at the ambient temperature. The phase difference between the two output beams can be approximated as:

$$\Delta\varphi = \frac{2\pi L}{\lambda}\Delta n = \frac{2\pi L}{\lambda}\frac{dn}{dT}_{T=ambient} \Delta T \tag{2.5}$$

This assumes that thermal expansion is negligible comparing to the thermo-optic effect and the thermo-optic coefficient is constant within the range of temperature variation. For instance, the thermo-optic coefficient (dn/dT, the temperature coefficient of the refractive index) of silicon is $1.8\text{x}10^{-4}$ K^{-1} at 300 K and a wavelength of 1550 nm [77]. The phase shifter waveguide with length of 10μm requires a significant, impractical temperature change (about 860 K) for a 2π phase shift. This calculation result is

consistent with the simulation result in [78].

From equation (2.5), given a wavelength, the required temperature change for a full range 2π phase shift is inversely proportional to the product of the length and the thermo-optic coefficient of phase shifters. Thus, the simplest way to reduce the required temperature change, hence the maximum input power, is to increase the length of phase shifters. In communication network applications, it is common to have the length of a phase shifter in optical switch a few hundred micrometers long in order to achieve low power operation. In contrast, for 2D OPAs with phase shifters inside the pixels, the length of a phase shifter is limited by the pixel pitch.

In holographic VR/AR display system, the spectrum of the input light beam is in the visible light range. The 1.12 eV indirect bandgap of silicon results in high absorption for visible spectrum; thus, it forbids us to use silicon as the material of the phase shifter waveguide core. The bandgap energy of silicon nitride (Si_3N_4) is around 5 eV (varying with manufacturing process flow) [79]; hence, it is transparent in the visible spectrum and suitable for VR/AR applications.

The thermo-optic coefficient of Si_3N_4 is reported to be between $3x10^{-5}$ K^{-1} and $6x10^{-5}$ K^{-1} at 620 nm (red light) [80]. This is 3 times less than for silicon at 1550 nm. To minimize the length of the phase shifter and the temperature change at the same time for large FOV and low power dissipation, we incorporate the "slow light" effect into the phase shifters. The phenomenon of slow light offers a way to fulfill the design requirement by increasing the effective refractive index at the expense of transmission loss [81,82]. The slow light in our NPA system is realized by incorporating Bragg gratings, an optical resonant structure in which light experiences slow group velocity, into the waveguide/phase shifter [71,72]. Our group achieved a slowdown factor of 6 and 10 with more than 80% transmission rate in Si_3N_4 and silicon respectively.

Therefore, the temperature changes are about 350 K and 90 K for a 2π phase shift in VR/AR NPAs with 620 nm wavelength and LIDAR NPAs with 1500 nm wavelength respectively.

2.3 Phased Array 3D Display

Over the past few years, research into virtual and augmented reality (VR and AR) has gained unprecedented attention. Numerous consumer wearable VR and AR displays for entertainment purpose, such as the Oculus Quest 2 [83], Samsung Gear VR [84], PlayStation VR [85] and Microsoft Hololens [86], are available on the market. This rapid growth has resulted from (a) advances in commercial light weight sensor technology and (b) high performance computing multicore processors enabled by advanced semiconductor technology. Besides entertainment, VR and AR technologies have great potential for applications in many different fields, such as military, healthcare and education, which will profoundly change the way we use information and interact with the world around us.

Stereopsis and holography are two widely used methods in current 3D display technology. Stereoscopic in 3D systems require some apparatus (e.g., a headset) which generate a pair of images of a scene from different angles for each eye to view a scene. These two similar images are merged in our brain to create an illusion of depth. This process may introduce psychophysical problems and distortion of the portrayed scene associated with their viewing [87]. Popular press reported people complaining about headaches, nausea, blurred vision, and other symptoms of visually-induced motion sickness after watching 3D films [88]. Studies have shown that the vergence-accommodation conflict is the main reason for psychophysical problems including fatigue [87], discomfort [89], induced binocular stress [90,91], difficulty in fusing two images into a stereo pair [91], and misperception of scene geometry [92,93].

Figure 2.2 illustrates the difference between viewing a real scene, (A and C), and a conventional 3D display (B and D). When viewing a real object, lenses in human eyes change their curvature according to focus cues. This is called accommodation and it allows the user to form a clear blur-free image near the object. As it is shown in Fig. 2.2 (C), the circle at the center is clear while the rest of the grid is blurred. In the meantime, the viewer's eyes move inward or outward to create a vergence distance by the depth cues. The focus and vergence distance are about the same in the case of viewing real object (Fig. 2.2 (A)).

Research also indicates that these two different processes are not independent, i.e. the focus cues will affect the vergence distance and vice versa. The response time of switching from focusing on one object to another is faster using both mechanisms than using only one of them. On the other hand, the focus cues of the object in a 3D stereoscopic scene force the viewer's focal distance be the distance between the eyes and the screen, while the depth cues tell viewer's eyes to converge either in front of or behind the screen depending on the object's position in the portrayed scene. Viewers must uncouple the natural neural response of vergence and accommodation. In Fig. 2.2 (D), both the center circle and the entire grid are clear instead of focusing at the center only and it causes the perceived scene differs from the real scene which it simulates. The perceived scene is flattened or exaggerated compared to the real scene. This vergence-accommodation conflict, which is inherent for all VR and AR displays available today, makes it difficult for viewers to fuse the perceived image binocularly, and is the main reason for visual fatigue and other psychophysical problems mentioned above. Therefore, a natural-to-senses VR and AR 3D displays is needed to solve psychophysical problems and degraded visual performance caused by vergence-accommodation conflict.

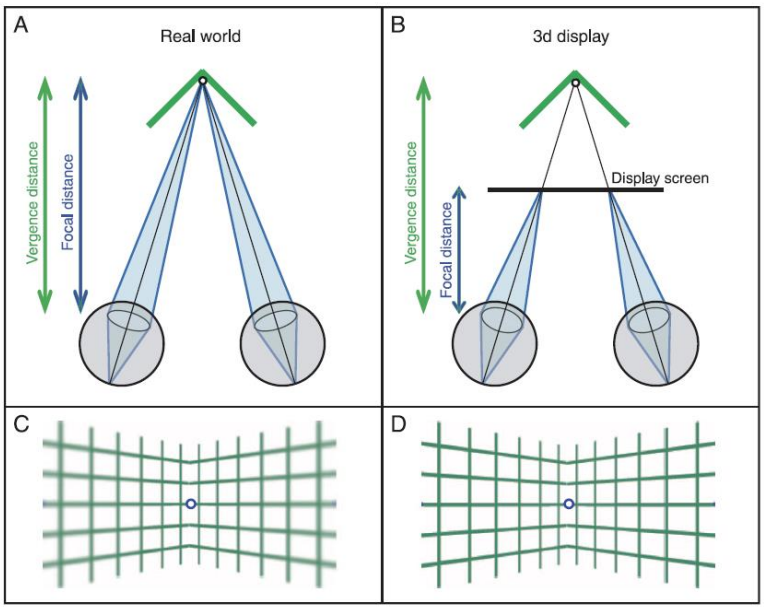


Fig. 2.2 Vergence and accommodation in real scene and conventional 3D display [87].

In natural viewing, because of the varying aperture and variable focus (adjustable pupil size and curvature of lens) of human eyes, a bundle of light rays impinges upon the retinas to create a clear image of the scene through the pupils and the lens. Therefore, to achieve a natural-to-senses display, one has to reproduce the entire light pattern stimulating the retina.

Holographic 3D displays simulate this process by recording and reconstructing both the amplitude and phase information of an object. Conventional holography recording and reconstruction processes are shown in Fig. 2.3. When recording an object, a coherence light source is split into two beams. One beam impinges to the object and the light reflected by the object is referred as object light. The other beam is called reference light. A photographic film is used to record the interference pattern of the reference and object light; thus, both the amplitude and the phase information are stored.

Assuming the transmittance of the recording film, $t(x, y)$ as a function of location x and y, after the exposure is linear related to the intensity of the interference of the object beam and the reference beam can be expressed as:

$$t(x,y) = \beta(|O(x,y)|^2 + |R(x,y)|^2 + O(x,y)R(x,y)^* + O(x,y)^*R(x,y)) \quad (2.6)$$

where β is the gain of the recording film, $O(x,y)$ and $R(x,y)$ are the complex wavefronts of the object beam and reference beam, respectively [94].

During the reconstruction process, the same reference beam passes through and is modulated by the recorded film. The mathematical expression of the modulated wavefront can be expressed as

$$\begin{aligned} t(x,y)R(x,y) &= \beta OR^*R + \beta(|O|^2 + |R|^2 + O^*R)R \\ &= \beta|R|^2O + \beta(|O|^2 + |R|^2 + O^*R)R \end{aligned} \quad (2.7)$$

The first term in equation (2.7) is the original object beam multiplied by a constant $\beta|R|^2$. If the light field represented by the second term in equation (2.7) is separated from the first term of interest, the radiation pattern of the object beam is recreated even though the original object is removed.

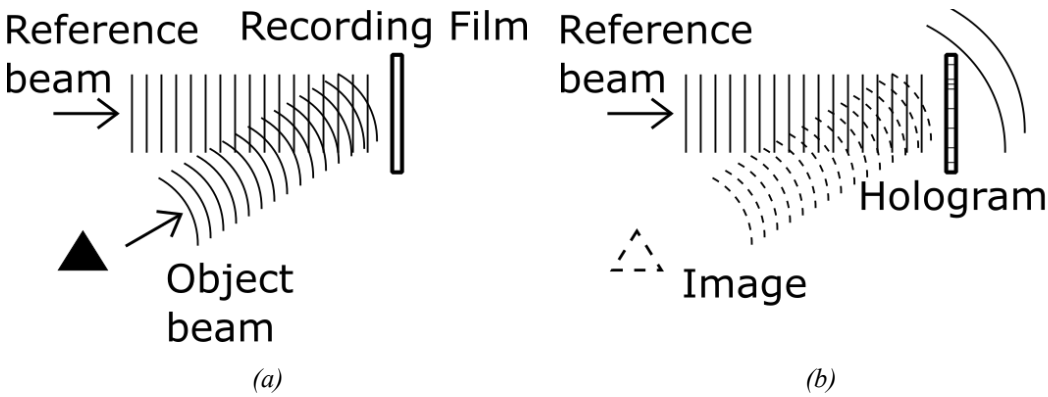


Fig. 2.3 Holography recording and reconstruction (a) recording and (b) reconstruction the image

A 2D OPA that is capable of generating any desired radiation pattern can replace the holographic recording film described above. Instead of recording an actual interference pattern, the phases of output beams from each OPA pixel which together render an approximate light field of the object beam can be calculated numerically using the Gerchberg–Saxton algorithm (GSA) [95]. In addition to portraying a still object, a 3D animated scene can also be created with an OPA by dynamically changing its output

phases. Therefore, OPAs are suitable for creating natural-to-senses VR and AR displays. Since the light pattern of the portrayed scene is faithfully reconstructed by an OPA, the focus cue is consistent with the depth cue as in the real scene. Therefore, the vergence-accommodation conflict problem is resolved by the phased array approach.

Clearly, the complex nature of the optical problems alluded to above makes the design of such a system non-trivial. Problems encountered include high power consumption (needed for adequate resolution in practical applications) and complicated circuit design for precise phase control.

2.4 Feedback Controlled Driver Circuit

In a TO OPA, thin film heaters are placed near the TO element [54]. The resistance of a silicon TO phase shifers is lowered by increasing the doping concentration of the silicon forming the integrated heater/TO phase shifer device [43]. A driver circuit containing an array of sources provides current to each of these heaters, maintaining a target temperature in each phase shifting element. Each heater and its current driver must be addressed independently to create an arbitrary interference pattern.

One possible straightforward implementation of the driver current source is to use a number, N, of parallel binary weighted current mirrors. Each current mirror is either on or off depending on the received N-bit binary input. Due to the high power density of TO OPAs, power loss at the driver current source and heat that transfers from an OPA to a driver circuit when they are tightly integrated can siginificanly increase the temperature at the driver circuit. Driver circuits need to include some kind of temperature compensation approach or a phase zero calibration for every cell in driver circuits should be performed to account for the devices' characteristics variations within the entire operation range.

A lookup table can be used to store the mapping between the N-bit input and the phase shift of the output beam. However, other temperature dependent parameters such as the heater's resistance, device mismatches from the fabrication imperfection and the thermal crosstalk between neighbor units complicate the calibration process. Furthermore, the electrical properties of the driver circuit may be continuously changing during the run time after calibration because of electromigration which results from high current density. All these effects together cause the actual temperature distribution in the OPA to deviate from the ideal one; hence, phase errors are introduced. Phase errors cause the output beam deflection in LIDAR applications or degenerate the image quality in display systems and the design quality sets the phase error tolerance.

As phase is determined by the temperature of the waveguide/phase shifter, this means that temperature must be controlled to this tolerance as well. A reliable control strategy for the driver circuit to meet the error tolerance constraint is needed.

Most current OPA designs appearing in the open literature focus only on the optical devices without driver circuits. External power sources are used to deliver the heating current. To the best of the author's knowledge, only three OPA systems with control circuits have been reported [55-57]. Chung et al. reported a monolighically integrated OPA system [56]. The optical devices and control electronics are fabricated on same chip using SOI CMOS process. The driver current source is similar to what we mention above. The reference current for the current mirror is temperautre independent and the output current source consists of four cascoded NMOS to equally distributed the heater-to-ground voltage drop. Hence, a higher voltage can be connected to the heater to minimize the output current and the power dissipation at the current sources. To avoid interconnection congestion between the OPA and its driver circuit, group instead of independent phase control is implemented.

Fatemi et al. presented a sparse 2D OPA with its control circuit mounted on a common printed circuit board (PCB) [57]. A high voltage pulse-width modulation (PWM) switch with soft turn on forms the output current source. Since the phase shifters are located outside the antenna array, they can be place far apart from each other to eliminate the thermal crosstalk. The driver circuits use a row and column address structure to access phase shifters in a time multiplexing fashion. As a result, only one column of phase shifters is heated up at the same time. Clearly, the sparse array prevents forming arbitrary radiation patterns without sacrificing the image quality. Also, the time multiplexing control scheme may cause large ripple errors.

An OPA system which is 3D integrated with its control circuit at wafer level using

a through-oxide via (TOV) has been reported [55]. Although the pitch size of the TOV can be as small as 7 μm, 10 TOVs in parallel form a single contact between the OPA and driver circuit to lower series resistance of TOV itself. The gates of a current source MOSFET is driven by pulse density modulation signals.

The above control circuits lack a method to dynamically adjust their outputs to balance the change of device properties during run-time. Some feedback control methods should be incorporated into the driver circuit to self-adjust the control signals of current source achieving the desired phase of the output beam as illustrated in Fig. 2.4. In this way, mismatches, thermal crosstalk, electromigration and other possible factors that degrade the system accuracy are all considered. Since the thing we want to control is the phase of output, ideally, we should measure and use the phase of the output as the feedback signal to control the amount of current.

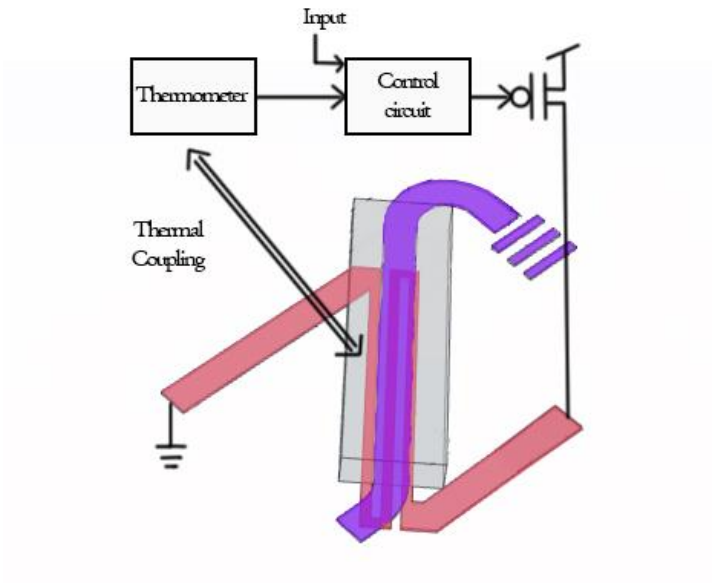


Fig. 2.4 Simplified diagram of phased array unit cell and driver circuit with feedback control.

Nevertheless, the phase of the output beams cannot be measure directly with current technology and their interference intensity profile is difficult to analyze at run-time. OPAs with feedback phase control using an external IR camera have been reported

[29,46,96].

In [29], the output beam power is measured by an IR camera as a feedback signal, so the heater currents are individually tuned to maximize the beam power. In [46], a computer compares the OPA's output image captured by an IR camera with the ideal image and tunes the output current to each p-i-n diode accordingly. In [96], the phase of each pixel's output beam is measured by an interferometer. The interferometer's output is again captured by an IR camera. The achieved control error is about $\pi/100$ when the target phase shift is π.

Obviously, the major disadvantage of these feedback control approaches is the speed of the feedback loop is limited by the IR camera frame rate. This system setup is slow and bulky and is only suitable to display still images. Thus, other measurement must be taken as an indirect indication of output phase to realize a fast and compact feedback system.

Recently, an on-chip calibration LIDAR NPA using transceivers has been reported [97]. The phased array can receive reflected light from a reference object. The received light signal is transformed to a voltage signal on-chip via a balanced photodetector followed by a transimpedance amplifier. The control circuit searches a set of output signals to the phased array to maximize the voltage signal. For a large array, the number of possible control states is large. This technique may need a long time to find an optimum solution. In addition, this method is not suitable for VR applications because the radiation pattern for an image does not have a peak beam for reflection.

As explained earlier, phase is modulated by changing the temperature of the waveguide/phase shifter, temperature can be used as an indirect indication of output phase under an assumption that the variations of the length and refractive index of waveguides/phase shifters are negligible so that the same temperature at different cells

result in almost the same output phases. This implies a sensor device that produces an electrical signal reflecting the temperature surround it is required to form a feedback control loop.

Intuitively, the sensor device can be located either on the MB or on the OC. For the sensor on the MB, a standard and sophisticated sensor circuit design can be adopted, and it is tightly integrated with other elements in the control circuit. The back-end-of-line (BEOL) metal vias in the MB and the flip chip bond act as a heat pipe that draws thermal energy from the heater strip in the OC. The energy drawn is proportional to the temperature of the adjacent heater. The downside is that it makes an "indirect" temperature measurement. A sensor device must be thermally isolated from the neighboring driver unit; otherwise, due to the high thermal conductivity of the MB silicon substrate, it will register the heat dissipated from the adjacent current sources more than it registers the target pixel heat.

On the contrary, because of the limitation of the OC fabrication processes, the sensor device in the OC must be simple. Also, since the OPA pitch size is on the same order as the state-of-the-art copper pillar flip chip bonding [98,99], only one electrical connection between the OC and the MB per pixel and it is occupied by the heater; hence, the sensor device should be integrated with the thin film heater to minimize the number of bonding connections. Even though it is possible to have multiple contacts per pixel with more advanced integration technology such as the wafer-level TOV bonding with minimum 7 μm pitch size, it is still preferable to maintain single connection by connecting multiple contacts in parallel to reduce the IR power loss.

The major challenge of designing an integrated heater/sensor device is that we cannot sense the temperature at the same time we drive the heater. This means that we will have to switch the operational mode of the controller using added circuitry on the

MB chip. The multiplexing operation will slightly increase the complexity of the control circuit. Also, we will introduce phase ripple at the switching frequency. But this approach provides a “direct” and full range temperature measurement of the waveguide/phase shifter.

2.4.1 Sensors on the MB

Two types of temperature sensor circuit have been applied to integrated circuit systems requiring temperature regulation: a diode based Proportional To Absolute Temperature (PTAT) circuit [100-102] and a ring oscillator based PTAT circuit [103, 104]. A standard diode based PTAT circuit is shown in Fig. 2.5. It utilizes the temperature dependence of the base-emitter voltage of a bipolar transistor. Two currents with ratio n are injected into two identical matched transistor Q1 and Q2. Two transistors are assumed to be at the same temperature. The voltage difference between V_{EB2} and V_{EB1} which eliminates common temperature dependence terms is in first-order approximation proportional to the absolute temperature. This circuit consists of only four transistors without the bias current generation circuit. This can be shared with other elements in the control loop such as an operation amplifier (op-amp). Since the sensor circuit must be repeated in every MB driver element, this simple circuit offering a minimal real estate overhead is a potential candidate for the sensor device realization.

The ring oscillator based PTAT circuit is made from a current starved ring oscillator whose current is supplied from a PTAT current source. The PTAT current source is usually created using the same principle as the diode based one. The frequency of the ring oscillator is proportional to the absolute temperature because of the PTAT current source. Hence, the number of the peaks at the output within a fixed interval is the temperature reading. The precision of the measurement can be increased by extending the counting interval. However, the measurement is not instant. Also, the

additional required peripheral circuit such as a counter makes this approach less favorable to a system with demanding space constraint.

Our system's thermal simulation results, which will be presented in the next chapter indicate the sensor on the MB is not suitable for our OPA system architecture because of the strong thermal coupling between the adjacent driver units explained above. Nevertheless, these methods may be useful for other TO devices with relaxed space constraint.

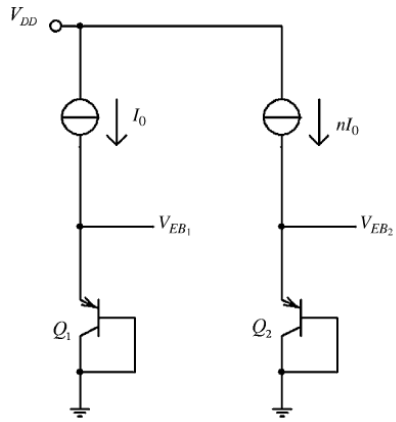


Fig. 2.5 Basic diode based PTAT circuit [102].

2.4.2 Sensors on the OC

Two possible approaches can be applied to measure the temperature using thin-film heater: (i) utilize the temperature dependence of the heater's resistivity; or, (ii) utilize the thermoelectric effect.

For case (i), the thin film heater acts as a resistance temperature detector (RTD). Two possible measurements scheme are as follows. First, by passing a known sensing current to the heater, we can get the heater's temperature by measuring the voltage across it. That means the heater is switched between two modes, heating mode (with a varying current depending on the target temperature at the phase shifter) and sensing

mode, with a fixed constant current. Without loss of generality (ignoring the higher order terms), the relationship between the temperature coefficient and the output voltage variation resulting from the rise of temperature is:

$$\Delta V_{out} = \alpha I R_0 \Delta T \tag{2.8}$$

where: I is the sensing current, R_0 is the heater's nominal resistance, α is the heater's first order temperature coefficient of resistivity and T is the temperature.

Given a heater material, the temperature coefficient of resistivity is fixed. Also, the heater resistance is kept small to deliver adequate power within some given voltage headroom. Therefore, the only design freedom left is in the sensing current. The sensing current should be large to get a measurable voltage variation. However, a large sensing current produces unwanted joule heating which affects the accuracy of the measurement. It is difficult to optimize the heater resistance for heating ability and sensing current for measurement accuracy at the same time. Moreover, the mismatch of the sensing current between each driving elements will introduce additional error. The limitations mentioned above make this method unattractive.

The second approach uses the temperature dependence of the resistivity to continuously measure the heater current and the voltage across it. The thin film heater can function as a heating element all the time without switching between modes. The measured voltage and current are mapped to the heater's resistance with an analog divider circuit. The divider's output is proportional to the temperature at the heater. However, an analog divider circuit is rather complex which is not able to fit within the allowed pixel space together with other driver unit components.

In case (ii), the heater is divided into two segments using different metals for each segment. Thus, the heater also serves as a thermocouple with the junction at the center where the maximum temperature occurs. The Seebeck effect generates a thermal electromotive force (emf) according to:

$$V_{out} = \int_{T_{end1}}^{T_{end2}} S(x,T)\, dT \tag{2.9}$$

where S is the Seebeck coefficient. If we assume that the two metals are uniform (that is, the Seebeck coefficient in the same segment does not vary along its length) and the temperature at the connection ends is able to reach an equilibrium close to the ambient temperature, equation (2.9) can be simplified as:

$$V_{out} = (S_1 - S_2)\Delta T \tag{2.10}$$

where S_1 and S_2 are the Seebeck coefficients of the two segments within the working temperature range of the positive and negative terminal respectively; ΔT is the amount of temperature rise at the heater center junction. To sense the temperature signal, the current driver is shut off and the same node is connected to a high input impedance device such as an opamp. The opamp's output voltage will depend on the temperature at the heater center.

Because the heaters also serve as temperature sensors, no additional contact is required between the OC and the MB. Moreover, the temperature sensor is in intimate contact with the heater element, and we believe that this will provide a more accurate sensing method, as compared with the "indirect" method.

Table 2.1 provides a list of the Seebeck coefficients at 20°C for common metals in bulk used in thermocouples [105].

Table 2.1 Values of Seebeck coefficient, S(μV K^{-1}), of common thermocouple material at 20°C

	Chromel	Fe	Au	Cu	Ag	W	Pt	Al	Ni	Constantan
S(μV K^{-1})	22.2	13.3	2.0	1.9	1.7	1.3	-4.7	-18.2	-19.5	-38.3

Chapter 3: NPA System Thermal Simulation

This chapter provides a comprehensive thermal study of our purposed integrated NPA system. It includes the followings.

- The steps and result of a thorough thermo-electrical simulation of a phased array unit with different thermal insulation structures are presented. The design choice of a trade-off between the heating power consumption and device complexity and the required maximum output current of the driver circuit are determined by this result.
- I used a small array to investigate the thermal crosstalk (thermal proximity effect). A system level thermal crosstalk matrix is extrapolated from this result and is used to conduct a further phase error analysis.
- A simplified system level thermal model was constructed and was used to explore the maximum realizable array size under different cooling conditions.
- The design and simulation of our thermocouple sample is explained. The thermocouple was used to determine the Seebeck coefficient of our choice of thin-film metal combination and was used to verify the temperature sensing approach.
- The actual thermocouple was fabricated and evaluated using infrared cameras

3.1 Thermal Model Description

From the background chapter, we know that the desired far field optical interference pattern is manipulated by changing the waveguide temperature and, subsequently, the refractive index of each individual phase shifters in the phased array

unit cell. The driver circuit injects electrical current into the resistive heaters raising the phase shifters' temperature to the target value. In this section, we present a thermal model for our NPA system and study the thermal and electrical coupling effect in the system. The simulation results provide insights for the driver circuit design: the maximum load current required for 2π phase shift, thermal crosstalk between neighboring unit cells and the control mechanism to minimize phase error and the cost of calibration.

Since we are interested in the interaction between the two different physical domains (electrical and thermal) of our system, COMSOL Multiphysics is used to perform finite element method (FEM) analysis for our system [106]. COMSOL Multiphysics provides several add-on modules for different physics domains and a built-in material library. It offers two types of solver: fully coupled and segregated. For the fully coupled solver, a large matrix with all the physics described in the model and their interactions is created and solved directly. On the other hand, the segregated solver decouples the different physics domains and solves them separately in sequence. The results in one physics domain in the previous step are used to modify the boundary conditions of the other physics domain in the next step. The computation stops when a converged solution with the error smaller than a preset tolerance is found for all physics domains. The fully coupled solver has fewer issues about reaching a converged solution but needs much more computer memory for the large single matrix and the segregated solver is the opposite.

Given the size and high aspect ratio of the NPA system geometry, the model must be simplified and solved by the segregated solver. In our model, we use the heat transfer module and the AC/DC electrical module. With these two modules, stationary and time dependent studies of joule heating in our system are performed.

In the steady state, the governing equations in the electrical and heat transfer domain are:

$$\nabla \cdot J = Q_j \tag{3.1}$$

$$J = \sigma E \tag{3.2}$$

$$E = -\nabla V \tag{3.3}$$

And

$$\rho C_p \mathrm{u} \cdot \nabla T + \nabla \cdot q = Q + Q_{ted} \tag{3.4}$$

$$q = -k\nabla T \tag{3.5}$$

respectively, where J is the current density, Q_j is the current source, σ is the electrical conductivity, E is the electric field, V is the electric potential, ρ is the density, C_p is the heat capacity at the constant pressure, u is the velocity field, T is the temperature, q is the heat flux, Q is the heat source per unit time, Q_{ted} is the thermoelastic damping and k is the thermal conductivity.

Equation (3.1) to (3.3) represent the continuity equation of electrical charges, Ohm's law and potential gradient respectively. Together, these equations result in electrical current conservation. In our case, the velocity field and the thermoplastic damping are zero, so equation (3.4) and (3.5) can be reduced and combined to yield:

$$\nabla \cdot (-k\nabla T) = Q \tag{3.6}$$

Equation (3.6) describes two things: (i) the heat flux is defined by the product of the material's thermal conductivity and the temperature gradient, and (ii), energy conservation. These two physical processes are coupled by the equation:

$$J \cdot E = Q \tag{3.7}$$

The heat dissipated from an electrical source is equal to the product of the amount of the electrical current and its voltage drop.

3.2 Model Geometry

The NPA system thermal simulations are divided into two parts: (i) detailed single

pixel and small array for thermal crosstalk and (ii) large array thermal interaction between the OC and the MB. Due to the limited computational resource and the size of FEM model geometry of our NPA system, the geometries for each part are simplified to different degrees. The simplification process takes several steps to ensure that the results are valid approximations.

For the first part, the 3D geometry model of a 3x3 NPA system in VR and AR applications is shown in Fig. 3.1. As stated earlier, to speed up the simulation and reduce the amount of computation resource required (e.g. memory), the model is simplified by only having 3x3 or 5x5 array unit cells instead of actually modelling a large phased array. The error introduced by this simplification is negligible for the following reasons: (i) The phased array and driver circuit only occupy a thin layer on the surface of the OC and the MB respectively. (ii) We are only interested in the heating response of a single unit cell and thermal cross talk between neighboring cells. A 5x5 array is sufficient for us to investigate these effects.

As shown in Fig. 3.1(a), the upper half and the bottom half of the drawing represent the MB and the OC respectively with several thin layers containing heaters, phase shifters, FC bonds, FC underfill and other materials in between. The thickness of the MB substrate and the OC substrate are both 1 mm. Silicon is used for the substrate and the front-end-of-line (FEOL) material of the MB in the model while a 10 μm thick silicon dioxide represents the BEOL and passivation layer. The actual detailed circuit structure is omitted here because the goal of thermal simulation is to derive the temperature distribution caused by the current flowing through the resistive heater located on the optical chip. Details of the semiconductor device simulation are discussed elsewhere. The detailed design, simulation and verification of the control circuit is performed by using a standard electronic circuit simulation tool (SPICE) and

layout tool (the Cadence virtuoso design suite). Those will be covered in the next chapter.

Our NPA, operating in the visible spectrum, is designed to be fabricated on an optical flat, so the material assigned to the OC substrate is silicon dioxide. The connections between two chips are the flip-chip bonds at the center of the model where the phased array cells are.

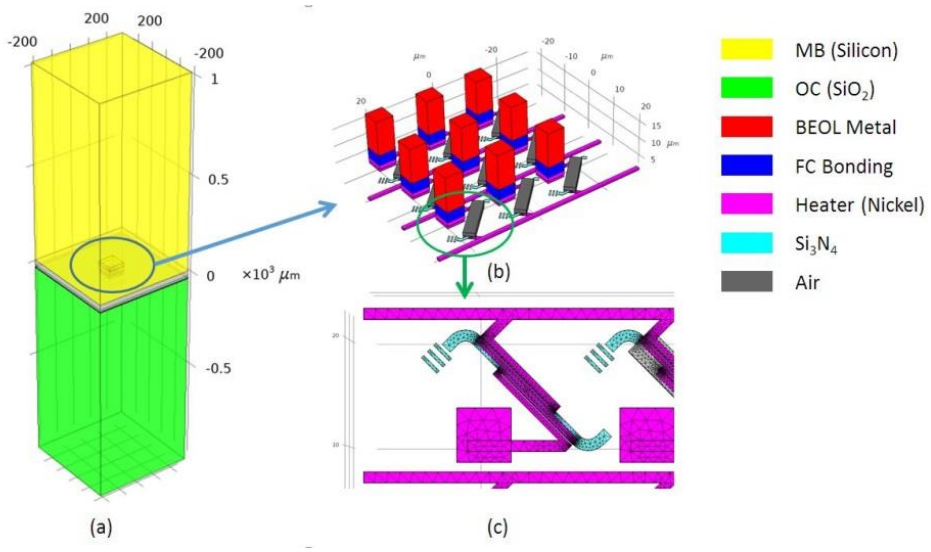


Fig. 3.1 3D geometry model of a NPA system in VR and AR applications.

A detailed view of 3x3 phased array cells' geometry is shown in Fig 3.1(b). The size of unit cell in the model is 15 μm x 15 μm with a 13 μm long phase shifter/waveguide at the diagonal (light blue). The gray rectangles enclose the waveguides represent the top air cladding and the three light blue bars structure at the end of waveguides are radiating antennas. The red rectangular columns are the back-end-of-line metal and vias of the circuit chip which are the electrical connections between the drain nodes of current source MOSFETs in the MB and the flip-chip bonds. Three horizontal pink lines are common ground buses tying one of each of the heater terminals to a common row together. All the ground buses are tied together and hooked

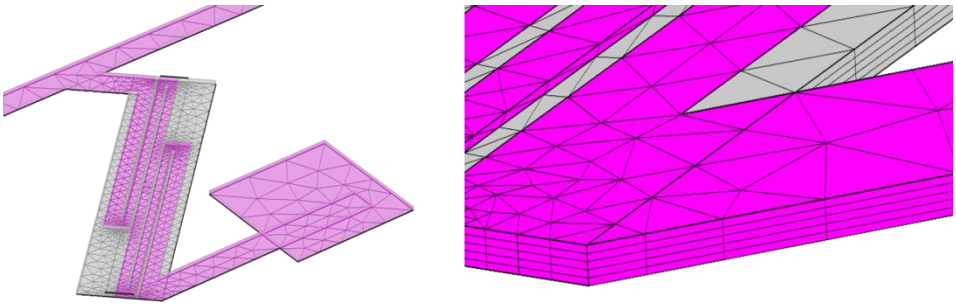
to the power supply, V_{DD}, outside the phased array. The entire structure is surrounded by SiO_2, except at the flip-chip bonds layer, which is surrounded by the FC underfill. A thin square box at the center top is the current source MOSFET. As mentioned before, the thermal model doesn't include any detailed circuit structure. This block is part of the MB FEOL and is uniform silicon in the model. We outline it in the figure to assign a secondary heat source whose function will be described below.

A close-up vertical view from the bottom of a single heater and phase shifter/antenna is shown in Fig. 3.1(c). The zig-zag resistive heaters are at the bottom with one end connect to their corresponding power contact leading to the MB current driver. The heater is designed to have its thinnest segments directly beneath the phase shifter for high heating efficiency by having the largest input voltage drop and heating power at these locations. The mesh grid on the heater and the phase shifter is also shown in Fig. 3.1(c).

The thickness of the heaters and the phase shifters are both 100 nm. Since the mode is well-confined, the distance between a heater and a phase shifter, and the thickness of the silicon dioxide bottom cladding, can be as thin as 500 nm for high heating efficiency while keeping the propagation loss negligible. The high geometric aspect ratio from these thin film structures leads to difficulties in mesh generation. The most common free tetrahedral 3D mesh elements are not suitable for thin layers and they either result in extremely small and dense mesh grids or cause mesh errors for not being able to insert appropriate mesh elements in some domains of the model.

COMSOL offers an efficient way to generate mesh grids for a thin layer called "the swept mesh approach," as shown in Fig. 3.2. The mesh generation processes are as follows. First, mesh grids on either the top or bottom surface, which is called the source plane, of the target thin layer are generated using 2D planar triangle mesh

elements. In Fig 3.2(a), the source plane is the bottom of the plane where the heaters are located. Next, the target thin layer is sliced by a number of planes parallel to the top and bottom surfaces uniformly. Finally, the mesh grids on the source surface are copied to all of these parallel planes toward the opposite surface, the destination plane, i.e. the mesh grids are swept from the source plane to the destination as shown in Fig 3.2(b). The resulted mesh grids of the source plane, the destination plane and all the equally spaced parallel planes in between are all the same. In Fig. 3.2(b), the heater layer is divided into five sections. A finer mesh can be obtained by having more vertical segments in the layer.



(a) (b)

Fig. 3.2 Swept Mesh on the heater plane

To apply the swept mesh method, two conditions must be met on the destination plane: (i) no mesh grids are created on the destination plane before constructing the swept mesh and (ii) all geometry shapes and edges on the destination plane must be included on the source plane. For the first condition, considering that our NPA model contains a stack of thin layers, a 100 nm heater layer, a 500 nm SiO_2 bottom cladding layer and a 100 nm phase shifter layer and so on from bottom to top, if the swept mesh approach is applied to the heater layer and the phase shifter layer first, it can't be used for the middle SiO_2 bottom cladding layer later, since the bottom and the top surface of the SiO_2 layer which are the top surface of the heater layer and the bottom surface of

the phase shifter layer has already been meshed. In this case, the SiO_2 layer must be meshed with free tetrahedral elements which do not work well on thin layers. To apply the swept mesh method to all stacked thin layers, the mesh sequence has to be in-order either from bottom-up or top-down.

For the second condition, if the destination plane contains some shapes that do not exist on the source plane, the triangle mesh grids generated on the source plane may not be fit into these shapes. Again, in our NPA model, the geometry of the heater is different from the geometry of the phase shifter and antenna. Thus, in a bottom-up swept mesh generation process, it is obviously that the mesh grids of the heater layer cannot be applied to the phase shifter directly. To solve this conflict, all geometries in the thin film layers above the heater layer are projected down to the heater layer. In this way, we create an extra fine triangle mesh grids on the heater layer which respects all the boundary constraints posed by the layers above. Therefore, this triangle mesh grid~~s~~ can be applied to all domains above seamlessly. In Fig. 3.2, one can see that there is a small rectangle inside the square terminal of the heater connecting to the current source MOSFET. The zig-zag part of the heater is partitioned into several pieces and these shapes originate from the geometry above the heater plane.

The input voltage is applied at the top surface of rectangular column of the target cell and zero voltage is assigned to the ground metal as boundary conditions. Hence, the voltage drop across the resistive heater of the target array unit cell induces a current flow. Equation (3.7) transforms the current flowing through the resistive heater into heat and this is the major heat source in the system. However, the current that flows through the resistive heater also flows through the current driver MOSFET. And so, a certain voltage drop must exist between the source and drain of the driver MOSFET to draw current from the power supply. The current sourcing transistor generates

unwanted heat in its own body (in the MB away from the waveguide). This secondary heat source cannot be neglected. Thus, we must explicitly assign a heat source to the region where the driver transistor is to get an accurate simulation of the actual temperature distribution. A secondary heat source is added on top of the rectangular column. The power that it dissipates is:

$$P = (V_{DD} - V_{in})I_{in} \tag{3.8}$$

where V_{DD} is the power supply voltage, and V_{in} and I_{in} are the input voltage and current on the top surface of the rectangular column of the target pixel. V_{DD} is set to a constant 2.5 V in the simulations according to the I/O voltage of the process technology, TSMC 65nm 1P9m, we used to fabricate our control circuit. V_{in} is the system input and ranges from 0 to 2.3 V. Here, we assume a 0.2 V drain-to-source voltage (V_{DS}) drop is required for the driver transistor at full load. The FEM solver will compute the value of I_{in} according to the input voltage V_{in} first, apply the secondary heat source to the model based on the equation (3.8) and evaluate its thermal effect.

The remaining model settings are the boundary conditions at the outer surfaces for cooling. Two types of cooling boundary condition are used in the model: fixed temperature and heat flux. These conditions must be applied carefully to achieve simulation accuracy as the model is simplified. The MB will be attached to a heat sink, so the top surface of Fig. 3.1(a) is fixed at 300 degrees Kelvin. When using a fixed temperature boundary condition to simulate a heat sink, the amount of heat generated in the system and the cooling capability of the heat sink must be considered. In this case, this assumption is legit because only a single unit is heated up. The total heat in the system is only a few milliwatts so the heat sink can easily extract the heat while maintaining the contact surface at the room temperature. However, in the simulations for finding the maximum allowable array size which will be presented and discussed in detailed in the later section, multiple heaters (rather than a single heater) dissipate

power at the same time. The amount of generated power may be challenging for a conventional heat sink or even exceed its cooling capability. The temperature at the contact surface may be raised above the ambient temperature. Assigning a higher fixed temperature to the contact surface is not suitable because the temperature is not uniformly distributed.

The rest of the five surfaces are assumed to transfer heat to the surrounding air. The outward heat flux per unit area is:

$$q_0 = h(T - T_{air}) \tag{3.9}$$

where h is the heat transfer coefficient, T_{air} is the air temperature, T is the temperature at the boundary. The air temperature is also taken to be 300 degrees Kelvin and the heat transfer coefficient is assumed to be 5 W/m^{2}K.

For the bottom surface, which belongs to the OC, this assumption is valid and matches the real system setup. This surface is in direct contact with air so the output light can emerge without blockage. On the contrary, the other four lateral surfaces are not actually surrounded by air as in the real system. For the geometry shown in Fig. 3.1(a), the length and width of the model are both 400 μm. It is at least an order of magnitude less than the dimension of the real system because the size of the actual OC is at least 1 cm by 1 cm. Thus, these boundaries are part of the OC and MB substrates in the real system. The actual outward heat fluxes at these surfaces are determined by the substrates' thermal conductivities and the temperature gradients based on the equation (3.5). However, the temperature gradients are unknown variables and using air convective heat flux as boundary condition is only an approximation which requires additional verification.

Furthermore, the total surface area of the model affects the simulation of the result. The total heat dissipated from these surfaces is the total area multiplied by the outward heat flux in equation (3.9). To dissipate the same amount of heat as is generated at the

heater, the temperature at the boundary with a smaller surface area must be higher than the temperature at the boundary with larger surface area as a higher temperature is needed to compensate for lower total area. Thus, the resulting temperature of the entire system may be higher than it should be if insufficient surface area is included in the simulation model. Ideally, the model geometry should be same as it actually is (i.e., having a geometry with a length scale on the order of 1 cm), which solves the problem of small surface area and false boundary conditions mentioned above at the same time. However, the high geometry aspect ratio due to the size of the phased array unit cell prevents us from doing this. Again, the size of unit cell is 15 μm by 15 μm and the thickness of the phase shifters and resistive heaters are both 100 nm.

The high aspect ratio of the model geometry requires large number of mesh points to arrive at a meaningful solution. The denser the mesh, the more computation resource is required. Also, the execution time of the simulation is longer. Therefore, the first step is to find an optimal size of the model geometry. Fig. 3.3(a) illustrates the process. The side length of the model geometry is gradually increased from L_1, 90 μm to L_2, 2070 um. These are 2 times and 46 times the length of a 3 x 3 array respectively. We ensure the number of mesh elements is dense enough to represent the system geometry at every step. A 2.3 V voltage is applied to the top surface of the rectangular column contacting the center unit cell of the 3 x 3 array. This condition corresponds to the maximum current flow through a single heater. The average phase shifter temperature and the minimum temperature at the bottom of the model geometry which is cooled by air convection are monitored. The simulation results in Fig. 3.3(c) and (d) summarize our analysis. When the side length is less than 500 um, both the average phase shifter temperature and minimum temperature at the bottom decrease rapidly as the side length increases. If the model size is too small, the simulation erroneously predicts a higher

temperature variation due to the boundary conditions artificially creating heat sinks close to the waveguide bodies.

One may underestimate the current requirement for full range phase shift and fail to design a driver circuit with proper current drive strength if the model size is too small. Above a side length of 500 μm, the average phase shifter temperature stabilizes at around 300 °C and the minimum temperature of the bottom surface stabilizes at room temperature, 27 °C and along the edge of the surface. It is worth mentioning that the temperature may not stabilize when the temperature gradients are too large. Furthermore, the amount of heat generated in the system is so large that the surfaces can't be cooled to room temperature by air convection. Hence, the entire system is above room temperature. A geometry that is same as the actual system must be used in the simulation for accurate result in this condition. Fortunately, the heat generated from a single heater is not large enough to significantly raise the temperature of the system as we move away from the waveguides into the substrate bulk.

The curves in Fig. 3.3(c) and (d) provide system designers with useful information about choosing the size of simulation model in different conditions. For simulating a single unit, a side length of 500 μm is enough. However, when currents flow through a small number of heaters at the same time for studying thermal crosstalk, the total generated heat will be more than that shown in Fig. 3.3. For simplicity, we choose to have a side length of 1050 μm, 30 times the size of a 3 x 3 array for the single pixel and the thermal crosstalk simulation. This allows room for us to study the heat generated by multiple heaters.

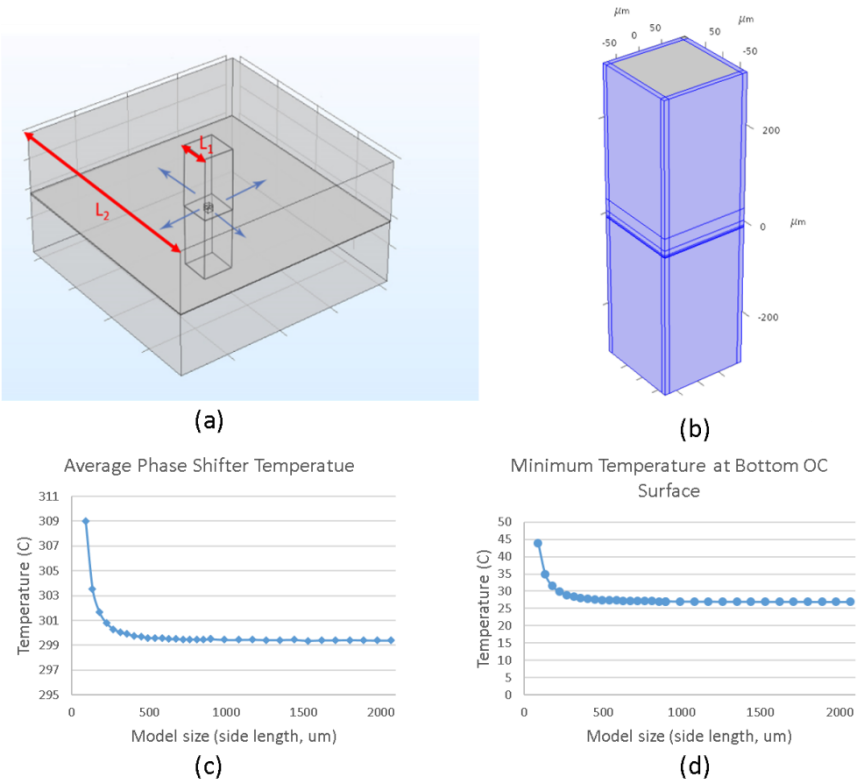


Fig. 3.3(a) Model geometry length sweep for determining optimal simulation size (b) model geometry using build-in infinite element at boundary, and the simulation result of geometry length sweep: (c) phase shifter average temperature and (d) minimum temperature at the bottom of the model

An alternative approach to solve the model size issue is to use the COMSOL built-in infinite elements. The model geometry is illustrated in Fig. 3.3(b) and the side length is only 180 µm. Thin layers highlighted in blue are created around the lateral boundaries and are assigned to infinite elements. The software stretches the coordinates in these layers using a scaling function with a parameter, t, from 0 to 1. The default scaling function, f(t), is defined as:

$$\mathrm{f(t)} = \frac{sw \cdot t}{sw \cdot (1-t) + p} p, \quad 0 \le t \le 1 \tag{3.10}$$

where sw is the scaled width which is the actual system size in our case and p is a user input which determines the shape of the scaling from the drawn geometry to the actual system size. If p is far less than sw, the scaling function increases slowly at the beginning and reaches sw abruptly as t is varied from 0 to 1.

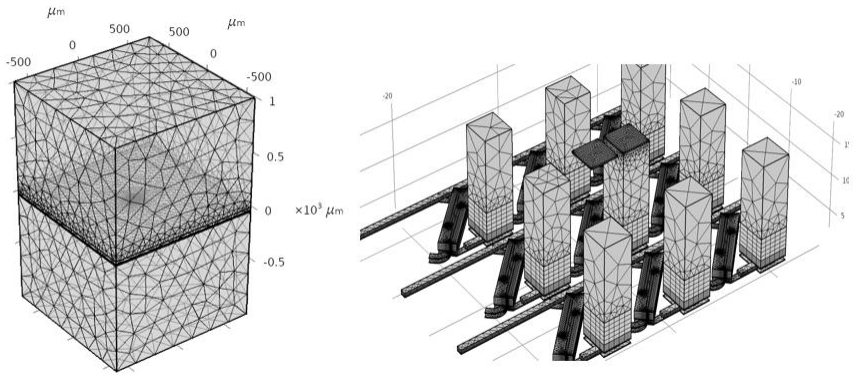
Using proper parameters, sw and p, the highlighted thin layers are scaled to represent the actual system size. The simulation result of the average phase shifter temperature is 299.87 °C which matches the results in Fig. 3.3(c). However, the meshes on the infinite element layers are highly dense so the total number of the mesh elements is comparable to the model with a side length of 1050 μm. The simulation accuracy is affected by the choices of scaling function and its parameters. The behavior of the scaling function must match the temperature distribution in the infinite elements layers for best results. For example, if the fixed temperature boundary condition at the heat sink contact surface is replaced by a convective heat flux boundary condition, the result of infinite element using the same scaling function and parameters deviates from the result of former model size analysis method. Furthermore, it does not provide the insights for creating a simulation model as shown in Fig. 3.3(c) and (d). Therefore, the former model simplification and approximation method is adopted.

For the maximum array size study in the later section, the model size must be same as the actual system for modelling cooling effect accurately. None of the above approach is used. The thin layer heater and phase shifter structures are simplified to keep the total number of mesh elements small so that the model can be fit in our computer's memory and the computation can be completed in a reasonable time.

The mesh result is shown in Fig. 3.4. The mesh partitions within the phased array thin layer are finer with both free triangular and swept mesh elements as described earlier. The overall mesh at the MB and optical OC substrate is coarser, with free tetrahedral elements. The total number of mesh elements is around 1,600,000.

Finally, according to equation (3.5), the governing material property of the temperature distribution is the thermal conductivity, k $(W/m \cdot K)$. Table 3.1 shows a list of the thermal conductivities of the materials used in the model at 300 degrees

Kelvin. The asterisk symbols following the values mean the parameters are temperature dependent.



(a) *(b)*

Fig. 3.4 The mesh grids of the finalized model (a) entire model (b) central 3 x 3 array

Table 3.1 Material thermal conductivities at 300 degrees Kelvin

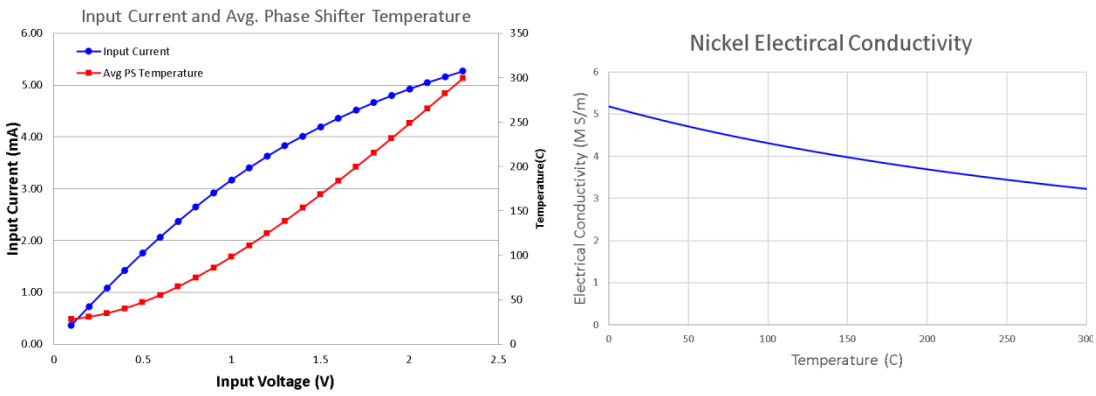
Material	Thermal Conductivity ($W/m \cdot K$)
Silicon (Si)	130
Silicon dioxide (SiO_2)	1.4
Silicon Nitride (Si_3N_4)	20
Air	0.0263 *
Nickel	90.58 *
Aluminum	238
Copper	400

3.3 Steady State Single Unit Simulation

With the side length of the geometry fixed at 1050 μm, a 0 to 2.3 V parameter sweep study for the input voltage (applied to the center unit cell) is performed. This simulation gives us a full understanding of the electrical and thermal interaction in the

system from a single phased array unit cell. The relation between the input voltage and the current passing through the resistive thin film heater and the input voltage and the average phase shifter temperature is shown in Fig. 3.5(a) as the blue and red curves.

The temperature at the phase shifter directly relates to the phase of the output beam, and we must ensure that the temperature variation is large enough for a full range (0 to 2π) phase shift. According to equation (2.5) with a 13 μm long and slow light slowdown factor of 6, the required temperature shift is 270 °C above the ambient temperature. The red curve in Fig. 3.5(a) shows that the average phase shifter temperature is around 300 °C at 2.3 V input voltage. Thus, the result validates our model showing that the input voltage indeed drives a current to the heater and raises the phase shifter temperature as designed.



(a) *(b)*

Fig. 3.5 (a) Input voltage vs input current flowing through the heater and the average phase shifter temperature (b) Temperature dependent electrical conductivity of nickel used in the model.

The blue curve in Fig. 3.5(a) shows the heater current. The value of heater current is obtained from a surface integration on the normal current density at the input surface. It is roughly linearly related to the input for voltages under 0.5 V as expected by Ohm's law. The current flattens out as the input voltage rises above 0.5 V. This is due to the temperature dependent electrical conductivity of the nickel and the parameter value in

the model from 0 °C to 300 °C is shown in Fig. 3.5(b). The electrical conductivity decreases as temperature rises. Therefore, when the input voltage increases, the temperature at the heater rises so its resistance increases and the current flowing through it is suppressed.

At 300 °C the electrical conductivity is about 3.2×10^6 Siemens per meter, only 60% of the value at 0 °C. One can find the current driving capability required for the driver from Fig. 3.5. In this case, a MB current driver must supply about 5.3 mA to the heater element to meet the goal of 2π phase shift. The normal current density across the heater is shown in Fig. 3.6.

The heater consists of five segments; the center segment is 13 μm long and 300 nm wide and the length and width of the other four segments are 9.1 μm and 300 nm for the two inner segments and 500 nm for the rest of two outer segments. Together, they provide appropriate resistance for heating the phase shifter to an average 300 °C temperature at 2.3 V input. The normal current density at the center section is about 1.75×10^{11} A/m^2. The current flowing through the heater is 5.25 mA by multiplying the normal current density with the cross-section area – 300 nm x 100 nm. The calculation result matches the value from surface integration on input surface mentioned above and verifies the correctness of the simulation result by electric current/charge conservation.

It should be noted that the simulated normal current density is so high that the device may have reliability issues because of electromigration. This does not affect the simulation results as the simulation mainly focuses on the relationship between the input power and temperature at the phase shifter. Instead, this result suggests that the 2π phase shift power must be reduced or the geometry of the heater should be changed, e.g., increasing the thickness to 1 μm or the width while keeping the resistance unchanged, to lower the current density at full load.

Next, we show the detailed temperature distribution resulting from a given input voltage. We are most interested in the impact of drive current on the phase change for light traversing the phase shifter, the heater and the MB. The highest temperature in the system is at the heater, and one must ensure the heaters and their connections can still function properly at the highest temperature. This temperature also dictates the range of available feedback signals and the precision requirement of the control circuit. We can also learn the heating efficiency of the system from the temperature difference between the phase shifter and its corresponding heater.

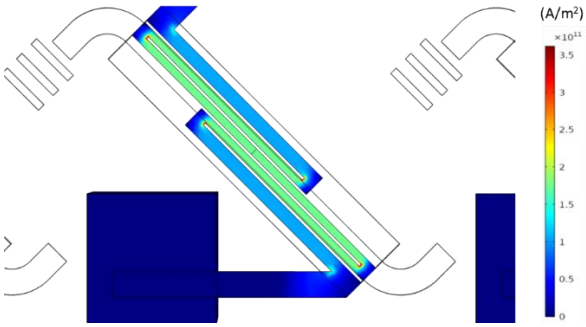
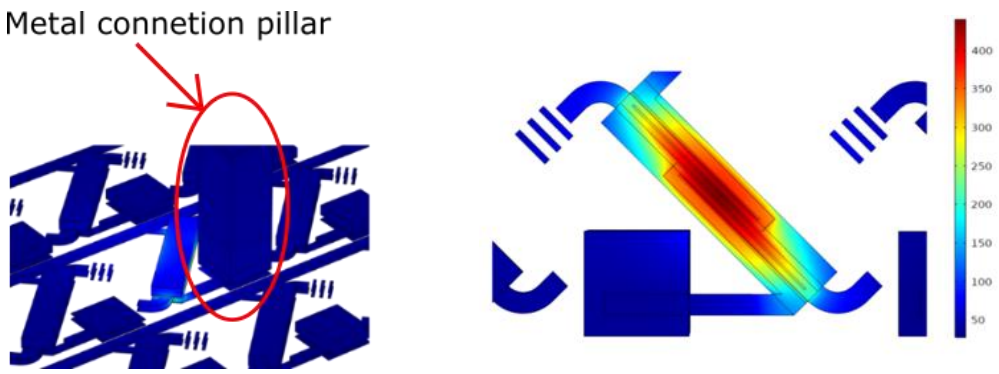


Fig. 3.6 Heater current density at 2.3 V input

The temperature distribution at the center array with a 2.3 V input applied to the center unit is shown in Fig. 3.7. A significant temperature rise happens at the center of the heater and the phase shifter/waveguide while the temperature of the other parts (such as BEOL metal in the circuit chip) and the heater ground connection are near room temperature. This is a near ideal performance situation. The heater is designed to be thin and narrow at the center serpentine part and be thick and wide at the contacts. An additional layer of a metal other than platinum is added to reduce resistance while keeping the cost low. Therefore, we ensure most input voltage drop occurs at the heater's center serpentine segment.

Figure 3.7(b) illustrates the successful heater design; the highest temperature is above 400°C at the center of the heater and the temperature decreases rapidly as we

move away from the center.



(a) *(b)*

Fig. 3.7 Center unit cell temperature distribution with 2.3 V input (a) side view. The rectanglar boxes are the air cladding of the waveguides. The center tall rectangle pillar is the metal connection between the OC and MB of the center pixel. The connection metal pillars of the rest of the pixels are hidden for clarity. The waveguides and heaters are blocked by the air claddings. The temperature at the center of the air cladding is raised slightly above the ambient temperature. The temperature at the center connection metal pillar is kept near the ambient temperature indicates the power loss due to the resistance is small. (b) bottom view focus on the heater.

Figure 3.8 gives a close look of the temperature distribution at the phase shifter/waveguide. The maximum temperature is about 400°C at the center and about 150 °C at both edges. The average temperature at the waveguide bending parts is about 65 °C which is 38 °C above the ambient temperature. Since the length of these two bending parts are shorter than the center phase shifter and do not contain slow-light grating structure, the introduced phase shift at these locations is negligible. The minimum temperature is 38 °C at the antenna part farthest from the heater center. The temperature distribution of the phase shifter is like the temperature distribution of the heater because the heater is right under the phase shifter.

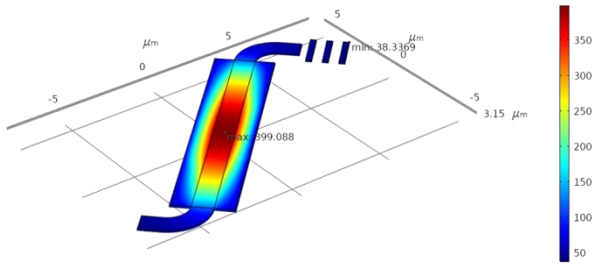
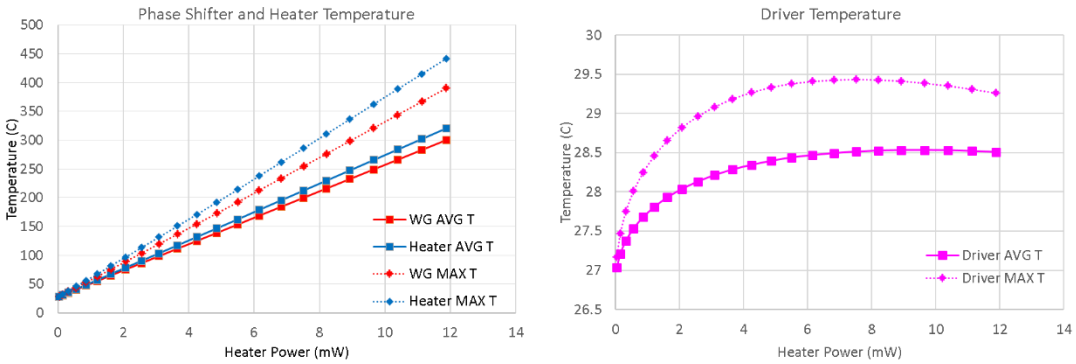


Fig. 3.8 Phase shifter/waveguide temperature distribution.

Figure 3.9 presents a simulation summary result for temperature at the heater, phase shifter and a 5 µm x 5 µm block next to the driver MOSFET in the MB with respect to the total power generated by the heater. As already shown in Fig. 3.7 and Fig. 3.8, the temperature along the heater and phase shifter is not uniformly distributed so we plot the maximum and average temperature.



(a) *(b)*

Fig. 3.9 Simulation result: heater power vs temperature (a) phase shifter and heater (b) driver circuit.

In Fig. 3.9(a), the average temperature at the phase shifter and at the heater are 300 °C and 320 °C, respectively, at full load (12 mW). These are linearly related to the heater power. A high degree of thermal coupling between the phase shifter and heater is achieved because the distance between them is only 500 nm. The highest temperature at the phase shifter and the heater are 400 °C and 440 °C and locate at the center of the devices as shown in Fig. 3.8.

The temperature at the corresponding driver circuit is shown in Fig. 3.9(b). Surprisingly, over a 400°C temperature changed in the OC, the temperature variation of the driver circuit inside the MB is less than 3 °C. This is mainly since silicon is an excellent thermal conductor. The heat generated from a single heater is so small that can be easily dissipated through the silicon substrate without raising the entire system temperature.

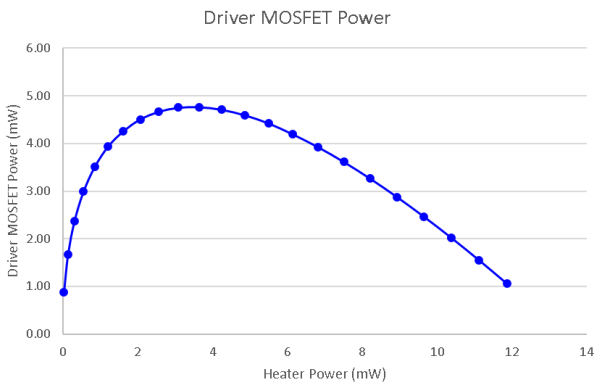


Fig. 3.10 Driver MOSFET power dissipation vs heater power

Furthermore, the temperature at the driver circuit doesn't increase monotonically with the heater power as the temperature at the phase shifter and the heater. The non-monotonicity property results from the secondary heat source, and unwanted power dissipated from the driver MOSFET. Fig. 3.10 illustrates the power dissipated from the driver MOSFET vs the power generated from the heater power. At light load, the current flowing through the heater is small so the driver MOSFET power is small, even though most of the supply voltage appears across the driver MOSFET. However, the driver MOSFET's power increases rapidly as the heater current increases while the driver MOSFET's V_{DS} is still large. In this region, the driver MOSFET dissipates more power than the heater does.

The maximum power dissipated by the driver MOSFET is around 4.8 mW and 3 mW by the heater at the same time. When the driver MOSFET's V_{DS} decreases faster

than the increase of the heater's current, the driver MOSFET's power starts to decrease and reaches a minimal value at full load when most of the supply voltage is across the heater. The driver MOSFET dissipates about 8 % of the heater's power, 1 mW, at full load. Combining Fig. 3.9(b) and Fig. 3.10, the maximum temperature at the driver circuit occurs when the heater power is 7 mW and the driver MOSFET's power is 3.8 mW. The situation is even worse when the thermal crosstalk between the neighboring driver MOSFETs is taken into account. A temperature sensor inside a driver circuit unit will register more heat from its adjacent driver MOSFETs due to the low thermal resistance of the MB substrate. As a result, this cannot faithfully reflect the temperature at its target phase shifter or heater. Therefore, it is impractical to implement a temperature sensor in the driver circuit. A temperature sensor must be located near its monitoring phase shifter or heater for high thermal coupling.

It is worth mentioning that there is an important design trade-off between the size of the driver MOSFET and the heating efficiency (power transfer ratio between the heater and the driver MOSFET). To maximize the heating efficiency and minimize the power dissipated by a driver MOSFET at 2π phase shift, a driver MOSFET's V_{DS} has to be kept as small as possible to ensure most supply voltage drops across a heater. The combination of the heater's resistance and the size of driver MOSFET determines the V_{DS_min} at full load. The heater's resistance at full load (R_{FL}), can be expressed as:

$$R_{FL} = \frac{(V_{DD} - V_{DS\_min})^2}{P_{FL}} \tag{3.11}$$

where V_{DD} is the power supply voltage, 2.5 V, and P_{FL} is the power required for 2π phase shift which we use 12 mW based on the simulation result. The heater's resistance is temperature dependent, and the value at room temperature can be approximated as:

$$R_{FL} = R_0\left(1 + \alpha \Delta T_{\text{heater\_2}\pi}\right) = R_0(1 + \alpha\beta P_{2\pi}) = R_0(1 + \beta' P_{2\pi}) \tag{3.12}$$

where α is the resistivity temperature coefficient of the heater and can be calculated

from the simulation result or the input material properties, and β is the slope of the heater power vs average heater temperature curve in Fig. 3.9(a). The input current at full load (I_{FL}) is:

$$I_{FL} = \frac{P_{FL}}{(V_{DD} - V_{DS\_min})} \tag{3.13}$$

The size of a driver MOSFET has to be large enough to supply I_{FL}. At full load, a driver MOSFET is fully on with gate-source voltage (V_{GS}) equal to V_{DD}, 2.5 V. Assuming a 0.7 V threshold voltage (V_{TH}), the V_{DS} has to be greater than 1.8 V to keep a driver MOSFET in its saturation region. Since the V_{DS_min} has to be small as possible to achieve high heating efficiency, a driver MOSFET must operate in deep triode region at full load. Using the MOSFET current equation in triode region, we have:

$$I_{FL} = K\frac{W}{L}((V_{GS} - V_{TH})V_{DS\_min} - \frac{1}{2}{V_{DS\_min}}^2) \tag{3.14}$$

where K is a fixed MOSFET parameter which is assumed to be 100 μA/V^2, W is the MOSFET's channel width and L is the MOSFET's channel length. Using the value mentioned above, the size of a MOSFET is:

$$\begin{aligned}\frac{W}{L} &= \frac{I_{FL}}{K\left(\left(V_{GS} - V_{DS\_min}\right)V_{DS\_min} - \frac{1}{2}{V_{DS\_min}}^2\right)} \\ &= \frac{120}{(2.5 - V_{DS\_min}) \cdot (1.8 \cdot V_{DS\_min} - \frac{1}{2}{V_{DS\_min}}^2)}\end{aligned} \tag{3.15}$$

and it is approximately inversely proportional to V_{DS_min} when V_{DS_min} is close to zero. We can evaluate the size of the driver MOSFET numerically with various values of V_{DS_min}.

For a given heating efficiency at 2π phase shift, one can determine the V_{DS_min} and use the equation (3.11) and (3.12) to find the heater's resistance at room temperature. The power dissipated by the heater and the driver MOSFET at any V_{DS} between V_{DD} and V_{DS_min} can be calculated as follows. Similar to equation (3.12), the heater

resistance at a given V_{DS} can be expressed as:

$$R = R_0(1 + \beta' P_{\text{Heater}}) = R_0\left(1 + \beta' \frac{(V_{DD}-V_{DS})^2}{R}\right) \tag{3.16}$$

Solving equation (3.16), leads to:

$$R = \frac{R_0+\sqrt{R_0{}^2+4R_0\beta'(V_{DD}-V_{DS})^2}}{2} \tag{3.17}$$

Thus, the heater power is:

$$P_{\text{Heater}} = \frac{(V_{DD}-V_{DS})^2}{R} = \frac{2(V_{DD}-V_{DS})^2}{R_0+\sqrt{R_0{}^2+4R_0\beta'(V_{DD}-V_{DS})^2}} \tag{3.18}$$

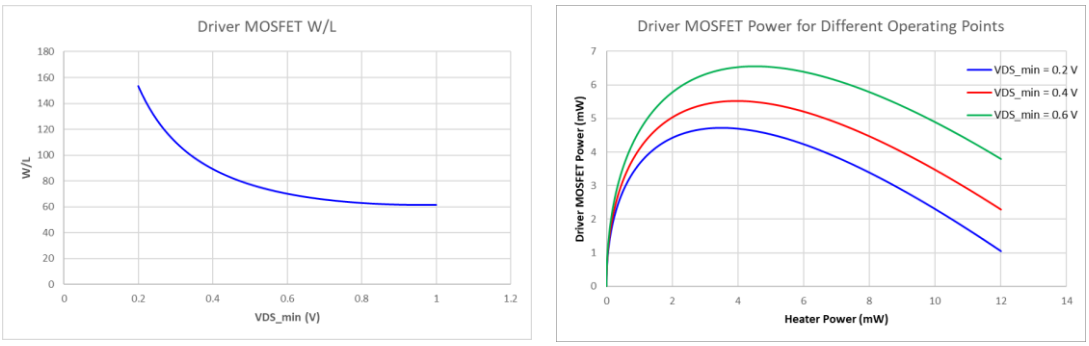
and the driver MOSFET power is:

$$P_{\text{Driver}} = \frac{(V_{DD}-V_{DS})}{R} V_{DS} = \frac{2(V_{DD}-V_{DS})V_{DS}}{R_0+\sqrt{R_0{}^2+4R_0\beta'(V_{DD}-V_{DS})^2}} \tag{3.19}$$

Figure 3.11(a) and (b) illustrate equation (3.15) for V_{DS_min} from 0.2 V to 1 V and the heater and the driver MOSFET power curve for V_{DS_MIN} equal to 0.2 V, 0.4 V and 0.6 V respectively. The required driver MOSFET W/L ratio increases steeply as the V_{DS_min} approaches zero. The blue curve, 0.2 V V_{DS_min}, in Fig. 3.11 (a), is similar to the power curve in Fig. 3.10 which means the analysis matches the simulation results. A small V_{DS_min} not only minimizes the driver MOSFET power at full load but also reduces this power loss for full range of phase shift. A driver MOSFET with larger W/L allows for a heater with higher resistance. For a fixed heater power, a heater with higher resistance leads to a larger voltage drop across the heater and smaller current. The power loss from the driver MOSFET connecting to this heater is smaller as the MOSFET is operated at less current and lower V_{DS}. Hence, a large driver MOSFET is preferred for achieving large array size by minimizing the power loss from the driver circuit.

Next, I investigated various structures to increase the thermal resistance in the NPA pixel to reduce the power required for a 2π phase shift. Based on the simulation result,

the power density of 12 mW on a 15 µm-by-15 µm pixel is about 5,300 W/cm^2. The cooling capability of most modern heat sinks is on the order of 100 W/cm^2 [107]. This means the surface area of a heat sink must be about 50 times greater than the size of an NPA. This is the case in our simulation as only one NPA pixel is powered and the surface of the cooling boundary is 1 mm^2. However, for a large size NPA, this high power density leads to gigantic (physically infeasible) heat sink. The concentrated high power from tightly packed pixels and the internal thermal resistance of a large heat sink results high temperature gradient which may drive the temperature at the control circuit above the safe operating limit. To realize a large size NPA, the power for a 2π phase shift must be further reduced.



(a) *(b)*

Fig. 3.11 (a) Driver MOSFET size W/L vs minimum V_{DS} (b) Driver MOSFET power dissipation vs heater power for different V_{DS}_min

One way to increase the thermal resistance connecting to the heater and the phase shifter is to remove part of the material (SiO_2) near these devices. The removed empty space is filled by air which acts as a thermal barrier, so the heat can be confined within the pixel. Fig. 3.12 illustrates the cross sections of various NPA pixel structures in this study.

Figure 3.12(a) shows the original structure. The top cladding of the phase shifter is air and a 500 nm thick SiO_2 is between the heater and the phase shifter. Heat can

travel in all directions toward the heat sink on top of the MB. The first thermal insulation layer is added by etching the SiO_2 along the two sides of the phase shifter vertically down to the heater as shown in Fig. 3.12(b). It eliminates the path for heat transferring upward directly. In Fig. 3.12(c), the thermal insulation is further enhanced by forming air trench into the OC substrate. It prevents heat from escaping in the lateral direction. Deep trenches result in high aspect ratio structures which may weaken the device's mechanical strength, complicate the OC fabrication processes, and reduce the yield. All of this would create a very expensive device.

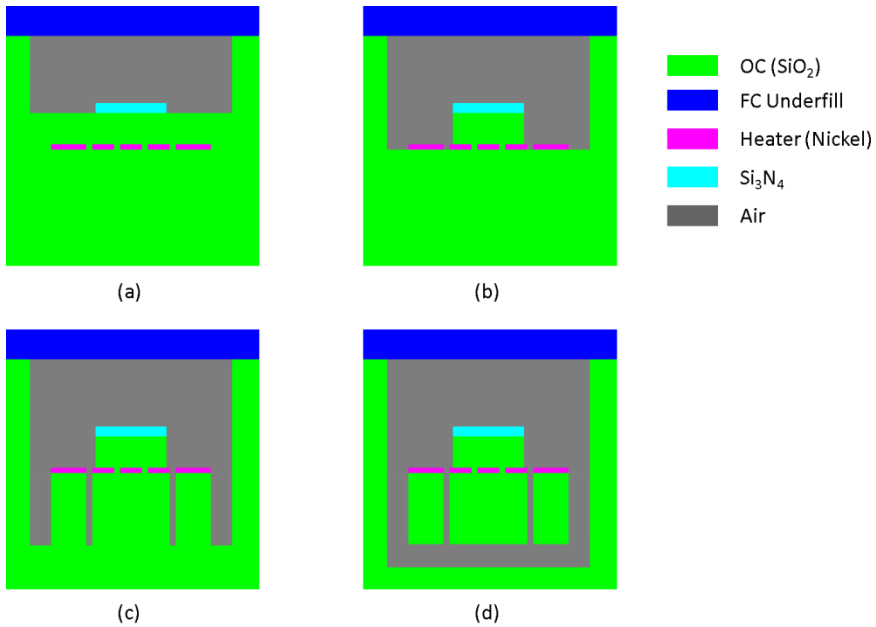


Fig. 3.12 NPA pixel cross section (a) original (b) SiO_2 removal between the heater plane and phase shifter plane (c) trench isolation (d) free standing

The trench depth in the study is limited to 4 μm so the ratio of the depth to the width of the SiO_2-heater-SiO_2-phase shifter stack is less than 5. Finally, a free-standing structure is possible [108]. The SiO_2-heater-SiO_2-phase shifter stack forms a free-standing thin film membrane and connects to the OC substrate in the direction normal to the Fig. 3.12(d). Heat can only propagate along the stack to the substrate and is

blocked in all the other direction. Again, the major disadvantage of this structure is the issue of low mechanical strength and complex fabrication processes.

Intuitively, increasing the thermal resistance between a device and a heat sink would increase the device's cool down time because the heat energy is well-confined and cannot escape to the heat sink easily. As a matter of fact, a commonly used figure of merit (FOM) for a phase shifter's performance is its 2π power requirement times its response time [78]. It indicates that there is a design trade-off between power and speed (as it appears above) by removing thermal leakage paths. Fortunately, unlike the optical communication applications for which ultra-high speed is essential, the respond time for NPA image applications only must be on the order of millisecond (30 to 120 frame rate per second). Special control technique can be used in LIDAR applications to speed up the scanning rate to offset the extra response time.

The results are shown in Fig. 3.13. The trench depths used in the simulation are 500 nm, 1 um, 2 um, 3 um and 4 um. The parameters for the free-standing structure are 2 μm/1 μm, 4 μm/1 μm, 4 μm/2 μm and 4 μm/3 μm where the first parameter is the trench depth, and the second parameter is the distance between the trench surface and the bottom of the free-standing stack. The difference between the two parameters is the height of the supporting SiO_2 under the heater. For simulation simplicity, the thickness of the heater is varied instead of changing the heater geometry to keep the average phase shifter temperature around 300 °C with 2.3 V voltage drop across the heater for all insulation configurations. The thickness is reduced as thermal insulation features are introduced so the heater provides less power at full load. Although it does not take into account the current capacity of the heater, the main purpose of this study is to learn how thermal insulation features in a NPA pixel lower the power requirement and one can apply the results to optimize the heater design based on any particular system

specification.

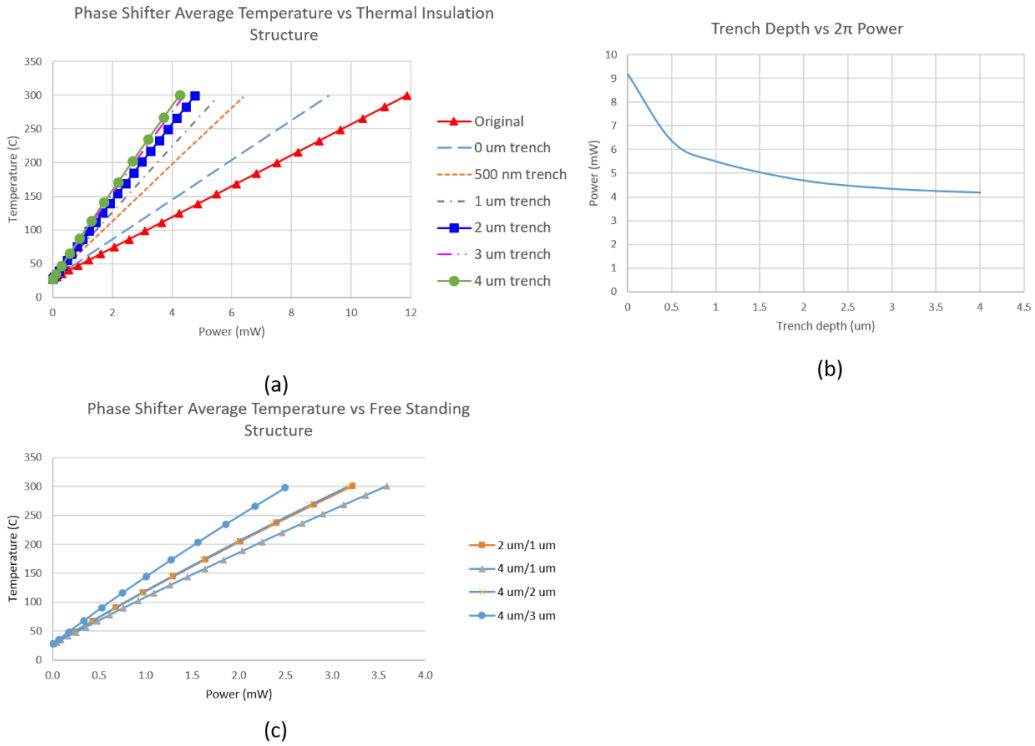


Fig. 3.13 Phase shifter power performance with various thermal insulation structures (a) air trench insulation (b) 2π phase power vs thermal insulation trench depth (c) free standing

The first step (SiO_2 etching) reduces the power requirement about 22%, from 11.76 mW to 9.16 mW. An additional 2.8 mW is saved with a 500 nm depth trench insulation. The 2π power continues to decrease as the trench depth increases but it begins to saturate at 3 μm depth as shown in Fig. 3.13(b). The 2π power reduction from 3 μm to 4 μm trench depth is only 0.16 mW. A high height-to-width center stack can be avoided because, as it provides negligible power saving. Free-standing structures further decrease the power requirement as expected. Although the differences of the required power are small, the 2π phase shift power of 2 μm/1 μm configuration is less than the power of 4 μm/1 μm but more than the power of 4 μm/2 μm and 4 μm/3 μm configurations. The trench depth and the height of the supporting SiO_2 together determine the power requirement. The cross section of the junction between the free-standing stack and the OC substrate through which the majority of the heat energy

transfers to the heat sink is proportional to the height of the supporting SiO_2. In summary, the lowest achievable 2π phase shift power of the etching thermal insulation method is about 2.5 mW. This is 21.2 % of the original simple structure requirement, i.e., about 80 % power saving.

The relation of the average temperature at a phase shifter and heater power with a 2 μm trench isolation depth structure and pixel pitch from 10 μm to 15μm is shown in Fig. 3.14(a). The length of the phase shifter is assumed to be 2 μm less than the pixel pitch, from 8 μm to 13 μm. The length is not directly proportional to the pixel pitch because not all devices' dimensions directly shrink with the pixel pitch; e.g. the widths of the waveguide/phase shifter are determined by the slow light Bragg grating design and the propagation mode. The result shows that for a same amount of heater power, the average temperature at a phase shifter is higher for smaller pixel size because heat energy is confined to a smaller volume. However, according to equation (2.5), a shorter phase shifter requires a higher temperature raise above the ambient for a 2π phase shift. The 2π phase shift power for different pixel sizes is presented in Fig. 3.14(b). The slow down factor is assumed to be a constant (6) not varying with pixel sizes. The power requirement increases as the pixel size reduces. The increase of the required temperature raise due to a shorter phase shifter dominates over the increase of power efficiency for smaller volume. Thus, higher power is needed for a NPA with larger FOV and finer beam width. It is possible to further reduce the 2π phase shift temperature with pixel size greater than 15 μm, because extra space allows a phase shifter to be bended 180° without significant propagation loss which at least double the effective length in equation (2.5).

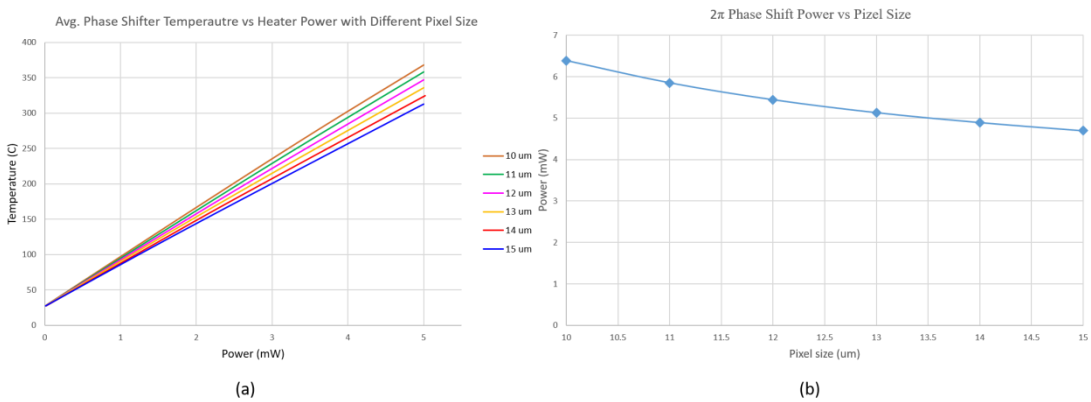


Fig. 3.14 Phase shifter power performance with various pixel size (phase shifter length) (a) power vs average temperature (b) 2π phase power vs pixel size (phase shifter length)

The above simulations are repeated for our LIDAR NPA. Since the wavelength used in a LIDAR NPA is 1550 nm, the core of the waveguide/phase shifter can be silicon instead of silicon nitride and SOI wafers are used to make the NPA. The cross section of a LIDAR NPA pixel is shown in Fig. 3.15(a) and (b). The major difference is the relative location of the heater and the phase shifter. The heater is on top of the phase shifter. The thickness of the top layer silicon of the SOI substrate is 220nm and is used as the core of the phase shifter. 2 μm buried oxide below the top layer is used as the lower cladding. A 500 nm thick SiO_2 is grown on top of the top layer silicon as upper cladding. Finally, the heater is deposited on the upper cladding. The distance between the heater and the phase shifter is same as the distance in the VR NPA and high thermal coupling is achieved.

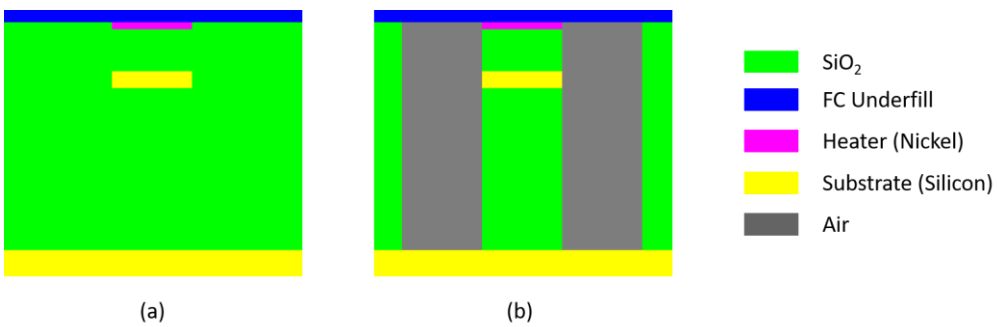


Fig. 3.15 LIDAR NPA pixel simulation model cross section (a) original (b) with trench thermal insulation

The thermo-optic coefficient of silicon is much higher than silicon nitride, and the slow down factor can be up to 10. Therefore, to reach a 2π phase shift, we need only elevate the phase shifter temperature 67 °C above the ambient temperature for a 13 μm phase shifter. This is a significant reduction compared to 270 °C in a VR NPA using silicon nitride. The results in Fig. 3.16 illustrate that the power requirement for a LIDAR NPA pixel without air insulation structure is about 6 mW. The power reduction is about 50 % comparing with the results of VR NPA. However, the expected power reduction is 75 % because the required temperature shift of our LIDAR NPA is only one-fourth of our VR NPA. This is mainly due to the high thermal conductivity of the silicon substrate and the thickness of the SiO_2 layer between the phase shifter and the silicon substrate to block the heat transfer is only 2 μm. The power required after etching down to the phase shifter plane and installing the 2 μm air trench thermal insulation structures is reduced from 6 mW to 5.5 mW and 4.2 mW respectively.

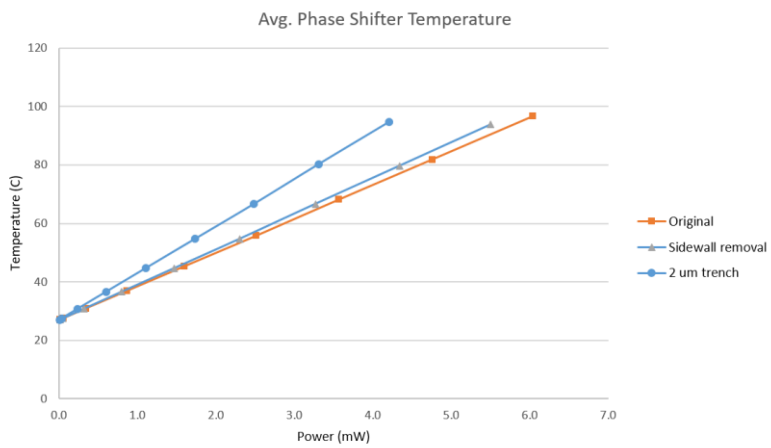


Fig. 3.16 LIDAR NPA pixel simulation result: average phase shifter temperature vs heater power with different pixel thermal insulation structures

To sum up, our LIDAR NPA consumes about half of the power (6 mW) that our VR NPA requires (11.76 mW) when no thermal insulation structure is present in the pixel. This is because of the higher thermo-optic coefficient of silicon and the use of slow light. When 2 μm thermal insulation air trenches are introduced, the 2π phase shift

power of our LIDAR NPA (4.2 mW) is only 0.56 mW less than the 2π phase shift power of our VR NPA (4.76 mW). This is due to the high thermal conductivity of the SOI substrate, thermal insulation structures work less efficiently, and high thermal leakage negates the advantages of using silicon as the phase shifter core.

3.4 Transient Response

Thermal transient simulation was performed based on pixels with a 2 μm trench structure. Fig. 3.17 shows the simulation results of two load conditions. Since the transient simulation requires more computational resource than the simulation discussed above, the simulation was done purely in the thermal domain.

Using our unique broken-loop feedback control method, this heater continuously switches between two modes mentioned above. No power is provided by the heater in the sensing mode and a near constant power is generated by the heater in the heating mode. Thus, a heat source of 100 kHz square wave with 90 % duty cycle was applied to the target heater directly. That is, the durations of the heating mode and sensing mode of each cycle are 9 μs and 1 μs.

At steady state, the driver output current to the heater is expected to repeat between two consecutive states as the amplitudes of the heater power in Fig. 3.17(a) and (c). Fig. 3.17(b) and (d) show the simulated maximum heater's temperature and the phase shifter's average temperature above the ambient for the heater output power conditions in Fig. 3.17(a) and (c). The temperature of the entire system is rest at ambient at time t = 0. For both load conditions, the phase shifter's temperature is stabilized within three cycles.

The phase shifter's temperature responses as a low-pass filter with respect to the heater power. Ripples exist as expected because the heater turns on and off. The amplitude of ripples is about one-third of the highest value. To minimize the ripple

amplitude, the sensing time must be kept small because of the small cooling time constant. For VR applications, the ripple does not lead to any performance issue because the frequency is far beyond 120 Hz framerate, the limit of human vision for most people.

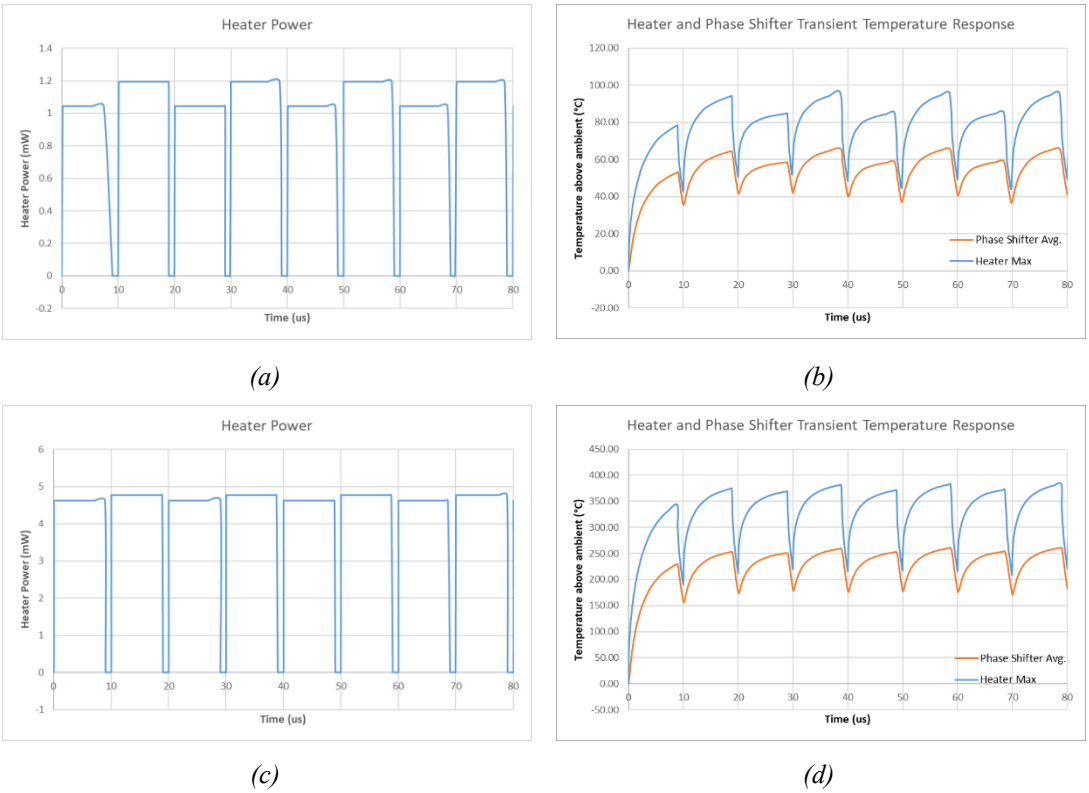


Fig. 3.17 NPA pixel thermal transient simulation result (a) heater power dissipation at light load, (b) results of (a), (c) heater power dissipation at high load, and (d) result of (c)

To sum up, the transient simulation result verifies that the broken-loop feedback driving mechanism is suitable for NPA pixels with 2 μm trench structure. The phase shifter's temperature can be raised successfully above the ambient at different levels and stabilized in a reasonable duration.

3.5 Thermal Crosstalk

In this section, the analysis of the simulation for inter-pixel thermal coupling is

presented. As explained earlier, to overcome the routing congestion in scalable 2D NPAs with independent phase control for all units, each pixel contains its own phase shifter and heater. A phase shifter and a heater may be surrounded by some local thermal insulation as introduced in the previous section. This has the potential for reducing the power needed by confining the heat energy. But due to the small pixel size, the heat can still leak to the adjacent pixels and inevitably raise their temperature. This proximity effect causes phase shifters to deviate from their ideal temperature and introduces phase errors. The purpose of this study is to quantify the thermal crosstalk and to use this information to explore the need of a feedback control mechanism and a proximity effect correction procedure for minimizing phase errors and preventing thermal runaway.

The array size in the simulation model is expanded from 3 x 3 to 5 x 5. The input voltage is not limited to the center unit and can be applied to multiple units simultaneously. The pixels are labeled by row and column indexes from 1 to 5. For example, the center pixel is referred to as pixel [3,3]. The average temperatures at each phase shifter in the array are measured. The applied voltages are either 0 or 2.3 volts, representing the off or fully on status of a current driver. From the single unit simulation results, the average temperature raise at a phase shifter is proportional to the power generated by a heater almost linearly. Thus, temperature for load conditions between off and fully on can be linearly interpolated. However, even with only 2 states, 25 array units lead to 2^{25} possible loading configurations. It is impossible to solve all the configurations so only certain representative cases are chosen.

First, pixels are turned on one at a time to see how these pixels affect the temperature at their adjacent pixels. The average temperature increases at phase shifters for pixels with 2 μm trench thermal barriers and without etching thermal insulation are illustrated in Table 3.2 and Table 3.3. The highlighted cells represent the pixels that are

turned on. Comparing Table 3.2 and Table 3.3, the temperature at the adjacent pixels is reduced by half when trench thermal barriers are presented in the pixels. Thus, it is essential to have the pixels surrounded by some form of thermal insulation structures. Not only does this reduces the power consumption, but it also suppresses the thermal crosstalk.

The temperature distribution is not symmetric with respect to the fully on pixel. For the same distances away from the fully on pixel, the temperature increase in pixels in the same row are higher than the temperature increases in pixels in the same column. For example, in Table 3.2(a), for one pixel pitch away from the fully on pixel [3,3] the temperature increases at pixels [3,2] and [3,4] are higher than the increases at [2,3] and [4,3]. This is because the pixel is not symmetric with respect to its center and the orientation of the heater placement. These differences reduce as one moves away from the fully on pixel. This is because the thermal coupling is reduced and the temperature variations caused by the asymmetric pixel geometry tend to be averaged out.

Comparing the results in Table 3.2, one can find that the temperature distributions centered on the fully on pixel are almost identical expect that pixels at the array boundary which are surrounded by fewer pixels have larger temperature increases. For example, the temperature increases at the boundary pixel [1,1] in Table 3.2(b) is higher than the increases at the pixel [2,2] in Table 3.2(a) which are both located one pixel left and one pixel up away from the fully on pixel of the respecting configurations. The difference is 0.486 °C which is about 30 % of the value of the pixel [2,2] in Table 3.2(a). On the other hand, the difference between the pixel [2,2] in Table 3.2(a) and the pixel [3,3] in Table 3.2(i) (which are in a same relative position to their fully on pixels as in the previous case) is only 0.016 °C. Similar patterns can be observed at other pixels in the table. The assumption of a uniform silicon dioxide layer, instead of the actual pixel

layer structure around the boundary pixels in the model, results in this error. Fortunately, the absolute error reduces as the distances from the fully on pixel increases. Therefore, combining the results of single fully on cell with a 5 x 5 array, a 7 x 7 estimated temperature distribution map with center cell being fully on can be obtained as shown in Table 3.4. Temperatures at cells outside the 7 x 7 map can be extrapolated if needed. This matrix is the impulse thermal response of a single pixel. The overall temperature distribution can be calculated by the convolution of the pixels input and this matrix if the thermal crosstalk response is linear.

Next, multiple pixels are turned on simultaneously. The results are compared with single pixel results to check whether the temperature distribution of multiple pixels can be approximated by the superposition of single pixel results. Table 3.5 shows the results of inner 3 x 3 pixels being all on with/without the center pixel and the estimations from superposition.

When multiple cells are turned on fully, each heater generates slightly less power than the power in the former simulations where only one cell is turned on at a time. This is because the simulation control input is voltage instead of power. With multiple cells being on, thermal crosstalk causes the heaters' resistance to go higher than the value when only single cell is on. Since the input voltage is still 2.3 V, the heaters' powers are reduced slightly. For example, the heater's powers of the pixel [3,3] are 4.66 mW and 4.7 mW in Table 3.5(c) and Table 3.2(a) respectively. To reduce estimation errors, the convolution kernel, Table 3.4, has to be scaled down by the ratio of the actual heater's power in the multiple unit simulations to the heater's power in the single unit simulations for each pixel. If the correction factors are neglected, the estimated temperature at pixel [3,3] for Table 3.5(c) is 298 °C which is 6.5 °C higher than the value in Table 3.5(d). Table 3.5 demonstrates a good agreement between the simulation

results and the estimations using Table 3.4. This proves the thermal crosstalk response is near linear with respect to the individual heater's power. It is unnecessarily to run simulations for all possible configurations. An impulse response matrix obtained by combining several simulations of single fully on cell at different locations is enough to calculate an approximated temperature distribution when the array size is small.

It is worth noting here that the impulse response matrix is generated under the assumption that the temperature at the heat sink contact surface is maintained at room temperature, 300 K. In other words, the array size is so small that the total heat dissipation from the system is much less than the heat sink's cooling capacity. As mentioned before, the surface temperature may be elevated above the room temperature to extract more power for a large NPA. Temperature gradients from the array's center to the edge must also exist for heat being extracted from the array's center. The existence of the temperature gradients is verified by the simulation results of array size. This is elaborated on in the next section. In additional to the proximity effect, a term representing the gradient is needed to model the thermal introduced phase error accurately. This will be discussed in more details in the phase error analysis section.

Table 3.2 Thermal crosstalk simulation results for 2 μm air trench: average temperatures at each phase shifters in a 5 x 5 array. Highlight cell represents the fully on pixel.

0.43	0.614	0.806	0.693	0.437
0.674	1.6	4.03	2.653	0.774
0.994	4.43	272.9	4.288	0.915
0.817	2.611	3.905	1.568	0.606
0.426	0.636	0.737	0.549	0.349

(a) Pixel [3,3] on

2.086	4.346	2.875	0.888	0.391
4.805	273.0	4.323	0.913	0.36
2.85	3.953	1.582	0.596	0.293
0.738	0.752	0.547	0.335	0.216
0.334	0.313	0.27	0.209	0.159

(b) Pixel [2,2] on

0.811	1.821	4.297	2.866	0.905
1.017	4.458	272.9	4.316	0.936
0.824	2.617	3.912	1.575	0.612
0.423	0.629	0.728	0.541	0.345
0.246	0.281	0.3	0.266	0.214

(c) Pixel [2,3] on

0.389	0.734	1.812	4.311	2.93
0.385	0.925	4.449	273.0	4.413
0.34	0.753	2.61	3.924	1.635
0.246	0.379	0.624	0.734	0.57
0.176	0.22	0.278	0.303	0.281

(d) Pixel [2,4] on

0.732	0.833	0.7	0.429	0.263
1.859	4.074	2.66	0.759	0.32
4.773	272.9	4.296	0.894	0.347
2.842	3.946	1.576	0.591	0.289
0.744	0.76	0.555	0.341	0.218

(e) Pixel [3,2] on

0.264	0.385	0.609	0.813	0.721
0.315	0.599	1.591	4.042	2.71
0.371	0.905	4.421	273.0	4.385
0.335	0.748	2.604	3.917	1.628
0.248	0.383	0.631	0.743	0.577

(f) Pixel [3,4] on

0.349	0.34	0.307	0.24	0.181
0.664	0.756	0.633	0.38	0.23
1.852	4.066	2.653	0.753	0.316
4.777	273.0	4.298	0.896	0.348
2.866	3.987	1.606	0.607	0.298

(g) Pixel [4,2] on

0.254	0.293	0.326	0.303	0.245
0.38	0.547	0.732	0.627	0.388
0.668	1.593	4.023	2.646	0.769
0.996	4.434	272.9	4.291	0.917
0.83	2.637	3.949	1.598	0.621

(h) Pixel [4,3] on

0.185	0.228	0.29	0.329	0.317
0.231	0.334	0.541	0.738	0.654
0.311	0.593	1.584	4.034	2.704
0.372	0.907	4.424	273.0	4.388
0.343	0.762	2.631	3.96	1.656

(i) Pixel [4,4] on

Table 3.3 Thermal crosstalk simulation results for pixel without thermal insulation structure: average temperatures at each phase shifters in a 5 x 5 array. Highlight cell represents the fully on pixel.

0.8286	1.1632	1.4795	1.2515	0.8181
1.3069	3.1224	7.5989	4.47	1.4076
1.9164	8.9719	272.44	8.7072	1.7797
1.4758	4.3875	7.4021	3.0777	1.185
0.7901	1.1442	1.3536	1.0447	0.6857

(a) Pixel [3,3] on

4.0778	8.2331	4.9078	1.6231	0.7554
9.7097	272.66	8.7614	1.7594	0.7088
4.8406	7.4929	3.1017	1.1611	0.5832
1.3281	1.3795	1.0404	0.6596	0.4373
0.6341	0.6027	0.5311	0.4216	0.3298

(b) Pixel [2,2] on

1.5719	3.5636	8.1354	4.8948	1.6639
1.9581	9.0225	272.49	8.7575	1.8167
1.4877	4.3993	7.4141	3.0892	1.195
0.7852	1.1328	1.3379	1.0315	0.6788
0.4774	0.543	0.5812	0.5253	0.432

(c) Pixel [2,3] on

0.7621	1.4261	3.545	8.1688	5.0366
0.7513	1.7765	8.9997	272.7	8.9775
0.6501	1.3573	4.3864	7.4414	3.2165
0.4793	0.7072	1.1239	1.35	1.0882
0.3537	0.4324	0.5381	0.5876	0.5536

(d) Pixel [2,4] on

1.376	1.5296	1.2647	0.8023	0.5148
3.6268	7.6849	4.4816	1.3736	0.6209
9.6519	272.62	8.7184	1.7275	0.687
4.8273	7.4824	3.0945	1.1546	0.5774
1.3397	1.3952	1.0555	0.6697	0.4424

(e) Pixel [3,2] on

0.5228	0.75	1.1551	1.4949	1.3088
0.623	1.1652	3.1056	7.6255	4.5951
0.7261	1.7422	8.9556	272.65	8.9268
0.6427	1.3497	4.3788	7.4314	3.2052
0.4825	0.7155	1.138	1.3669	1.1006

(f) Pixel [3,4] on

0.6694	0.6533	0.5901	0.4714	0.3646
1.2466	1.3818	1.1367	0.7094	0.4530
3.6133	7.6705	4.4677	1.3625	0.6131
9.658	272.62	8.7171	1.7274	0.6878
4.8709	7.5535	3.1438	1.1792	0.5906

(g) Pixel [4,2] on

0.5014	0.5737	0.6276	0.5825	0.4801
0.7370	1.0361	1.339	1.1272	0.7265
1.2971	3.1108	7.5863	4.4586	1.3986
1.9206	8.9777	272.45	8.7133	1.7835
1.4973	4.4344	7.4789	3.131	1.2101

(h) Pixel [4,3] on

0.3738	0.4565	0.5684	0.6352	0.6106
0.4618	0.6553	1.0269	1.3505	1.1802
0.6150	1.1546	3.092	7.6113	4.5825
0.7276	1.744	8.955	272.66	8.9311
0.6547	1.3724	4.424	7.5055	3.254

(i) Pixel [4,4] on

Table 3.4 Estimated 7 x 7 temperature distribution with the center pixel being fully on (impulse response)

0.1576	0.2116	0.268	0.299	0.2778	0.2193	0.1748
0.2138	0.334	0.541	0.732	0.633	0.38	0.2446
0.2916	0.599	1.6	4.03	2.653	0.759	0.3388
0.3447	0.905	4.43	272.87	4.288	0.894	0.3658
0.3137	0.753	2.611	3.905	1.568	0.591	0.3073
0.2292	0.379	0.629	0.728	0.541	0.335	0.2292
0.1794	0.2390	0.3053	0.3242	0.2888	0.2272	0.1750

Table 3.5 Thermal crosstalk simulation results for multiple pixels being turned on simultaneously, inner 3x3 on except the center (a) simulation and (b) estimated by convolution; inner 3x3 all on (c) simulation and (d) estimated by convolution

5.023	8.8776	11.063	10.046	5.8873
9.561	280.45	282.57	280.54	9.9029
11.85	282.54	24.722	282.54	10.935
10.604	280.51	282.42	280.22	8.6163
5.7096	9.2122	10.079	7.9293	3.9782

(a)

3.8363	7.9777	10.174	9.2372	5.2381
8.5584	281.29	283.76	281.45	9.6819
10.811	283.68	24.774	283.73	10.742
9.7244	281.4	283.69	281.08	8.4294
5.1972	9.1322	10.047	7.8736	3.8813

(b)

5.4314	9.4564	11.818	10.691	6.2968
10.205	281.5	285.24	282.31	10.635
12.795	285.28	289.36	285.19	11.807
11.369	282.23	284.99	281.24	9.1887
6.108	9.8035	10.769	8.4463	4.3097

(c)

4.1536	8.4887	10.863	9.8319	5.5961
9.1240	282.47	286.77	283.46	10.397
11.663	286.78	291.48	286.72	11.584
10.433	283.35	286.58	282.23	8.9876
5.5543	9.7230	10.732	8.3848	4.1994

(d)

3.6 NPA Scalability Analysis

The above simulations focus on the performance of a local single pixel. Here, we shift the focus to the large-scale system thermal response. Because our thermo-optic NPA is power hungry and the NPA (OC) and the control driver circuit (MB) are tightly integrated, the local power density can be one order of magnitude higher than the cooling capability of the most commonly used modern heat sink. Even though several high-performance cooling techniques have been reported, e.g. cooling device with an average power density of 1.16 kW/cm^2 [109] and 1.25 kW/cm^2 [110], the cost of these devices is too high to use in commercial products; the temperature at the control driver circuit is greatly affected by the total power consumption of the OC.

For reliability and performance of the driver circuits, the maximum temperature at the MB must be kept below 125 °C. This restricts the size of a NPA because a larger

phased array results higher congregated heat energy. That stresses the heat sink and creates a high temperature gradient across the entire system. A set of thermal simulations for characterizing the function between the maximum temperature at the MB and the array size is essential and provides insights for system designers.

The FEM geometries used in the single pixel thermo-electric and small array thermal crosstalk studies are not suitable for system level simulation as using a detailed pixel structure at system level makes the total number of mesh elements prohibitive large (The 5 x 5 array geometry used in the thermal crosstalk study is merely fitted in our workstation with 64 GB memory and the required memory grows quadratically with the array size). Thus, a system level model geometry which replaces the exact pixel structure with simple shapes is crucial to conduct the study with limited computation resource.

Figure 3.18 demonstrates a series of simplification process with increasing levels of abstraction. Same set of operating conditions are applied to all models and the maximum temperature at the MB in each model are compared to verify the validity of the process. In Fig. 3.18(b), the phase shifters, flip-chip bonds and interconnections are omitted comparing to the original model in Fig. 3.18(a). Each zig-zag heater is replaced by a 13 μm x 1 μm rectangle. The current driver remains unchanged - a 5 μm x 5 μm square. The layers between the heaters and current drivers are hidden in the figure. The simulation of the simplified model is purely in the thermal domain. No voltage or current boundary condition is applied to the model. Instead, uniform power sources with values corresponding to the loading condition are applied to the heater and the current driver blocks. In Fig. 3.18(c) and (d), the 5 μm x 5 μm current driver and 13 μm x 1 μm heater blocks are further removed and uniform average power density conditions are applied to each whole NPA pixel and control circuit blocks.

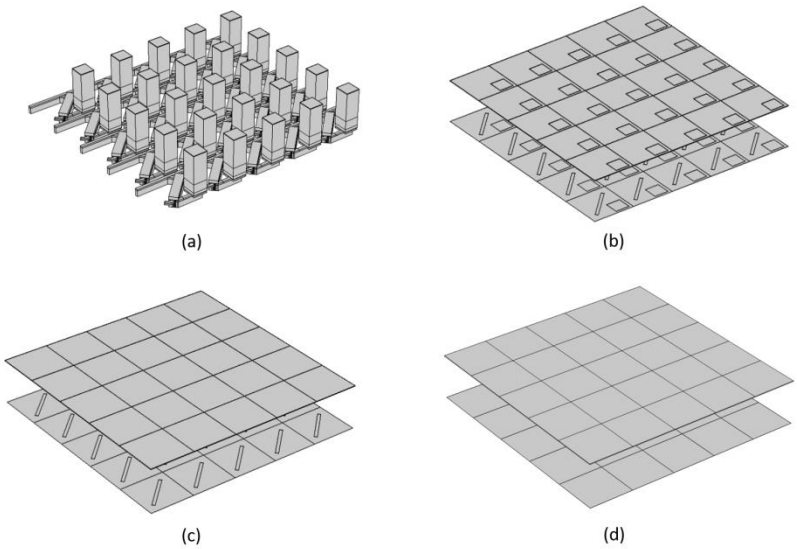


Fig. 3.18 Simplified 5 x 5 simulation model. The model is gradually step-by-step simplified from (a) to (d).

The simulation is performed by turning more pixels fully on sequentially and measured the maximum temperature at the MB. The fully on power of the heater is 4.8 mW based on the result of pixel with 2 μm trench thermal insulation. Fig. 3.19 summarizes the simulation results. The x-axis represents the number of the turned-on pixel in the simulation. The maximum temperature at the MB of the simplified models is less than the original model in all cases because the metal interconnects and FC bonds between the heaters and current drivers are omitted. They provide high thermal conductance paths for transferring heat energy from the heaters to the MB directly. Nonetheless, the error is small, and the simplification is valid for two main reasons. First, the temperature distribution at the MB is dominant by the heat generated locally from the current drivers. Neglecting detailed NPA pixel structure and interconnection in the simulation model does not affect the MB much. Second, most heat energy from the OC still travels upward and raises the temperature at the MB, because the heat sink is attached to the MB's top surface. The maximum temperature is slightly reduced because the energy spreads laterally as it goes upward. It results a less concentrated but

more uniform temperature profile at the MB.

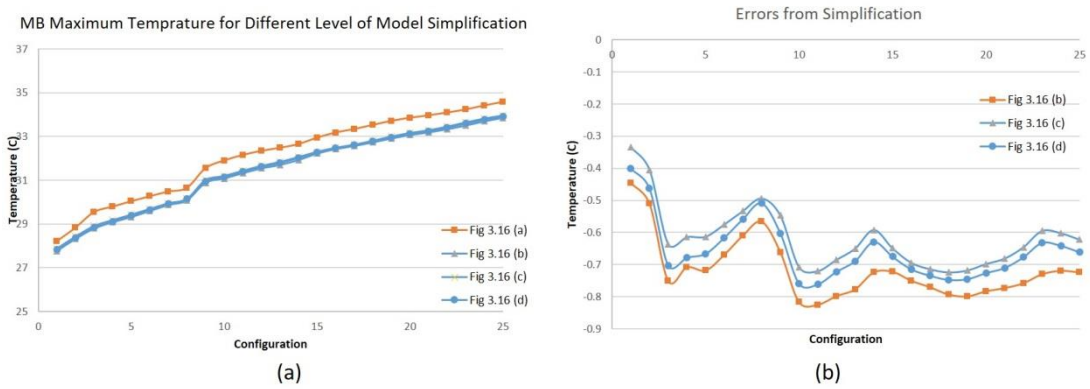


Fig. 3.19 Simplified 5 x 5 model simulation result, (a) MB maximum temperature (b) difference of the maximum MB temperature between the original model and the simplified models.

The errors due to the model simplification are mostly from step a to step b. The resulting curves of step b to step d in Fig. 3.19 are almost overlapping, because each step only modifies the model of the previous step slightly. The error of the step d is about 0.4 °C when only one pixel is on. The error increases as more pixels are turned on at the beginning but settles between 0.6 °C and 0.8 °C when most pixels in the 5 x 5 array are on. As more pixels are on, the lateral heat spread in the MB from each heater due to the simplification process are overlapped. This compensates the indirect heat transfer effect between the OC and the MB because of not including the metal connections in the model and reduces the errors. The result shows that for a large size NPA, the error of the maximum MB's temperature predicted by the model in step d, using uniform blocks to model heaters and current drivers does not exceed 1 °C.

The model geometry and the heaters' output power are demonstrated in Fig. 3.20. A metal plate (Al) which is larger than the NPA chip is attached to the MB. This is an indirect cooling method. The metal plate spreads the heat flux and provides a larger surface area where the coolant touches and takes away the heat energy. A convective heat flux boundary condition is applied to the top surface with the coefficient being 1

W/cm^{2}K or 10 W/cm^{2}K, typical value for force liquid convective and two phase convective boiling cooling techniques [107]. In Fig. 3.18(d), a simplification method is used to instantiate the pixels and current drivers in the array.

Fig. 3.20 Large size NPA (a) model geometry (b) heaters' output power

To simulate a real case, we use the butterfly picture in Fig. 3.20(b) to assign heat sources to the heaters and current drivers. We don't power all pixels fully on because it corresponds to zero phase shifts among all the output beams. The picture is resized to match the size of the simulated NPA and inverse Fourier transform is performed to obtain the output phase distribution of the NPA. The phase shifts are linearly converted to the heaters' output power distribution as shown in Fig. 3.20(b) and the current drivers' output power distribution is calculated using equation (3.18) and (3.19). Again, the maximum temperature at the MB is measured.

To further reduce the model complexity and speed up the simulation, the pixels away from the center are merged so the total number of the elements in the model is reduced. This method is validated with the process in Fig. 3.21. The simulation starts with a 64 x 64 array. The power sources assigned to the current driver and the temperature distribution result are shown in the left-most column in Fig. 3.21. Next, the center 32 x 32 pixels are unchanged, and for the remaining pixels, every 4 pixel group (a 2 x 2 array) forms a larger equivalent pixel. The average power density of the original pixels is assigned to the equivalent pixel. The number of the pixels can be reduced

further by having fewer pixels at the center and creating a larger equivalent pixel. For example, by having a 16 x 16 array at the center, we can have two layers around the center. The first layer consists of 2 x 2 equivalent pixels and the second layer consists of 4 x 4 equivalent pixels as shown in the center column of Fig. 3.21. Fig. 3.21 also shows that the average power dissipated by the current driver is around 1.2 mW.

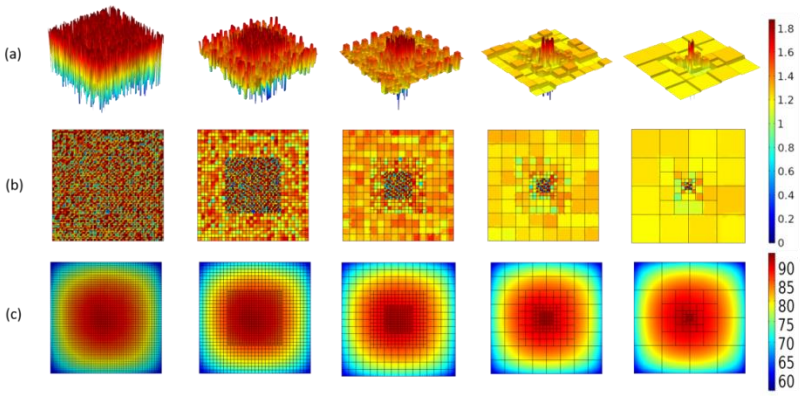


Fig. 3.21 Model simplification by merging pixels away from the center (a) and (b) applied heat sources at the current driver array (c) temperature distribution at the current driver array

The simulation results show that the temperature distribution patterns of each case is similar, and the maximum temperature is about 91.7 °C for all cases. For the rest of the simulation, the center array is fixed at 8 x 8 as this is the minimum size of the following simulation. The number of the equivalent pixel layer increases as the array expands.

The array size in the simulation starts from 8 x 8 (64 pixels) to 128 x 128 (16,384 pixels). The other two simulations parameters are the convective heat flux coefficient at the heat sink surface, 1 W/cm^{2}K or 10 W/cm^{2}K as mentioned earlier, and the thickness of the MB, 300 µm or 25 µm. To increase the cooling efficiency, we assume the MB chip is thinned down to 25 µm so the internal thermal resistance is reduced. Fig. 3.22 shows that thinning down the wafer reduces the MB's maximum temperature effectively when the array is small. However, one can observe that the slopes of the curve with a same heat flux coefficient, but different MB thickness are about the same

when the array size is big. The curves of 1 W/cm^{2}K with the MB thinned, and 10 W/cm^{2}K without the MB being thinned crossover around 7000 total pixels. When the array is large, to lower the MB's temperature effectively, the cooling capability of the heater must be increased.

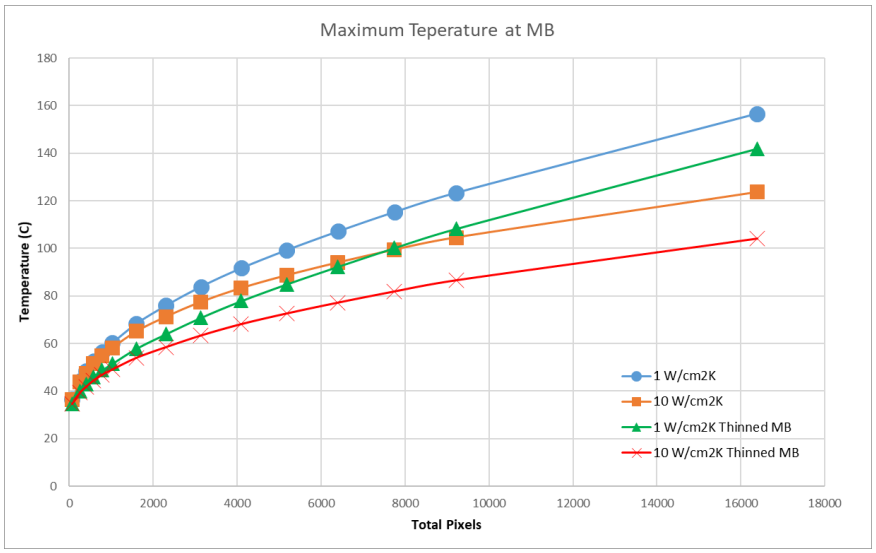


Fig. 3.22 Simulation result of maximum temperature at MB vs the number of total pixels with different assumed cooling schemes

In summary, to keep the MB's temperature below 125 °C and prevent the circuit from being damaged, the achievable total pixel number is about 9,000 with a 1 W/cm^{2}K heat sink and exceeds 128 x 128 (16,384) with a 10 W/ cm^{2}K and thinned MB chip. By incorporating slow light mechanism and thermal barriers around the heaters, the power required for our TO-based NPA is reduced significantly and the array size with a moderate heat sink can be about 9 times greater than the true 2D NPA reported in [43]. Larger array sizes are possible with advanced cooling technique which provides direct heat extraction from the MB.

3.7 Thermocouple Design

To achieve accurate phase control, I designed an integrated heater and

thermocouple element to measure the local pixel temperature without any extra connection between the OC and the MB. The heater consists of two segments of metal, one gold and one nickel. A thin film gold/nickel thermocouple with 10 μV/K Seebeck coefficient has been reported [76]. However, the Seebeck coefficient may vary with the thickness of the metal and the junction size. The design of a thermocouple sample is aimed at characterization of the Seebeck coefficient of the thermocouple in our NPA pixel using the same metal thickness and a similar junction size. This allows us to estimate the range of the feedback signal for driver circuit design. The simulation result is presented here.

Figure 3.23 shows the top view of our thermocouple testing device. The thin line at the center is the thermocouple sample and the width is 1 μm. A thickness of 10 μm SiO_2 is grown beneath the thermocouples and the heaters to minimize the heat loss to the substrate. The temperature will be measured directly by an IR camera. Because the focus length of the IR camera's microscopic lens is only 4 cm, the temperature at the testing device must be kept low enough to protect the lens. This limits the output voltage we can measure.

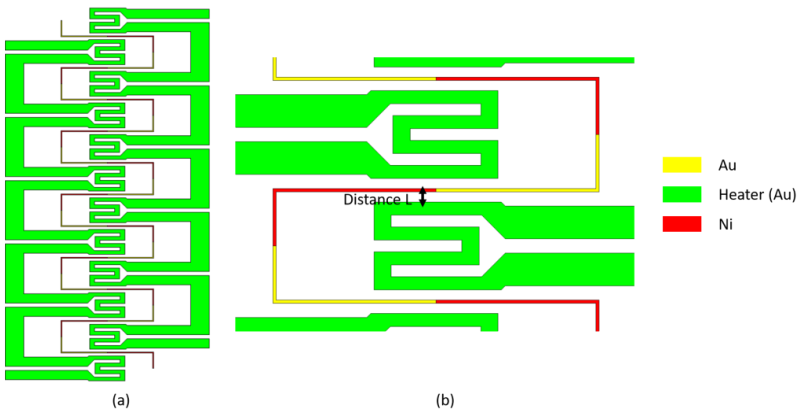


Fig. 3.23 Design of a thermocouple sample for Seebeck coefficient measurement (a) whole thermocouple (b) close-up view at the center

To reduce the measurement error due to the low level output signal, our testing

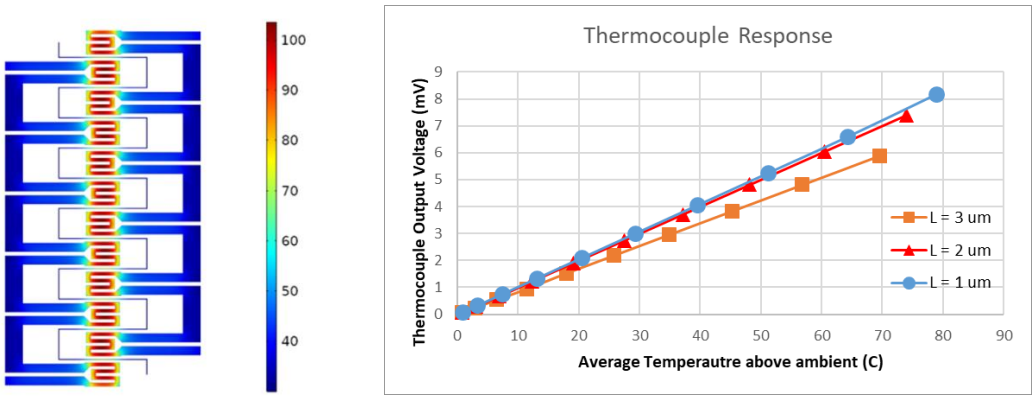
device consists of 11 thermocouples in series. The output voltage is the sum of the 11 thermocouples' response. The heaters are placed next to the thermocouple. The width of the heater segments that lie at the center line is reduced to ensure most voltage drop occurs at these parts. Voltage is applied to the heaters to increase the temperature at the thermocouple junctions along the center line above the ambient temperature. The voltage across the entire thermocouple array is recorded.

A major challenge regarding the temperature measurement is the spatial resolution of the IR camera which is 20 μm/pixel with a microscopic lens so the actual temperature at the junction cannot be measured directly. It is worth mentioning that resolving the temperature of the fine features of our thermocouple sample is beyond the limit of the resolution of any IR camera. To overcome this issue, we control the distance between the thermocouple junction and the heater (the parameter L Fig. 3.23(b)) so the junction temperature matches the average temperature of a wider area surrounding the junction. In the simulation, we evaluate the average temperature of the 20 μm x 300 μm rectangle area located at the center covering those junctions symmetrically. This corresponds to 15 pixels of the IR camera and we use the reading of these 15 pixels to estimate the junction temperature.

The simulation result of the temperature distribution with 5 V being applied to the heaters is shown in Fig. 3.24(a). Fig. 3.24(b) shows the simulation thermocouple output voltage vs the average temperature above ambient temperature with different values of the distance L. The voltage applied to the heaters was swept from 0 to 5 V.

From Fig. 3.24(a), the highest temperature occurs at the heater center segments and the temperature of each segment and for each thermocouple junction is about the same. Single thermocouple junction response can be calculated by dividing the output voltage by 11. The slope of the curves in Fig. 3.24(b) will corresponded to the extracted

Seebeck coefficient. The extracted Seebeck coefficients are 9.4 μV/K, 9.1 μV/K and 7.7 μV/K for L being 1 μm, 2 μm and 3 μm respectively. The results show that the distance L affects the extracted result seriously. The actual Seebeck coefficient used in the model is 10 μV/K. The extracted Seebeck coefficient approaches the ideal value as the L, distance between the thermocouple junctions and the heaters, reduces. When the distance L is too far, the actual junction temperature is much lower than the measured average temperature. As the result, it is important to ensure that the junction temperature is close to the average temperature. The sample measurement result will be presented in the next chapter



(a) *(b)*

Fig. 3.24 Thermocouple sample simulation result (a) temperature distribution (b) Thermocouple output voltage vs average temperature above ambient of a 20 μm x 300 μm area at the center

Chapter 4: NPA Driver Circuit Design, Simulation and Layout

In this chapter, I present the detailed analysis of our NPA driver circuit design. The driver circuit realizes a non-continuous feedback mechanism to regulate the phase shifter temperature at pixel level resolution. The key points of this chapter are:

- Full disclosure of the driver circuit system design and the broken-loop feedback operation overview.
- Description of two distinct realizations (analog and mixed-signal) of the broken-loop feedback methodology. The pure analog approach meets the circuit area constraint: average area is 15 μm x 15 μm per pixel, as would be suitable for VR applications. The circuit area of the mixed-signal approach slightly exceeds the established area constraint but is useful for LIDAR applications. These do not require independent phase control for every pixel.
- Design, constraints, layout challenges of meeting space requirement, and post-layout simulation results of individual circuit components in the driver system.
- System-level circuit layout and post-layout simulation results
- A 2-by-2 driver unit to demonstrate the compact design that fits in the circuit area design constraint.

4.1 Driver Circuit Overview

Figure 4.1 shows the simplified diagram of our driver circuit. The most important feature of this circuit is that the system does not require an extra connection between the NPA pixel and its driver circuit for temperature (phase) regulation. This is realized by using a novel integrated thin film heater/thermocouple device. This brings two major advantages. First, the size of a NPA pixel can be small because no extra temperature sensing device is needed. Second, it reduces the cost and the area constraint of the

heterogeneous integration between the OC and the MB as the number of the connections is minimized.

The resistor in the upper right of the diagram is the integrated thin film heater/thermocouple device. The dot at the center of the resistor indicates the presence of the thermocouple junction. The driver circuit consists of four stages: sense amplifier, comparator, gate driver and current source NMOS. The circuit cycles in two states: heating and temperature measuring. In the heating state, the switch S_1 is open to break the connection between the input of sense amplifier and the heater/thermocouple. The gate driver circuit uses the measuring result from the previous cycle to setup the output voltage to the gate of current source NMOS. In the temperature measuring state, the gate driver pulls its output voltage to ground to turn off the current source NMOS completely. The switch S_1 is closed and the input voltage of the sense amplifier is V_{DD} minus the thermocouple's output. The comparator compares the amplified thermocouple's output with an input control signal. The gate driver uses the digital output of the comparator to increase or decrease the gate voltage in the next heating cycle.

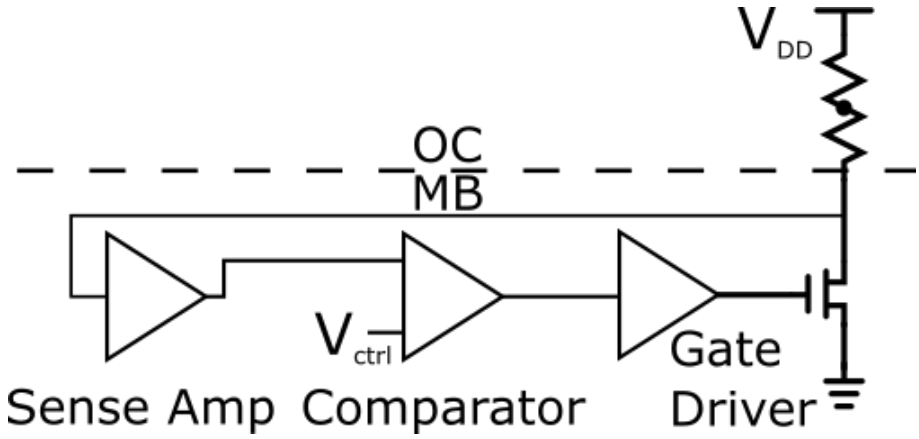


Fig. 4.1 Simplified driver circuit diagram

Even the circuit in Fig. 4.1 is simple at the first glance, it is difficult to implement the circuit directly in an area of 15 μm x 15 μm for matching the size of NPA pixel.

Thus, the driver circuit in Fig. 4.1 is modified into two versions for LIDAR NPAs and VR NPAs respectively as shown in Fig. 4.2(a) and (b). The idea is to share some components among a few neighboring units, so the "average" size of a driver circuit unit is below 15 μm x 15 μm.

For beam steering in LIDAR application, independent phase control for all pixels is not required. The relative phase differences between two adjacent pixels in the same row/column do not vary. If the light travels through all the phase shifters in the same row/column successively and is coupled to the antennas after each phase shifter, all these phase shifters introduce same amount of phase delay. Therefore, pixels in a same row are split into several groups equally. The sense amplifier, comparator and gate driver are share among the pixels in the same group. As in Fig. 4.2(a), the gate driver is used to control multiple current driver NMOSs in the same group and the feedback signal is only sensed from one of the heaters/thermocouples. If the current driver NMOSs and the heaters/thermocouples are matched, the phase delays introduced by the phase shifters corresponding to those driver units are the same and a constant phase shift between two successive pixels in the same row is achieved. In this way, the temperature is regulated at the group level to offset the spatial process variations and reduce the phase errors.

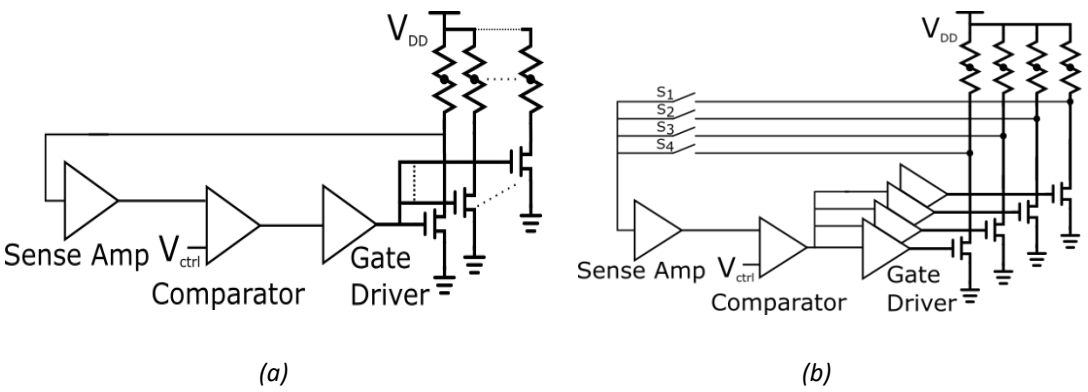


Fig. 4.2 Driver circuit diagram for (a) LIDAR NPA and (b) VR NPA

Unlike beam steering, there is no specific phase relation between the output beams of a NPA for forming a VR scene. To generate any arbitrary scene, the phase shift of each pixel must be controlled fully independently. Fig. 4.2(b) shows the diagram of VR NPA driver circuit where the sense amplifier and the comparator are shared among four driver units. Each current source NMOS has its own dedicated gate driver so that their output currents do not depend on the state of the other units sharing a sense amplifier and a comparator. The input of the sense amplifier is connected to the thermocouples state by the switches S_1 to S_4 in time-multiplexing fashion.

The control waveforms of the switches S_1 to S_4 is shown in Fig. 4.3. The switching frequency is 100 kHz with 10 % duty cycle. The switch S_1 is closed for 1 μs, t_1, connecting its thermocouple output to the sense amplifier and is opened for 9 μs, t_4, handing over the sense amplifier to the other units. The gate drivers are synchronized with the switch control signals. To prevent the overlapping between the sensing and the heating intervals, for example, the current source NMOS connected to the switch S_1 is turned off and on slightly before and after the t_1. Thus, the duty cycle of the heating is slightly less than 90 %. Because there are four drivers in a group, the sensing period are $\pi/4$ out of phase with respect to their following one. The interval between two successive sensing periods, t_2, is 1.5 μs. During this period, the sense amplifier keeps the output constant for a short amount of time so the following stages can process the signal unambiguously and resets itself for the next sensing. It is possible to include more driver units in a same group to save more space for implementing more sophisticated control algorithm, but it requires shorter period of sensing and resetting.

Next, the design of each stage is presented.

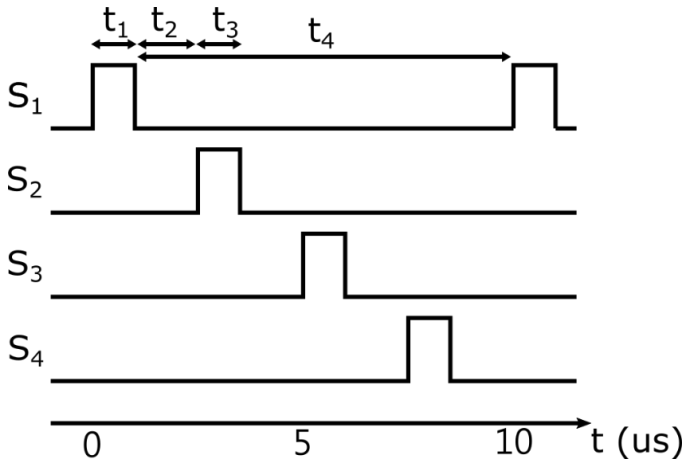


Fig. 4.3 VR NPA driver switch control timing diagram

4.2 Driver Circuit Design and Simulation

4.2.1 Heater/Thermocouple SPICE model

To design the driver circuit, a SPICE model of the heater/thermocouple device is needed to run circuit simulations. As stated earlier, the driver circuit system operates in two modes with several switches turning on and off, so we are most interested in transient simulation. A behavioral SPICE model of heater/thermocouple device which characterizes the steady state and transient responses from COMSOL FEM simulation is created using Verilog-A. The model has to reflect the temporal temperature change at the heater/thermocouple device caused by heating and cooling and generates a corresponding thermoelectric output voltage to the sense amplifier. It is important to note that the temperature mentioned here is different from the SPICE built-in temperature parameter. This built-in parameter is the temperature at the driver circuit and sets the simulation parameters of the transistors.

The schematic of the heater/thermocouple behavioral model is shown in Fig. 4.4. The model consists of three parts. The first, left-most, part models the thin-film heater. The resistance value of the R_{heater} is extracted from the FEM simulation steady state results. It is a function of V_T to simulate the temperature dependence. The V^+ is tied to power supply, V_{DD}, and V^- is tied to the drain terminal of the current source NMOS.

The model calculates the power generated at the heater, P_{OUT}, and uses it to evaluate the temperature at the heater/thermocouple.

The RC network in the middle part is used to model the transient temperature response. The voltage V_T across the capacitor C models the highest temperature at the heater which is designed to be at the thermocouple junction. In previous chapter, we showed the temperature at both the phase shifter and the heater is almost related to the heater's output power, so a controlled voltage source of which the value is proportional to heater's output power P_{OUT} is used to charge the capacitor C. In heating mode, the current source NMOS draws a current flowing through the heater which generates P_{OUT}. The controlled voltage source charges the capacitor C through R, simulating the transient temperature rise. In sensing mode, the current source NMOS is off so the P_{OUT} is near zero. The capacitor C is discharged through R and this models the cooling. The RC time constant is extracted from the FEM transient simulations. In this model, the absolute values of R and C do not matter. One of the values can be assigned arbitrarily as long as their product equal to the heating and cooling time constant.

The last stage is a controlled voltage source with the value being the product of the Seebeck coefficient and the thermocouple junction temperature, V_T. The output is connected to one of the input terminals of the sense amplifier.

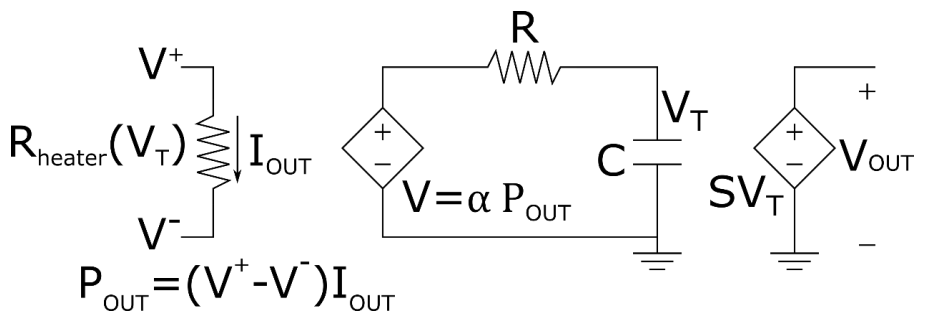
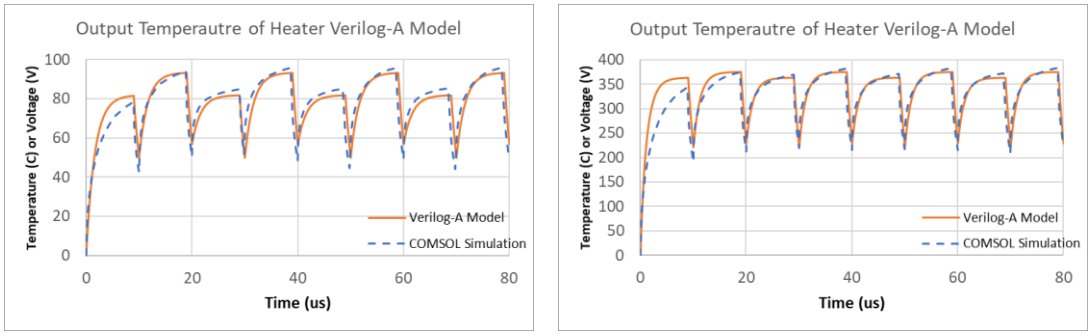


Fig. 4.4 Schematic of heater/thermocouple behavioral model

Figure 4.5 shows the caparison between the simulation results of the heater temperature from COMSOL FEM in Fig. 3.17 and Verilog-A behavior models

described in Fig. 4.4. The error is acceptable for driver circuit design. The SPICE model file can be found in Appendix A.



(a) *(b)*

Fig. 4.5 Temperature response of the heater Verilog-A model. Heater power dissipation: (a) Fig. 3.17(a) and (b) Fig. 3.17(c)

4.2.2 Sense amplifier

The first stage of the driver circuit is an amplifier. The output voltage of the thermocouple is extremely low. The maximum voltage is about 4 mV. This assumes an estimated 10 μV/K Seebeck coefficient multiplied by 400 K above ambient temperature at the junction for 2π phase shift of the output light. If this signal is used directly, the smallest increment value of the control signal also has to be small to distinguish different levels of junction temperature. For example, if the control circuit is designed to limit the phase error within π/16, the resolution of the control signal has to be at least smaller than 125 μV. A sense amplifier maps the small thermocouple signal to a signal with larger dynamic range so the requirement of the following signal processing circuit can be less stringent.

Another important function of the sense amplifier is to hold the amplified thermocouple signal for a short amount of time. The temperature at the heater keeps dropping during the sensing mode. If the amplified thermocouple signal also decreases continuously during this period, the timing of comparing the amplified thermocouple

signal and control signal must be the same exactly in every cycle for consistent results. However, in practice, the comparison may take place slightly early/late because of the clock uncertainty. The comparison result may be incorrect because the amplified thermocouple signal is bigger/smaller than the value at the correct moment. In addition, the speed of the comparator has to be fast enough so that the fluctuation of the amplified thermocouple signal does not affect the result. These design problems can be solved by having the sense amplifier to be able to lock its output. In this way, the comparator is less susceptible to timing issues by working with definite input signals.

Several sense amplifiers exist in the driver system depending on the total number of the driver group. It is preferred that the gains of these amplifiers are matched as close as possible. Otherwise, the range and the step size of the control signal has to be calibrated with respect to the gain of the sense amplifier for each driver group. This complicates the design of circuit for generating of control signal. Therefore, the sense amplifier itself should contain an internal feedback network to suppress the open-loop gain sensitivity to the transistors' intrinsic parameters, e.g. transconductance gm or output resistance for better matching between the driver groups.

Two possible realizations of the sense amplifier are shown in Fig. 4.6, (a) a switched capacitor (SC) sample-and-hold circuit with a fixed gain and (b) a RC integrator with an extra switch between the input and the resistor R. The features of these two circuit topologies meet all the requirement described above. The gain variation among the amplifiers in different group is minimized by using a capacitor-based feedback network. A controllable definite closed-loop gain can be obtained. In addition, as indicated by its name, the operations of a sample-and-hold amplifier perfectly matches the natural of the feedback mechanism of our driver circuit which also switches between two modes, temperature measuring and heating.

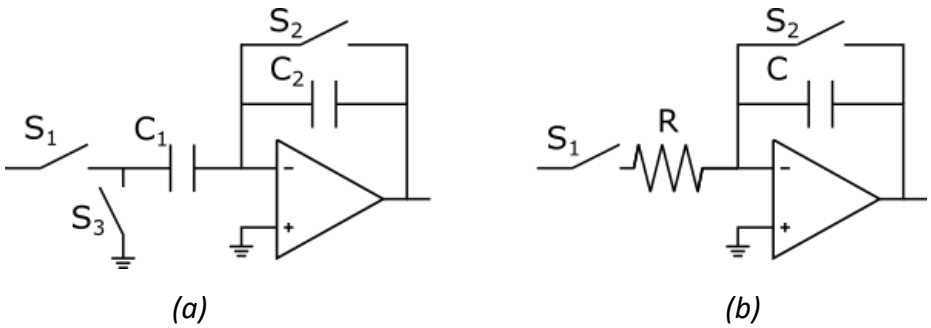


Fig. 4.6 Schematic of possible topologies of sense amplifiers (a) switched capacitor sample-and-hold (b) integrator

Next, brief descriptions of the working principle of both circuits followed by a detailed analysis of the pros and cons of them being in our system, actual circuit implementations, design considerations and simulation results are presented.

During the sampling mode of a sample-and-hold amplifier, the switch S_1 and S_2 is close and S_3 is open. The equivalent circuit of the amplifier in this mode is shown in Fig. 4.7(a). The voltage of the amplifier's inverting input and output are at the virtual ground because of the negative feedback which is formed by closing the switch S_2. The left plate of the capacitor C_1 is connected to the output of the thermocouple and the other plate is connect to the inverting terminal of the amplifier, holding its potential at the virtual ground. Thus, the voltage across the capacitor C_1 tracks the output voltage of the thermocouple.

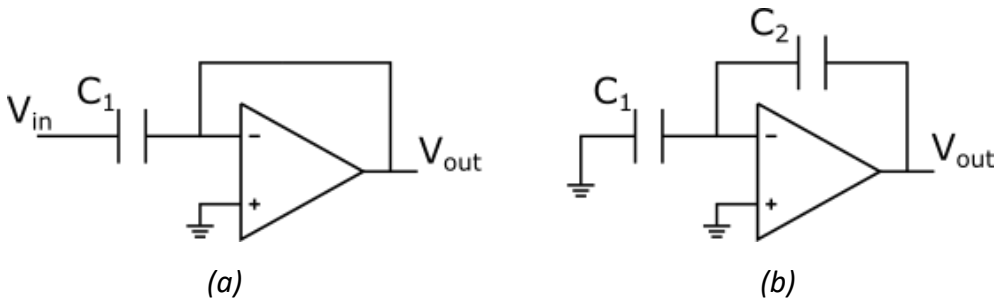


Fig. 4.7 Operation of a sample-and-hold amplifier (a) sampling (b) holding (amplifying) mode

During the holding (amplifying) mode, the states of the switch S_1, S_2 and S_3 are reversed. The capacitor C_2 forms a negative feedback loop around the amplifier. The voltage across the capacitor C_1 at the beginning of the holding mode is the

thermocouple output voltage at the time immediately before the switch S_1 being open. The amplifier provides a current to discharge the capacitor C_1 through the switches S_3, pulling the potential at the left plate of the capacitor C_1 to ground. Since same amount of current flows through the capacitor C_1 and C_2, the voltage across C_2, which is the output voltage of the amplifier, is the ratio of the capacitance of C_1 to that of C_2 multiplied by the sampled thermocouple output (V_{in}) after the C_1 is fully discharged. Thus, the thermocouple signal is sampled and amplified by a gain, C_1/C_2, and the output voltage is maintained for the comparator as the amplifier can compensate the leakage current of C_1 and C_2.

It is worth noting here that the ground connected to the non-inverting terminal of the amplifier and the capacitor C_1 is the small signal ground. This means the DC voltages at these two terminals can be different. It is an important property when considering the bias point of the amplifier and the voltage level of the thermocouple output. It is one of the unique features that differentiate this circuit from the integrator sense amplifier and have to be taken into account for deciding the implementation of the sense amplifier.

The operations of an integrator are shown in Fig. 4.8. In the first stage, the output of the amplifier is reset by having the switch S_2 to short the capacitor C_2. The potential of the output and the non-inverting input of the amplifier are at the virtual ground. Next, the switch S_1 is close and S_2 is open. The resistor R converts the thermocouple output into a current, charging the capacitor C. The output of the amplifier rises while sensing the temperature. The gain depends on the value of the resistor R, the capacitor C, and the sampling time. When the sampling is done, the switch S_1 is open, leaving the left terminal of the resistor R floating. Since no current flows through the capacitor C, the output voltage of the amplifier remains constant. After the comparator generates a result

for the gate driver, the switch S_2 is close again. The integrator is back to the resetting mode, clearing the charge in C and preparing for the next temperature measurement.

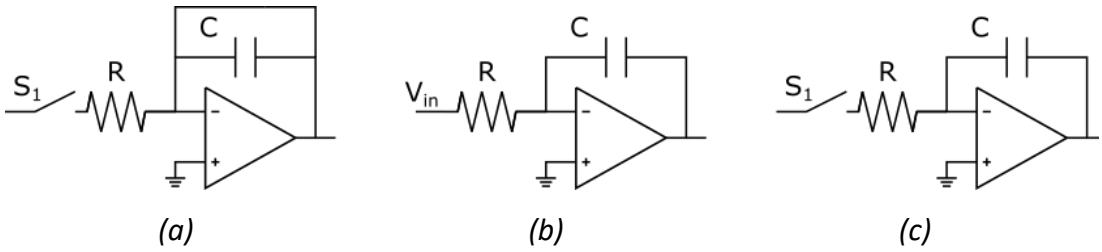


Fig. 4.8 Operation of an integrator amplifier (a) resetting (b) sampling (c) holding mode

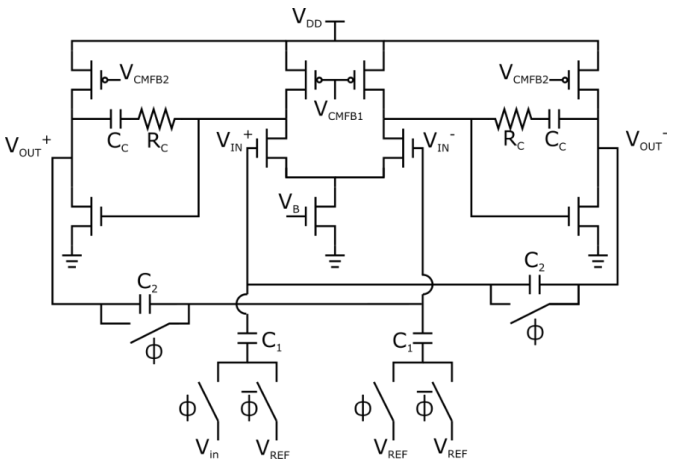
One of the undesired features of these two design approaches is the noise introduced when the switches are turning on and off. The noise can be categorized into clock feedthrough and charge injection [111]. Clock feedthrough is the switch control signal coupling to the amplifier from the switches' parasitic capacitor. Charge injection happens when the switches turn off. The charges in the channel of the switches transfer to the capacitor network. Because the area limitation, the capacitors used in the amplifier must be kept small. This makes the capacitors be more sensitive to the switching noise and degrades the precision. One way to minimize the error is using fully differential operation to cancel the switching noise.

Figure 4.9 shows the schematic and layouts in TSMC 28 nm technology of a two-stage fully-differential sample-and-hold sense amplifier. This design uses core transistor with 0.9 V rail voltage. Two common mode feedback (CMFB) circuits, not shown in the schematic, provide the gate bias to PMOS current source loads in both the first and second stage and define the common mode of the outputs. The common mode at the inputs of the core amplifier is the same as that of the outputs because they are shorted by the switches during the sampling mode. The common mode at the capacitor inputs is V_{REF}. The value of the input voltage, V_{in}, is the V_{REF} minus the thermocouple output. The difference between the common mode at the amplifier's inputs and the V_{REF}

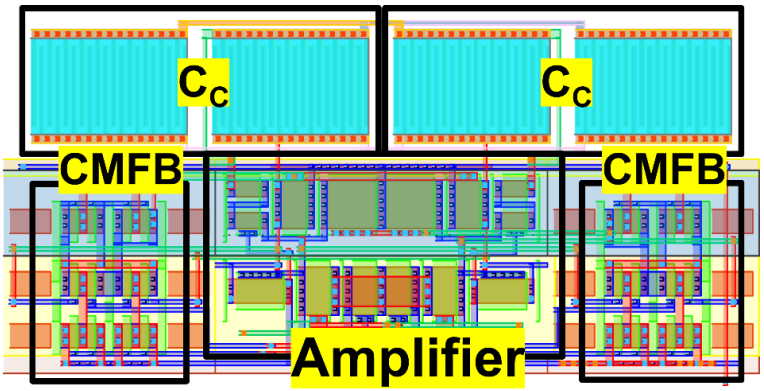
does not change the output common mode during the amplifying mode. Only the difference between the V_{in} and the V_{REF} is amplified. Thus, the common mode of the amplifier and the V_{REF} can be chosen independently for design optimization.

The major drawback of this approach for our driver design is it needs large space. Besides two frequency compensation capacitors, C_C, the amplifier contains two sampling capacitors, C_1, and two feedback capacitors, C_2. To get a large gain, the ratio of C_1 to C_2 has to be large. This means either C_2 has to be small, C_1 has to be large or both. The minimum effective value of the C_2 is limited by the design rule and the parasitic components. The actual capacitance of C_2 is its nominal design value plus all the parasitic capacitors that connect the outputs and the inputs of the amplifier. If the nominal value of C_2 is comparable to or even smaller than the parasitic components, the capacitor C_1 has to be made larger to offset the gain reduction due those unwanted parasitic components.

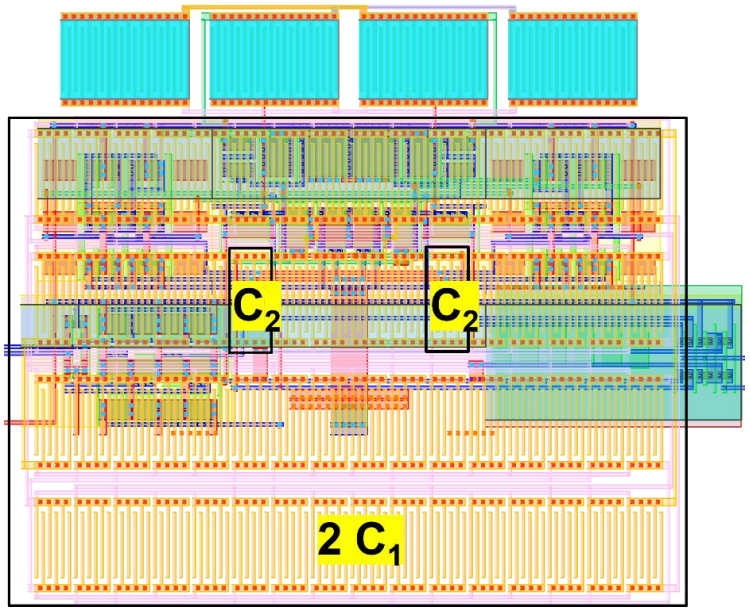
Figure 4.9(b) and (c) shows the layout of the design. The capacitors C_1 and C_2 are hidden in Fig. 4.9(b). Two CMFB circuits are located next to the core amplifier both side symmetrically to bias the current source PMOSs in the amplifier. The layout is symmetric to the center to cancel layout dependent effects (LDEs) to the transistors. All capacitors are implemented using metal-oxide-metal (MOM) capacitor. A MOM capacitor is form by interdigitated metal fingers in the same layer. A deeply scale technology allows small distance between two parallel metal lines, to high capacitance density can be obtained from this type of capacitors. Same MOM capacitor structures can be made in different metal layers vertically and be connected in parallel to get a large capacitor.



(a)



(b)



(c)

Fig. 4.9 two-stage fully-differential sample-and-hold sense amplifier (a) schematic (b) layout with capacitors hidden (c) full layout

The compensation capacitors, C_c, contains six metal layers, M1 to M6 while C_1 and C_2 only use metal layers from M4 to M6. To get good device matching of C_1 and C_2, C_2 is set as the unit capacitor with 6 fingers. The ratio of C_1 to C_2 is 31 and C_1 is connected in common centroid style. It is obvious that the upper part of the C_1 and C_2 unit capacitor array is on top of the core amplifier in Fig. 4.9(c). Not shown in the figure, the comparator and gate driver circuit will be placed beneath the bottom part of the capacitor array to fully utilize the space. It is against the generally recommended layout rules for analog circuit because noisy signals from the circuit underneath the capacitor array would be coupled to the input terminals of the amplifier and worsen the accuracy and performance. The issue can be mitigated by placing a grounded metal plate, M3, between the capacitor array and the circuit below it. The metal plate shields the amplifier from the noisy signals, but it comes with some drawbacks. It introduces extra parasitic elements between nodes in the amplifier to the ground which slow down the circuit. The M3 metal plate also limits the circuit below it to use only M1 and M2 metal layer for routing. This may lead to interconnection congestion and increase the coupling parasitic capacitance between the signals.

The circuit in Fig. 4.9(a) can be modified to a fully differential integrator by replacing capacitors C_1 with resistors and removing one of the switches connected to C_1. However, fully differential integrators are not suitable for our driver circuit. Unlike using capacitors to sample an input voltage, the difference between the common mode voltage at the input terminals of the core amplifier and V_{REF} generates currents flowing through the resistors. Therefore, this common mode difference is also integrated by the amplifier and experiences a same amount of the integration gain as the desired input. Although measuring the output voltage differentially can cancel the output variation, the common mode voltage mismatch may drive the output beyond the operation range

which reduce the accuracy. Again, the common mode of input terminals of the core amplifier is defined by the CMFB circuit and is not related to the V_{REF}. Since the thermocouple voltage is small, a small common mode mismatch is enough to saturate the output of the integrator. Hence, additional circuit is required to match the common mode across the resistors and it may cause the total circuit area exceeds the constraint. As a result, single-ended output integrated is used. The positive input terminal of the amplifier can be tied to V_{REF} directly. The negative input terminal is virtually shorted to V_{REF} by the feedback network, so ideally the common mode voltage across the integration resistor is zero.

Figure 4.10 shows the schematic and the layout in TSMC 65 nm of the single-ended integrator for our driver circuit design. All transistors are I/O devices with a 2.5 V voltage rating. The core amplifier is a single-stage cascode amplifier using current mirror load. Since the thermocouple output is V_{DD} minus the voltage generated at the thermocouple junction, a source follower circuit is needed to shift down the thermocouple output for setting the DC bias points of the amplifier properly. The source follower also isolates the thermocouple from the integrator's input. It avoids the loading of the thermocouple. The input current to the integrator is provided by the source follower, not the thermocouple itself.

Each switch is formed by two NMOSs. One of the transistors is dummy with its source and drain terminals being shorted. The size of the dummy switch is half of the size of the actual switch. Their gates are controlled by a pair of complementary signals. The function of the dummy switches is the clock feedthrough and charge injection cancellation. Complementary switch method for the noise cancellation is not used here because PMOS switches result in higher area overhead.

A dummy switch, S_{1_Dummy}, is inserted between the shifted down V_{DD} and the

positive input node of the core amplifier. This switch matches the drain-source voltage source drop along the input signal path to further reduced the DC voltage error across the integration resistor. The integration resistor is realized by a NMOS operated in triode region. This assumption is valid for the entire range of operation because the maximum expected input signal is less than 5 mV.

Both the integration capacitor and the frequency compensation capacitor are implemented by metal-insulator-metal (MIM) capacitor. A MIM capacitor utilizes the two extra metal layers between the M7 and M8. To obtain large gain, the integration capacitor has to be small. Due to the minimum space design rule of the MIM layers, two MIM capacitors are connected in series to get the desired capacitance. The node where two MIM caps share is tied to a n-plus diffusion region in the layout as shown in Fig. 4.10(b). It is a workaround for satisfying the antenna design rule of the MIM device. Antenna design rule dictates that the structure of a MIM device must be balanced, i.e. if one metal plate is connected to a gate or an active region, the other metal has to be connected to a same type of region. Because one node of the MIM capacitors in series are connected to the amplifier's output and input (active region and gate), the center node is connected to an n-plus region to form balance structures. This creates a parasitic pn junction between this node and the substrate. The voltage at this node is always above ground, so this parasitic pn junction is reverse biased all the time. It only introduces an equivalent parasitic capacitor to ground and does not affect the function of the circuit.

The area of the integrator without the MIM capacitors is about 11.5 μm by 5.5 μm. All the NMOSs and PMOSs are enclosed by a p-type and n-type guard ring (substrate contact and N-well contact) for better noise and latch-up immunity. The comparator, not shown in the Fig. 4.10(b), is to the right of the integrator. The frequency

compensation capacitor, C_C, is on top of the integrator, and the integration capacitor is overlapped with the gate drivers. The design rule forbids M7 to be overlapped with MIM bottom plate. Thus, grounded M6 metal plates are used to block the noise from the comparator and the gate drivers. The remaining metal layers for routing under the overlapped regions are M1 to M5.

Next, the effect of the non-ideal core amplifier's circuit performance, finite gain or transconductance is examined. For simplicity, the voltage at the positive input of the amplifier is assumed to be V_{REF}, and an input voltage, V_{REF} - V_{in}, is applied to the integration resistor. The current flowing through the integration capacitor is provided by the core amplifier. Because of the finite transconductance, a voltage difference between the positive and negative input of the core amplifier is required to output the current. Assume the voltage difference is ΔV. The actual current flowing through the integration resistor would be

$$I = \frac{(V_{REF}-\Delta V)-(V_{REF}-V_{in})}{R} = \frac{V_{in}-\Delta V}{R} = gm\Delta V \tag{4.1}$$

where gm is the transconductance of the core amplifier. Rewriting equation (4.1), ΔV can be related to the input voltage, V_{in}, as:

$$\Delta V = \frac{1}{1+gmR} V_{in} \tag{4.2}$$

Assuming the integration time is T, the actual integrator output is

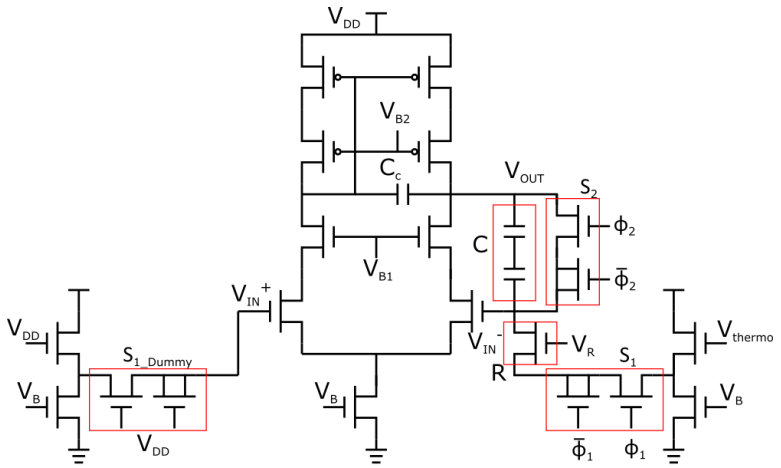
$$V_{out} = \frac{gmR}{1+gmR}\frac{1}{RC}\int_0^T V_{in}(t)dt \tag{4.3}$$

The output voltage is reduced by a factor $\frac{gmR}{1+gmR}$. To minimize the gain reduction, the value of gmR should be large. Thus, when using a smaller value of R to achieve a larger output voltage by having a smaller product of RC, the value of transconductance gm has to be increase accordingly. The result is as expected. Reducing the value of integration resistor increases the input current. The input current is supplied by perturbing the DC bias current of the amplifier. To keep the same ratio of the input

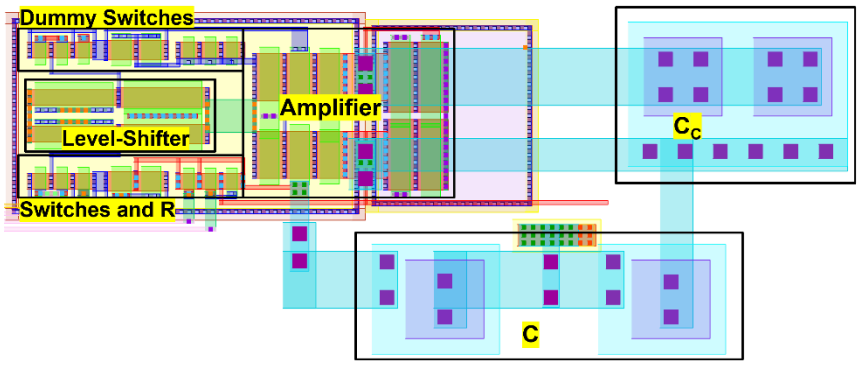
current to the DC bias current, the DC bias current needs to be larger hence a higher transconductance gm.

The post-layout simulation of the integrator is presented in Fig. 4.11. The post-layout netlist includes lump and coupling capacitors and transistor parameters for LDEs. In Fig. 4.11(a), the input voltage is a constant 2 mV below the V_{DD}. The common mode of the amplifier's inputs and output is set at 1.2 V. The operation frequency of the switches is 100 kHz. The switch S_1 is on during T_1 to perform the integration. The switch S_2 is off slightly before T_1 so the on time of two switches does not overlap. The integration time is 1 μs. The output voltage increases about linearly because the input voltage is constant. During T_2, both the switch S_1 and S_2 are off. The output voltage is hold constant for 1 μs, giving the following comparator and gate driver enough response time. During T_3, the switch S_2 is on, shorting the output and the negative input node. The integrator capacitor is fully discharged. The reset is finished instantaneously. It is obviously that three more integration operations can be fit in T_3 so this integrator can be share between four driver units as discussed before.

The output voltage as a function of input voltage being constant from 0.5 mV to 5 mV below V_{DD} is shown in Fig. 4.11(b). The curve is nearly linear for input voltage less than 3 mV and begins to saturate above 3 mV. This is because the output voltage is so high that the current source PMOSs enter the triode region. The voltage at the negative input node starts to deviate from the virtual ground potential. The output current is reduced and less charges are stored in the capacitor during the integration. The actual thermocouple output is not constant but falls rapidly. Again, the expected maximum voltage is less than 5 mV. Therefore, the output voltage saturation will not happen when an actual input is used.

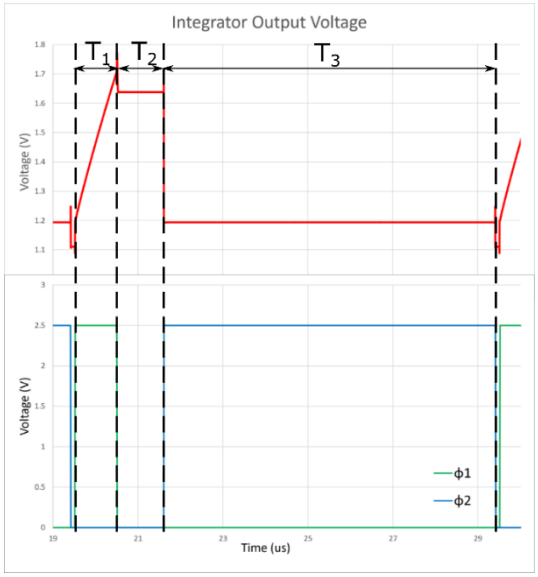


(a)

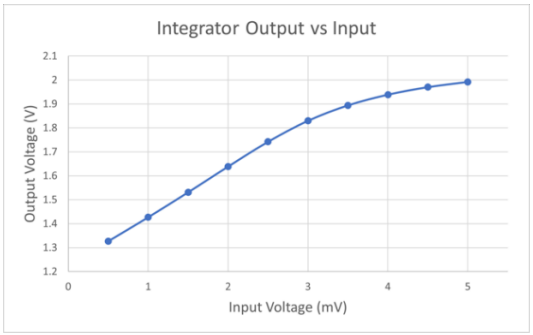


(b)

Fig. 4.10 Single-ended output integrator amplifier (a) schematic (b) layout



(a)



(b)

Fig. 4.11 Post-layout simulation results of the integrator sense amplifier

Two circuit implementations of the sense amplifier used in the driver circuit were demonstrated. The major advantage of the sample-and-hold amplifier approach is that the switching noise can be suppressed by fully differential operation without having to match the common mode of thermocouple's output and the input terminals of the core amplifier. There are two drawbacks of this approach. First, the number of the required capacitor is doubled for differential operation. The physical size of a unit capacitor must be small so a large number of unit capacitors can be fit in the allowable space to achieve a high gain. Thus, the design has to be realized by a deeply scale technology. Second, a unique characteristic of the input from thermocouple is that it decays rapidly in the sensing mode of the driver circuit because the temperature at the thermocouple junction decreases when no electric current flow through the heater. To fully utilize the range of the thermocouple output, the sampling should take place immediately after the driver circuit leaves the heating mode. However, a short period is required for the sampling capacitor to track the thermocouple's output voltage. The available sampled value is lower than the peak voltage when the driver circuit changes states. Moreover, the uncertainty of the switches closing time affects the measurement accuracy because the input is not stable.

The single-ended integrator approach suffers more from the switching noise but requires less space. The entire range of the thermocouple's output can be utilized because the resistor can respond instantly. The integrator samples the input for a fixed interval and the input has fallen to a smaller value when the switch is turned off at the end of the sampling period. Therefore, the switch uncertainty has less influence on the measurement accuracy.

Another major disadvantage of the integrator approach is using a MOSFET in deep triode region to simulate a resistor. The amplifier's gain is more susceptible to the

process variations and calibration is required to cancel the gain error.

Since our driver circuit is fabricated using 65 nm technology, the integrator amplifier is chosen for implementing the sense amplifier because of less required circuit area. Nonetheless, the sample-and-hold amplifier approach is better for noise immunity and achieving process independent gain. It can be realized in more advanced semiconductor technology such as 28 nm or below.

4.2.3 Comparator

After the temperature signal from the thermocouple is measured and amplified, the next step is to adjust the current flowing through the heater according to the difference between this signal and a reference control signal like any other feedback system. As a result, a straightforward way to accomplish the task is to use a feedback system as in Fig. 4.12. The circuit in Fig. 4.12 is similar to the one in Fig. 4.1 but it contains only three stages without a gate driver. The error amplifier in Fig. 4.12 drives the NMOS current source directly. When the temperature at the heater is less than the desired value, the output of the sense amplifier is smaller than the control reference, V_{ctrl}, the output voltage of the error amplifier, V_{error}, is increased and the NMOS current source draws more current to raise the heater's temperature, and vice versa. Ideally, if the gain of the error amplifier is large and the feedback loop is stable, the amplified thermocouple signal will be close to V_{ctrl} which means the heater's temperature reaches the target value.

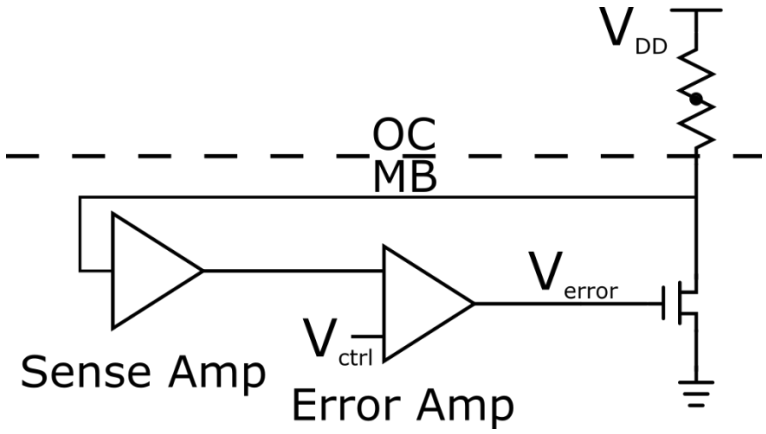


Fig. 4.12 Driver control circuit without gate driver

The simulation result of the amplified feedback signal at different values of the control signal of a driver system in Fig. 4.12 is shown in Fig. 4.13. The output of the sense amplifier in the simulation is a differential signal so as the control signal. The switching frequency is boost to 2 MHz and the heating duty cycle is 80%. It is obviously that the amplified signal is not close to the control signal. For the control signal being 60 mV or 80 mV, the amplified feedback signal only meets the control signal every three cycles at equilibrium. The error is big for the rest of time. For a larger value of control signal, 100 mV or 120 mV, the amplified feedback signal is not close to the target at any time. The polarity of the amplified feedback signal with respect to the target switches at every cycle.

In all cases, the control loop regulates the “average” of the amplified feedback signal instead of the signal itself to the desired value. It is because the control loop is not closed all the time but is broken during the heating mode. The control loop is not able to adjust the current to respond to the temperature at the heater instantaneously in the heating mode. A large difference between the feedback and the control signal results in overheating. The current source is turned off too hard in the following cycle, so the temperature falls far below the target value and the entire process is repeated.

This control fashion causes high temperature fluctuation. A possible way to reduce

the ripple is to increase the switching frequency so the feedback loop can respond earlier to prevent the overheating or the overcooling. To maintain the same heating capability, the heat duty cycle has to be the same. This means the sampling time decreases as the switching frequency increases. Therefore, increasing the switching frequency not only leads to a higher power loss, the switching loss of the current source, but also introduces design challenges because of the shorter sampling time.

Another reason that a standard error amplifier is not suitable for our driver system is it does not hold a constant output voltage with the input being floated. In the heating, in order to have the NMOS current source to supply a steady current mode, a constant voltage difference between the two input terminals of the error amplifier's must be presented. The sense amplifier needs to maintain its output for the entire heating cycle so it cannot be shared among different driver units by time-multiplexing. Each driver unit needs a dedicated sense amplifier and error amplifier so the entire system cannot be fitted into the allowable space.

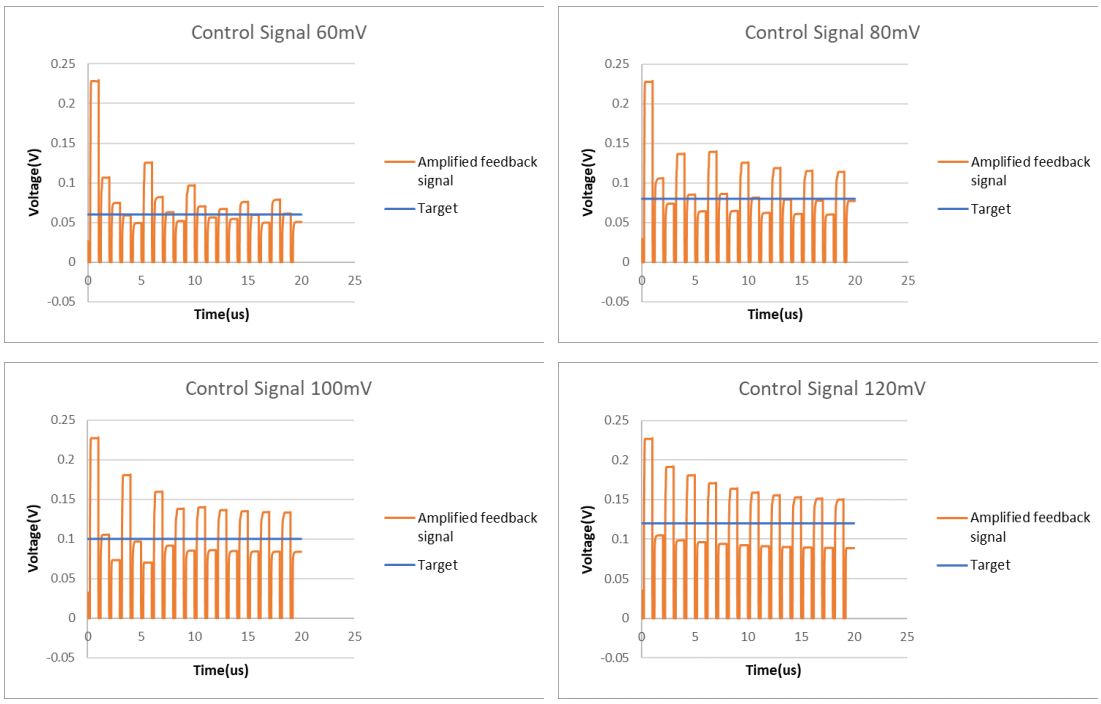


Fig. 4.13 Simulation result of the voltage of amplified feedback signal at different level of control signal

To solve the above design challenges, our system uses a comparator to test the amplified feedback signal against the control reference. The comparator generates a digital signal which is the input of a gate driver circuit. The gate driver updates its state by evaluating the comparator's output at some specific time frame and its output remain unchanged for the rest of the time. The output of the gate driver changes gradually in each state transition so the temperature ripple can be minimized. The gate driver only reads the result of the comparator at a particular interval, multiple gate drivers can be controlled by a same sense amplifier and comparator consecutively.

The comparator consists of three stages as shown in Fig. 4.14. The first stage is a differential pair with diode-connected, M_{P1} and M_{P4}, and cross-couple M_{P2} and M_{P3}, loads. Loads M_{P2} and M_{P3} create a positive feedback to speed up the output transition. The small-signal output impedances of the M_{P2} and M_{P3} are negative. A special case happens when (M_{P1}, M_{P4}) and (M_{P2}, M_{P3}) are match. The small-signal output impedances of (M_{P1}, M_{P4}) and (M_{P2}, M_{P3}) are equal in magnitude but opposite in sign, so they cancel each other out. The total equivalent output impedance of the load becomes infinity. The small-signal voltage gain of the first stage reaches a maximum value, depending on the transconductance and the output impedance of the input NMOS transistors.

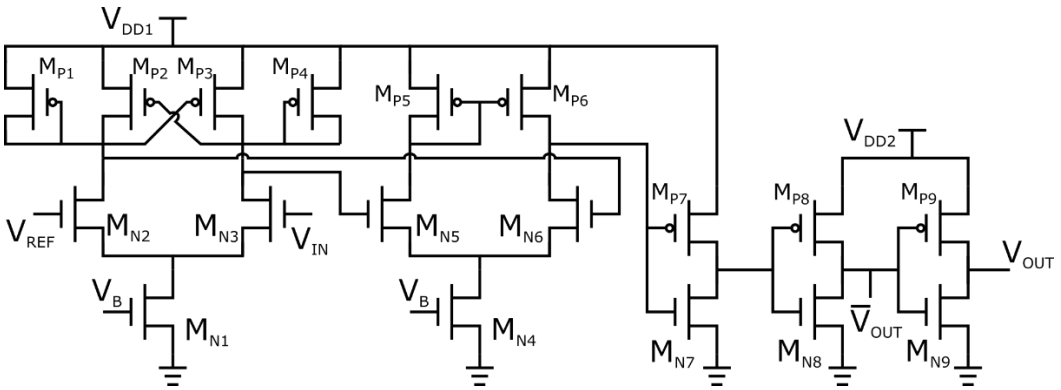


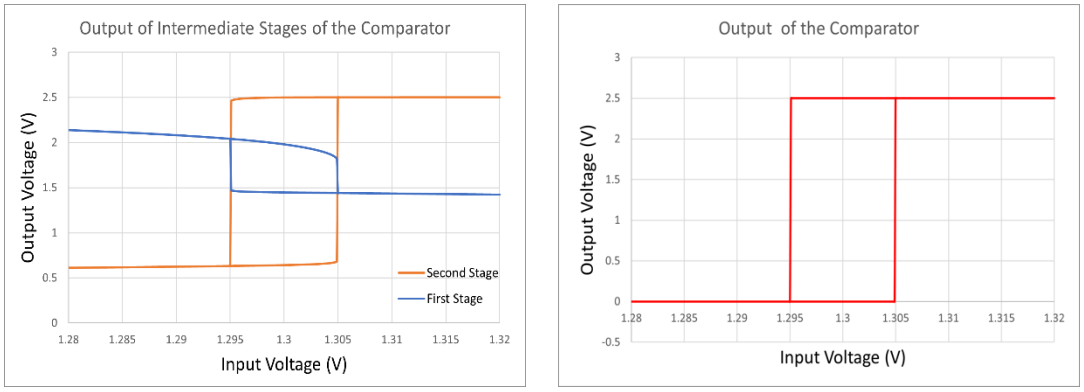
Fig. 4.14 Schematic of the comparator used in the driver circuit

A common problem of comparators is the stability issue. The output may oscillate because of the high gain and the signal coupling between the input nodes and the outputs by unwanted parasitic elements. For our driver circuit, the devices and interconnections are placed tightly together to fulfill the space requirement. It is impossible to eliminate all the parasitic signal paths between the comparator's outputs and V_{in}. When the integrator switches from the integration to holding mode, its output experiences a drop because of the switching noise as shown in Fig. 4.11. If the drop crosses the comparison threshold, V_{REF} in Fig. 4.14, and the final value of the integrator's output is very close to V_{REF}, the comparator changes the state and large voltage swings at its outputs are coupled back to the integrator's output which increases slightly and across V_{REF}. This forces the comparator to transit to the opposite direction. The above coupling is repeated in the opposite direction endlessly, so the integrator's output varies around V_{REF}, making the comparator oscillate. Introducing hysteresis to the comparator can solve this problem. Therefore, instead of matching (M_{P1}, M_{P4}) and (M_{P2}, M_{P3}) to maximize the voltage gain of the first stage, the width-to-length ratio of (M_{P2}, M_{P3}) is made larger than that of (M_{P1}, M_{P4}).

The outputs of the first stage are hard to be pulled low because of the diode-connected loads unless the bias current is large which results in high power consumption and large transistor width of M_{N1} or channel length of M_{P1} to M_{P4} is long. Both cases require large circuit area. The first stage generates a differential output feeding to the second stage which is a differential pair with current mirror load. The second stage provides additional voltage gain, and its output can be low enough to have the output of the inverter M_{N7} and M_{P7} be high without large circuit area or bias current. The inverter M_{N7} and M_{P7} generates a digital output at voltage level V_{DD1} and is followed by successive inverters which can be operated at different voltage level, V_{DD2}.

The last two inverters provide level-shifted complimentary outputs for the gate driver.

Post-layout simulation results are shown in Fig. 4.15. The threshold voltage V_{REF} is held at 1.3 V and the input voltage V_{IN} is swept between 1.28 V and 1.32 V for both directions. The comparator switches the state when V_{IN} is 5 mV above/below V_{REF} in upward/downward transition, so the maximum voltage variation tolerance at V_{IN} is 5 mV for the worst-case scenario. The output low voltage of the first stage is about 1.43 V which is not low enough to turn off the NMOS in an inverter. The differential output voltage of the first stage is more than 330 mV if V_{IN} deviates from the equilibrium more than 0.3 μV and it is large enough to saturate the second stage differential pair. The output low voltage of the second stage is lowered to 630 mV and a clean digital output is produced at the final output of the comparator. The hysteresis does not deteriorate the circuit precision. As the output voltage of the integrator drops when the integration is finished and the amount of the voltage drop is larger than the hysteresis range, 10mV, the comparison threshold is the lower limit of the hysteresis. Thus, the hysteresis only adds a constant offset to V_{REF} and the effect can be compensated easily.



(a) (b)

Fig. 4.15 Post-layout simulation results of the comparator

4.2.4 Gate Driver

The gate driver provides a bias voltage to the gate of the current source NMOS to control the heating current. It has two key functions. First, it gradually adjusts the heating current based on the measured temperature signal from the previous cycle to avoid abrupt temperature variations due to over or insufficient heating. Second, it maintains a steady gate control voltage without needing a continuous input signal from the comparator, so the integrator and the comparator can be used by the other driver units. The gate driver circuit is realized by two opposite approaches, digital and analog. They are tailored to meet the circuit area constraint for the LIDAR and VR driver systems in Fig. 4.2 respectively.

4.2.4.1 Digital Gate Driver

The digital implementation uses a counter to represent the state of the current driver. The current source for the digital gate driver is formed by an array of small unit transistors. The gate voltage of each unit transistor is either at ground or V_{DD}, and the counter determines how many unit transistors are on. The counter is either incremented or decremented by one every cycle based on the comparator's output. To save the circuit area, the design is simplified so the clock signal is not gated. The counter never stops updating but switches between two consecutive states when the temperature at the heater approaches the target value. More sophisticated and efficient update algorithm is possible, but the circuit will occupy more area. The unit transistors are binary encoded in our design. A major advantage of using a counter to represent the state of current source is that different encoding for compensating the nonlinearity relationship between the control refence voltage and the heater temperature can be easily implemented if needed.

Core transistors can be used to realize the counter, so the digital gate driver design

can fully take advantage of the minimum devices offered from submicron process nodes. The size of the counter can be further scaled down with future advanced nano-semiconductor technologies effortlessly. However, level shifter circuits are required to boost the counter's outputs to I/O voltage level to fully turn on the current source transistors. Moreover, the design rules demand the transistors of the counter to be surrounded by guard rings. The drain terminals of the current source NMOS array are connected to an I/O pad and are called OD injectors. To comply with the rules for latch-up prevention, transistors located within 20 μm from OD injectors have to be surrounded by guard rings. This distance is bigger than the average size of control driver unit, so all transistors need to be placed in guard rings. It introduces a significant area overhead, especially for the counter using core devices. Other latch-up prevention rules require a minimum spacing between PMOS transistors and the current source NMOS array. Special device placement must be taken to minimize the size of the circuit layout. Because of the latch-up design rules, it is difficult for our VR driver with four units in a group to meet the space specification using 65 nm technology. This will be shown in more details when the layout of the counter is presented.

Each bit of the counter is represented by a master-slave J-K flip-flop as shown in Fig. 4.16(a). Because the internal propagation delays of the logic gates are extremely short, the edge-triggered pulse has to be short and precise to avoid the race condition and assure the functional correctness for gated latch flip-flops. To circumvent the issues caused by the edge-triggered pulse uncertainty, the master-slave implementation is used at the expense of more space.

The inputs J and K of the flip-flops are tied together, so the outputs of each flip-flop switch or remain unchanged if the input is 1 or 0. The outputs of a flip-flop representing bit n, Q_n and $\sim Q_n$, need to switch when the counter counts up/down and

the state of bit 0 to bit n-1 are all 1s/0s respectively. The input of the bit n flip-flop does not depend on its current state but the output of all its previous bits and the comparator's output and can be written as:

$$Input_n = (Q_0 \& Q_1 \ldots \& Q_{n-1} \& UP) | (\sim Q_0 \& \sim Q_1 \ldots \& \sim Q_{n-1} \& DOWN) \qquad (4.4)$$

where UP and DOWN are the commands from the comparator's complimentary outputs for n > 0.

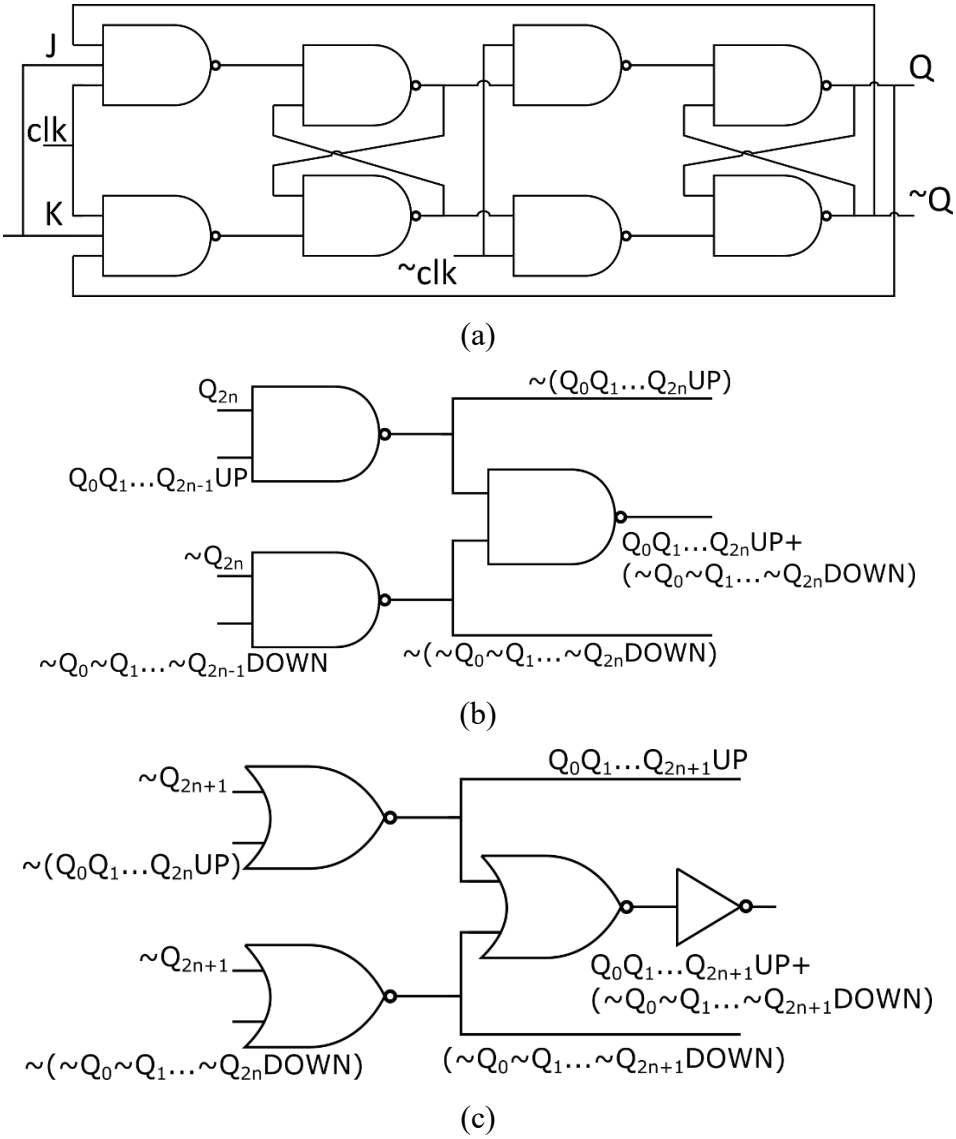


Fig. 4.16 Schematic of counter's components (a) master-slave J-K flip-flop (b) and (c) intermediate stage to determine the flip-flops' inputs.

The input of the first flip-flip, bit 0, is connected to V_{DD}, as this bit flips every

cycle for both counting up and down cases. A straightforward way to implement equation (4.4) is to use two n-input AND gates and one 2-bit OR gate. The n-input AND gates can be replaced by 2-input AND gates to reduce the gate count by reusing the intermediate results of setting the input of the preceding flip-flop. Thus, for a circuit expressing the equation (4.4) for the bit-n flip-flop, it receives the outputs of the bit-(n-1) flip-flop, Q_{n-1} and Q_n, and intermediate results of the bit-(n-1) flip-flop's input setting circuit, (Q_0&Q_1…Q_{n-2}&UP) and (~Q_0&~Q_1…~Q_{n-2}&DOWN). It generates three outputs, one for the input of bit-n flip-flop, and two signals, (Q_0&Q_1…Q_{n-1}&UP) and (~Q_0&~Q_1…~Q_{n-1}&DOWN) to determine the input of bit-(n+1) flip-flop.

Although all the flip-flops' input can be expressed as equation (4.4), odd and even bits of flip-flops are implementing differently using the circuit in Fig. 4.16(b) and (c) respectively because standard CMOS logic gates circuit only offer inverting functions. Three NAND gates in Fig. 4.16(b) can express equation (4.4) but the intermediate results are inverted. Therefore, the logic circuit for the following stage uses three NOR gates and one inverter to accommodate the inverted intermediate results from the previous stage.

To ensure the functional correctness, when the counter is updated, the outputs of flip-flop bit-0 to bit-(n-1) have to be hold constant before the flip-flop bit-n latch its input. If single stage gated latch flip-flops are used and the internal delay is short, the clock pulse needs to arrive to the flip-flops in the sequence from the most significant bit (MSB) to the least significant big (LSB) to prevent race conditions. Violating this rule may cause a flip-flop latches an incorrect value because one or more of the previous flip-flops update their outputs too early. This again justifies the extra circuit space from using master-slave flip-flop for relaxing the timing constraints.

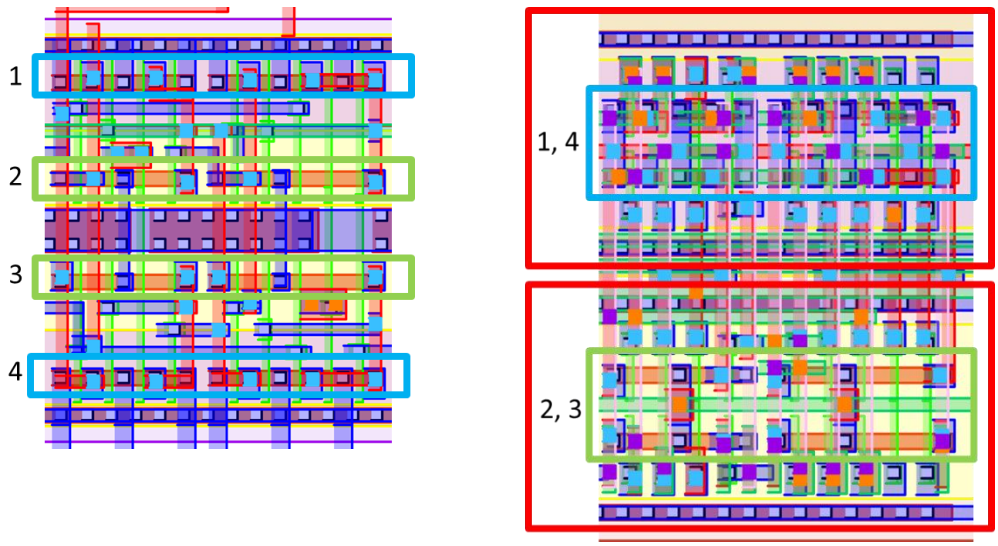
The major challenge of the counter design is to minimize the space occupied.

Logic gates from the standard cell library IP and automatic placement and routing are not suitable for creating the layout of the counter because these tools can hardly optimize the layout based on our unique design limitations and introduce huge area overhead. The layout of the counter is created manually for circuit area optimization.

Figure 4.17 shows the layout of a master-slave J-K flip flop with and without guard ring. The layers above M2 are hidden for clarity. In both layouts, all transistors are abutted at least on one side to share their source or drain terminals to save area and reduce the overhead of spacing between active regions in horizontal direction. The flip-flip consists of two rows of NAND gates as in Fig. 4.16(a). Intuitively, the simplest way of the transistors placement is to follow the schematic, having two rows of NAND gates and connecting the gates of NMOS and PMOS pairs of those NAND gates with poly line directly as in Fig. 4.17(a). The vertical distance between the NMOSs and PMOSs in a same NAND gate is longer than the sum of the minimum distance between a N-well and N-type active regions and the minimum N-well enclosure for P-type active regions required by the design rules. Thus, the design rules do not introduce area overhead at these regions as they are fully utilized for placing horizontal interconnection metal lines.

However, the transistors in Fig. 4.17(a) are not surrounded by any guard ring so it cannot be placed within 20 μm next to the driver current source. The vertical ploy lines forbid horizontal substrate and N-well contacts to be placed between the NMOS and PMOS pairs and closed contacts cannot be formed around the NMOSs and PMOSs. To form guard rings, the gates of the transistors in NAND gates are connected by metal layers instead and one end of each gate needs to be expanded to have a poly contact as demonstrated in Fig. 4.17(b). Instead of having a symmetric devices placement for the two rows of NAND gates with respect to the horizontal axis, having one NMOS cluster

at the center and two PMOS clusters at the top and bottom or vice versa, only one NMOS cluster and one PMOS cluster are used to minimize the number of guard rings so the area overhead. This arrangement requires that the gates and drains of the inner NMOSs and PMOSs of the first row of NAND gates are connected by M2 layer and the gates and drains of the outer NMOSs and PMOSs of the second row of NAND gates are connected by M4 layer. Hence, two rows of NAND gates are not match and plenty of parasitic capacitors are formed because of the overlap between the M2 and M4 vertical connections. Fortunately, the circuit is digital and does not operate at ultra-high speed so these non-ideal effects on performance are negligible.



(a) *(b)*

Fig. 4.17 Layout of a J-K flip-flop in Fig. 4.15 (a) without (b)with guard ring. The transistors in the blue rectangles are PMOSs (labeled 1 and 4 in (a) and labeled 1,4 in (b)) and the transistors in the green rectangles are NMOSs (labeled 2 and 3 in (a) and labeled 2,3 in (b)). The transistors in (a) have to be rearranged as shown in (b) so they can be surrounded by guard rings (the two outermost red rectangles in (b))

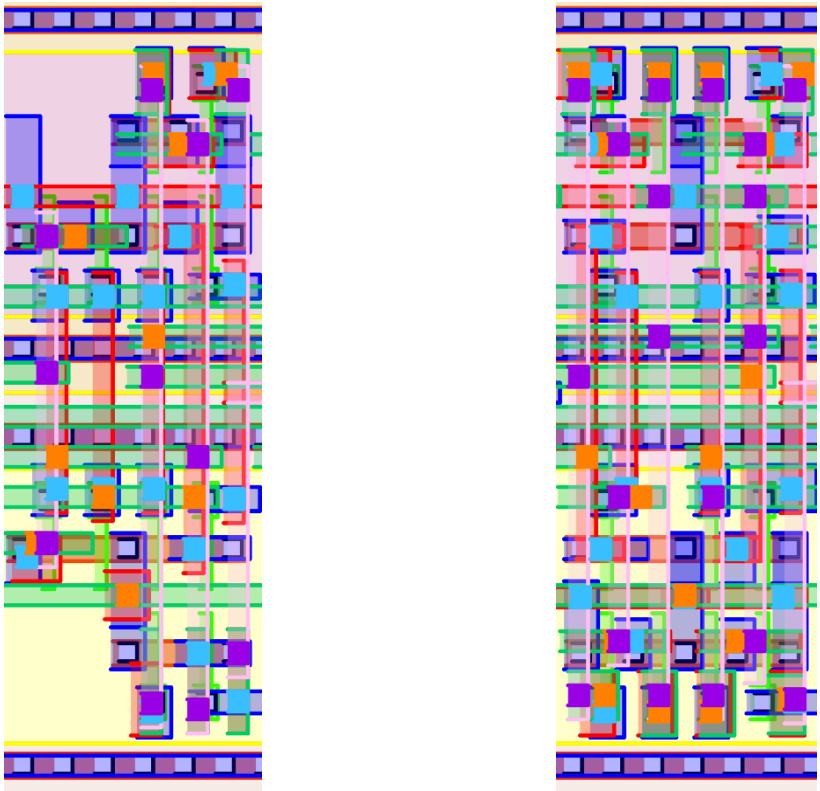
Using M4 layer is not as easy as it seems because of the design rules. The design rules of an advanced technology such as 65 nm in our case specifies the minimum area for each metal piece depending on the metal layer. For example, the M1 rectangles

covering contacts and VIA1s in Fig. 4.17(b) are larger than the area to satisfy the minimum enclosure rule of a single contact or VIA1. The M4 connecting lines do not cause problems because they are long. The difficulty lies in the placement of the metal pieces and vias connecting the poly gates and M4 lines. In addition, the minimum spacing rule of metal and via layers which is not a single number for most older microscale technologies but varies with the dimension of the metal pieces and the local metal density further increases the complexity of the placement and routing of the interconnections. As stated earlier, the layout has to be drawn manually to minimize the total circuit area and to comply with all the design rules at the same time.

Comparing Fig. 4.17(a) and (b), even a thick substrate contact strap is at the center in Fig. 4.17(a), it is obvious that the height of the circuit with guard ring is longer than that of the circuit without guard ring. Fig. 4.18 shows the layout of circuits for setting the input of flip-flop in Fig. 4.16(b) and (c). All metal layers used in the layout are shown. The circuit in Fig. 4.18(a) has three NAND gates, with two empty spaces near the top-left and bottom-left corner. It is possible to utilize these spaces by putting two of these circuits together, but the interconnection congestion would be worsened because the device placement in the layout does not follow the bit sequence of the counter. There is no empty space in Fig. 4.18(b) because the circuit in Fig. 4.16(c) contains one more inverter.

A simple four-transistor latch is used to shift the voltage level of the counter's output from the core voltage (1 V) to the I/O voltage (2.5 V) as in Fig. 4.19(a). The drawback of this circuit is the asymmetric gate overdrive voltage of NMOS and PMOS. For example, at the moment of V_{IN} rising from 0 V to 1 V, M_{P1} is fully on and its overdrive voltage is 2.5 V minus a threshold voltage while the overdrive voltage of M_{N1} is much lower, 1 V minus a threshold voltage. During transitions, M_{N1} and M_{P1} fight

against with each other in pulling the voltage at their common drain node low. The W/L ratio of M_{N1} must be larger enough so that during transitions the drain current of M_{N1} can be higher than that of M_{P1}. If the W/L ratio of M_{N1} is not big enough, two cases may happen. First, the output of the level-shifter may be stuck in an intermediate undefined voltage level because M_{N1} cannot pull down the gate voltage of M_{P2} low enough to shut down M_{P1}. Second, transition time is long, so the switching power loss is large.

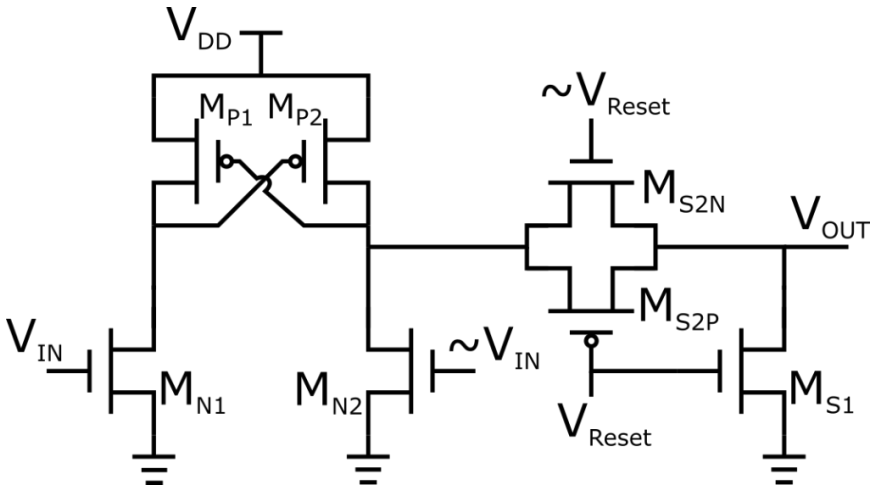


(a) *(b)*

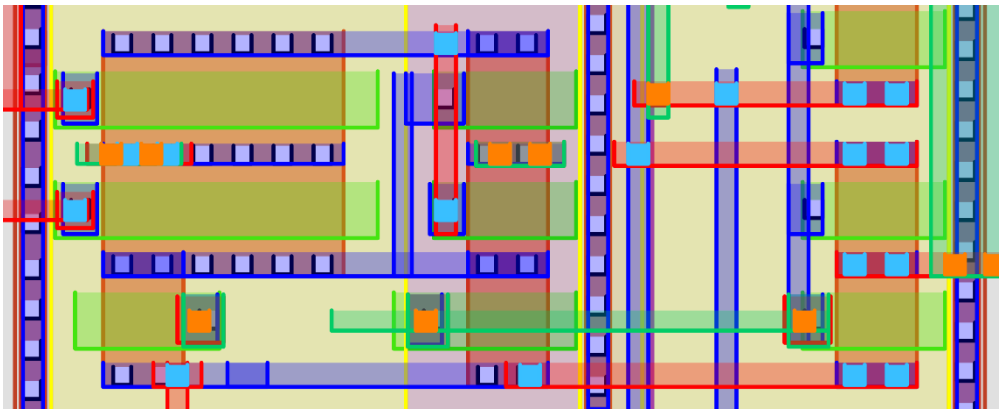
Fig. 4.18 Layout of circuits in Fig. 4.15 (b) and (c) with guard ring (a) three NAND gates (b) three NOR gates and an inverter

A single switch, M_{S1} and a complementary switch, M_{S2N} and M_{S2P} control the voltage at the gate of a driver current source. In the temperature sensing mode, M_{S1} connect the gate to ground to fully shut down the current source so the voltage at the thermocouple junction can be measured. M_{S2N} and M_{S2P} are off to break the connection between the level shifter and current driver. It prevents the output of the level shifter

being pulled to ground when V_{IN} is high. In the heating mode, switch M_{S1} is off and M_{S2N} and M_{S2P} are on. The voltage at the gate of the driver current source depends on the output of the level shifter. Because the output of the level shifter is either 0 V or 2.5 V, a complimentary switch is used to minimize the switch voltage drop for both cases.



(a)



(b)

Fig. 4.19 Level-shifter and switches for driving the gate of driver current source (a) schematic (b) layout

Figure 4.19(b) which is rotated by 90° show the layout of the circuit in Fig. 4.19(a). The level shifter and switch M_{S2N} and M_{S2P} are on the left. As shown in the figure, the width of the NMOSs in the level shifter is three times longer than that of the PMOSs. Three switches, M_{S1}s, for three different bits are on the right.

A trial layout of a control driver group with four units for VR applications is shown

in Fig. 4.20. The sense amplifier and comparator are located at the center. Four units of Level-shifter and 6-bit counter are above and below the sense amplifier and comparator and each side contains two units. Two units of level shifters are placed alternatively along a row to match the width of a 6-bit counter and minimize the distance between the outputs of the counters and their corresponding level shifters' inputs. The width of the 6-bit counter is 24.5 μm and the driver current source arrays fill the remaining width of two pixels, 5.5 μm. Binary encoding of 6-bit can control 63 unit transistors, so each array driver current source consists of 4-by-16 transistors with one dummy. Additional four arrays belonging to the adjacent driver groups are placed on sides to simulate the edge effect of an actual bigger control driver circuit. Because of the LDEs, the characteristic of drain current vs gate-to-source voltage of each unit transistor varies. To minimize the output current variations of each counter step, transistors with current characteristic higher and smaller the average value are grouped to form a higher order bit. This increases the routing density between the level-shifters and the current source unit transistors.

The size of the layout is exactly the total area of 4 NPA pixels, 30 μm-by-30 μm, but the internal devices are not surrounded by guard rings, so the design cannot pass the design rules for latch-up. The height of the two-counter cluster at the top and bottom is 6.6 μm. The height is increased to 7.7 μm if the design with guard ring in Fig. 4.17(b) is used. Thus, a layout which includes guard rings in the counters exceeds the length restriction by at least 2.2 μm and this number will be larger if guard rings are also added to the level-shifters. The layout in Fig. 4.20 also violates the rule of the spacing between the internal PMOSs and n-type OD injectors. The PMOSs in the counters and level-shifters are highlighted in red rectangles. The PMOSs near the left and right boundary of the red rectangles are too close to the array of current source NMOSs which are

connected to I/O pads. Therefore, the placement must be rearranged to increase the distance between the internal PMOSs and current source NMOSs without adding more area overhead.

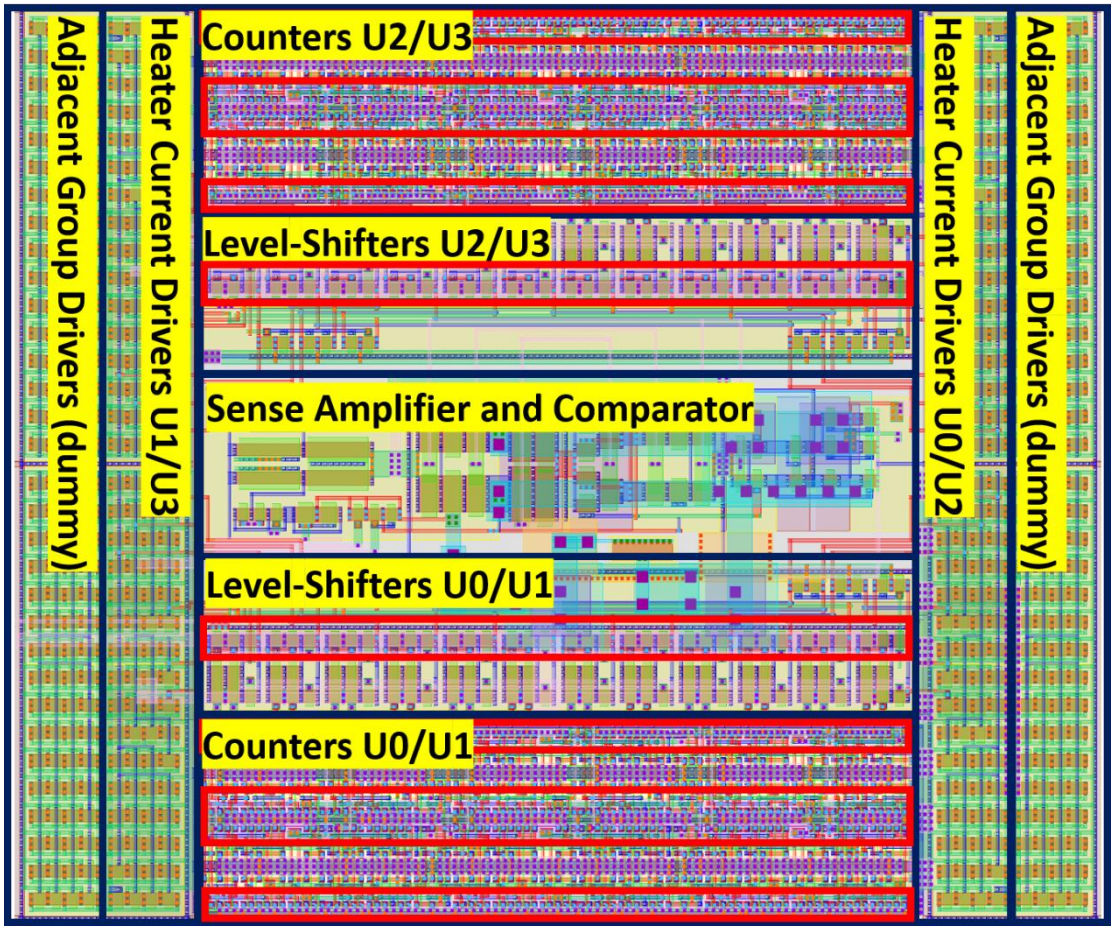


Fig. 4.20 A trial layout of a 2-by-2 control driver group without guard ring

Figure 4.21 shows a trial layout of a single driver unit using the TSMC 28nm technology node. The circuit area is about 12 µm x 12 µm. Unlike the circuit topology of the Fig. 4.20, by taking the advantage of small feature size in 28 nm, this driver control unit can have its own dedicated sense amplifier and comparator and still meet the design area constraint. However, the current driver MOSFET array is implemented by 0.9 V core transistors. Although a further design optimization step of replacing core

transistors with I/O transistors at the current driver MOSFET is needed to improve the power efficiency, this trail layout shows that the proposed mixed-signal gate driver design can meet the circuit area constraint by shrinking the digital counter with a more advance technology node.

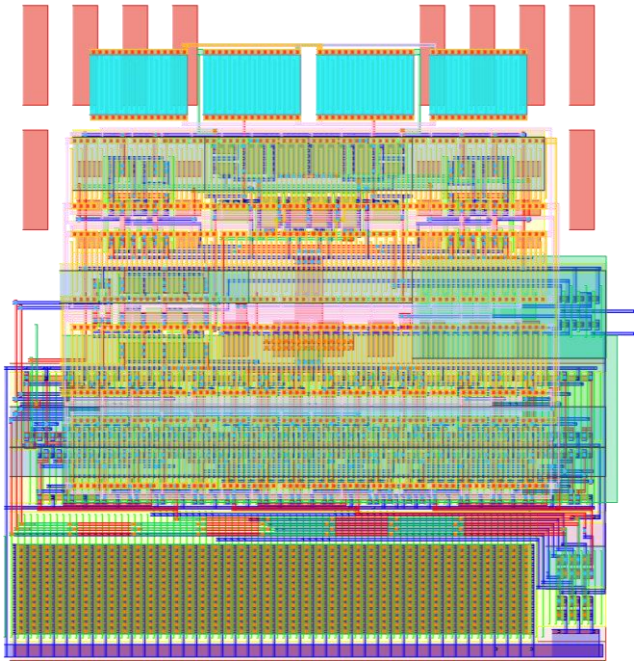


Fig. 4.21 A trial layout of a single control driver group without guard ring in TSMC 28nm node. The circuit area is about 12 µm x 12 µm.

To sum up, the mixed-signal approach for implementing a VR NPA driver group with 4 units in 65 nm technology cannot meet the area requirement. This method is suitable for LIDAR driver because only one counter and one set of level-shifters are required. Other solutions for applying mixed-signal method in a VR NPA driver design include using more advanced technology so the size of counter can be reduced as shown in Fig. 4.21 or having more units sharing the sense amplifier and comparator. Both solutions provide extra room for guard rings and rearranging the placement of the I/O current source NMOSs.

4.2.4.2 Analog Gate Driver

Alternatively, an analog gate driver which does not contain a digital counter has been designed to make the circuit in Fig. 4.2(b) meet the area limitation in 65 nm. The concept of the design is illustrated in Fig. 4.22. The counter is replaced by a big capacitor, C_{state}. The charges stored in the capacitor represents the state of the driver current source, NMOS M_{N1}. Each cycle, depending on the comparator's output, the switch S_1 may be closed for a fixed period so the current source transfers a fixed amount of charge to C_{state}. The function of the switch S_2 and S_3 is same as the switches in Fig. 4.19(a). S_2 connects C_{state} to the gate of driver current source during the heating mode and is open before each temperature measuring mode to prevent C_{state} being discharged to ground through S_1 during. At this period, the upper plate of C_{state} is floated, so the charges in C_{state} is held if the leakage current is small and can be neglected. Thus, the state of the driver current source before entering the temperature measuring mode is recorded.

The main reason that we can't implement the simplified circuit in Fig. 4.22 directly is the charge sharing between C_{state} and C_g when switch S_2 is closed. Because the equivalent capacitor at the gate of the driver current source, C_g, is pulled to ground every cycle for temperature measurement, when S_2 is closed, the voltage across C_{state} is reduced, depending on the ratio of the capacitance of C_{state} to C_g. Assuming C_{state} is m times bigger than C_g, and the current source increases the voltage at C_{state} by ΔV each cycle. The maximum voltage at C_{state} is $(1+m)\ \Delta V$. For example, if m is 100 and ΔV is 10 mV, the maximum voltage at C_{state} is only 1 V and it is too low to fully turn on the driver current source. Clearly, higher maximum voltage at C_{state} can be achieved by increasing the ratio m or ΔV. However, a large NMOS is required to provide enough heating current so C_g is large too. The allowed space prohibits a high m ratio. A high ΔV means big step size which lowers the control precision.

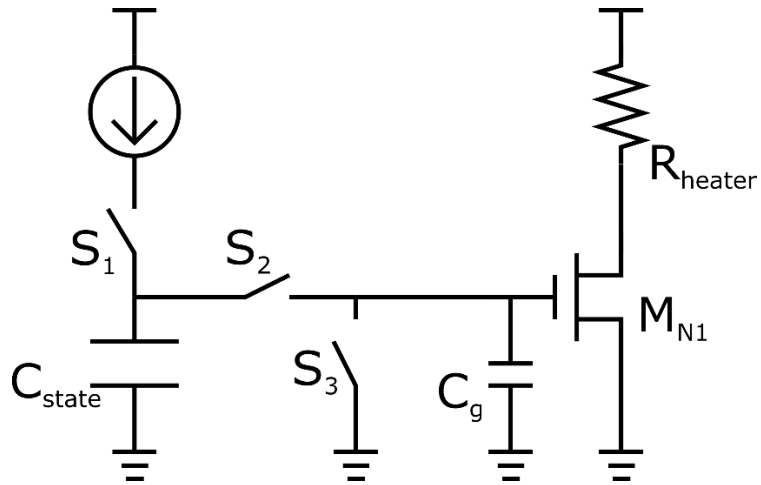


Fig. 4.22 The simplified schematic of the analog gate driver for design concepts illustration

In addition, an actual current source implemented by MOSFET transistors is not able to charge C_{state} to V_{DD}. The driver current source should be at least operated in subthreshold region to draw a current which can raise the temperature at a NPA pixel. The circuit in Fig. 4.22 may waste a number of cycles at the beginning of setting a new target temperature because the voltage at C_{state} is far below the threshold voltage. Thus, a voltage buffer between C_{state} and the driver current source is essential. This buffer isolates C_{state} from C_g so the voltage at C_{state} does not drop when S_2 is close and shifts the voltage at C_{state} approximately one threshold voltage up so the driver can always be partially on.

The complete schematic of the analog gate driver design is shown in Fig. 4.23. In Fig. 4.23(a), M_{P2} and M_{P3} forms a source follower, shifting the voltage V_C up to V_{C_SF} and providing the current to bring V_G from ground to V_{C_SF} when the driver turns on the switches M_{S2N} and M_{S2P}, switching to the heating mode from temperature measuring mode. An important design consideration of the source follower is the transient response, which is the time required to charge V_G from 0 to V_{C_SF} or the slew rate, especially when V_{C_SF} is high. A longer transition time reduces the effective heating

duty cycle so the design requires a higher peak heating current or a larger driver current source to set a 2π phase shift at the NPA pixel. The transition time depends on the bias current of M_{P2} which is the maximum current available for charging the node V_G. The difficulty lies in that the ideal source-to-gate voltage of M_{P3} when V_C is near ground is around 500 mV so the driver current source can be operated at the boundary of subthreshold and saturation region. Therefore, a high bias current at M_{P2} needs the W/L ratio of M_{P3} to be much greater than that of M_{P2}. This contradicts to our primary design goal, a compact driver.

To overcome this issue, the transistors M_{N1}, M_{N2} and M_{P1} set the DC bias current of M_{P2} dynamically. When V_C is far below the threshold voltage of M_{N1}, the bias current of M_{P1} and M_{P2} is dominant by the by the bias current of M_{N2}. For a given transition time, the bias current of M_{P2} can be smaller when V_C is low. Hence, in this operating region, V_{B1} sets a minimum current flowing through M_{P2} to minimize the size of M_{P3} that shifts V_{C_SF} about 500 mV above V_C and satisfies the speed requirement. When V_C rises above the threshold of M_{N1}, M_{N1} starts to draw a current and pulls down the gate voltage of M_{P1}. When all transistors are in the saturation region, the bias current of M_{P2} increases quadratically with V_C so the transition time can be reduced even a larger voltage swing is at V_G.

A design concern is that the dynamic bias current may increase the variation of the amount of voltage shifted because intuitively the difference between V_{C_SF} and V_C increases as the gate voltage of M_{P2} decreases so. Fortunately, the variation can be compensated partially by the body effect of M_{P3}. When M_{P2} and M_{P3} are in the saturation region, the bias current at steady state can be expressed as:

$$\frac{1}{2}K\left(\frac{W}{L}\right)_{M_{P2}}\left(V_{SG\_M_{P2}}-V_{TH\_M_{P2}}\right)^2=\frac{1}{2}K\left(\frac{W}{L}\right)_{M_{P3}}\left(V_{SG\_M_{P3}}-V_{TH\_M_{P3}}\right)^2 \qquad (4.5)$$

$$\rightarrow V_{C\_SF} - V_C = \sqrt{\left(\frac{W}{L}\right)_{M_{P2}} \Big/ \left(\frac{W}{L}\right)_{M_{P3}}} \left(V_{SG\_M_{P2}} - V_{TH\_M_{P2}}\right) + V_{TH\_M_{P3}}$$

A constant up-shift voltage means the equation (4.5) does not vary with the dynamic bias setting. For a first order estimation, assuming the threshold voltage of M_{P2} does not vary with the gate voltage of M_{P2} and the ratio of (W/L) of M_{P2} to M_{P3} is m, we want to satisfy:

$$\Delta V_{SG\_M_{P2}} + \sqrt{m}\, \Delta V_{TH\_M_{P3}} = 0 \tag{4.6}$$

The threshold voltage of M_{P3} is at its maximum when V_C is 0. As V_C increases, V_{C_SG} increases so the body-to-source voltage of M_{P3} and the body effect is reduced. The change of the threshold voltage of M_{P3} is indeed opposite to the source-to-gate voltage of M_{P2}. By adjusting the size ratio of M_{P2} to M_{P3}, the relation in equation (4.6) can be realized.

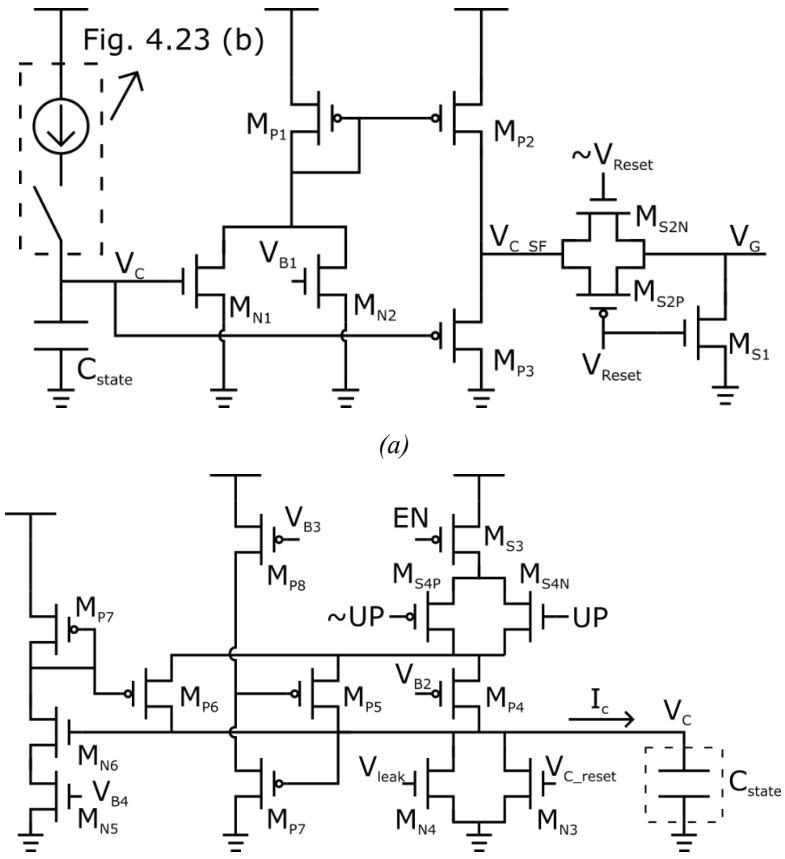


Fig. 4.23 The complete schematic of the analog gate driver (a) source follower buffer (b) current source for charging C_{state}

The post-layout simulation results of the dynamic bias point setting are shown in Fig. 4.24. A constant voltage source, V_C, is connected to C_{state}. Fig. 4.24(a) to (b) show the results of DC sweeps of V_C from 0 to 2.5 V. In Fig. 4.24(a), the output voltage of the source follower is about 510 mV when V_C is 0 as designed. The gate voltage of M_{P2} is nearly constant around 1.95 V when V_C is below 0.6 V because M_{N1} is off. When V_C is above 0.6 V, M_{N1} starts to conduct a current and pulls the gate voltage of M_{P2} down. For V_C from 0.6 to 1.9 V, the gate voltage reduces roughly 500 mV. The amount of the voltage shifted by the source follower is shown in Fig. 4.24(c). It decreases at the beginning from about 510 mV to 470 mV when Vin is less than 0.6V because of the reduction on the threshold voltage of M_{P3} for less body effect. For V_C from 0.6 V to 1.9 V, because of the reasons mentioned above, the body effect reduction of M_{P3} and the size ratio of M_{P3} to M_{P2}, the variation of the shifted voltage is less than ±30 mV. It shows the dynamic biasing current control circuit for M_{P2} only has a minor effect on the accuracy of the source follower.

Figure 4.24(d) shows the bias current of the source follower. The current starts from less than 100 nA and increases about two orders of magnitude when V_C is around 1.8 V. Fig. 4.24(e) to (h) show the transition time for different value of V_C. For the simulation setup, the switch S_2 is turned on at 21.7 μs to charge the gate of the current source driver. The output of source follower is pulled down deeply to the gate voltage of the driver current source at this moment. When V_C is 0 V, the transition time for the source follower charging its output and the gate of the driver current source to its final value is about 400 ns. The transition time increases to 1.1 μs when V_C is 0.5 V, because the bias current of the source follower remains the same (M_{N1} is off) but the final value becomes 1 V instead of 510 mV. If the bias current is fixed, when V_C is 2 V, a rough estimate of the transition time is 3.2 μs which is more than 30 % of the heating time per

cycle, 9 μs.

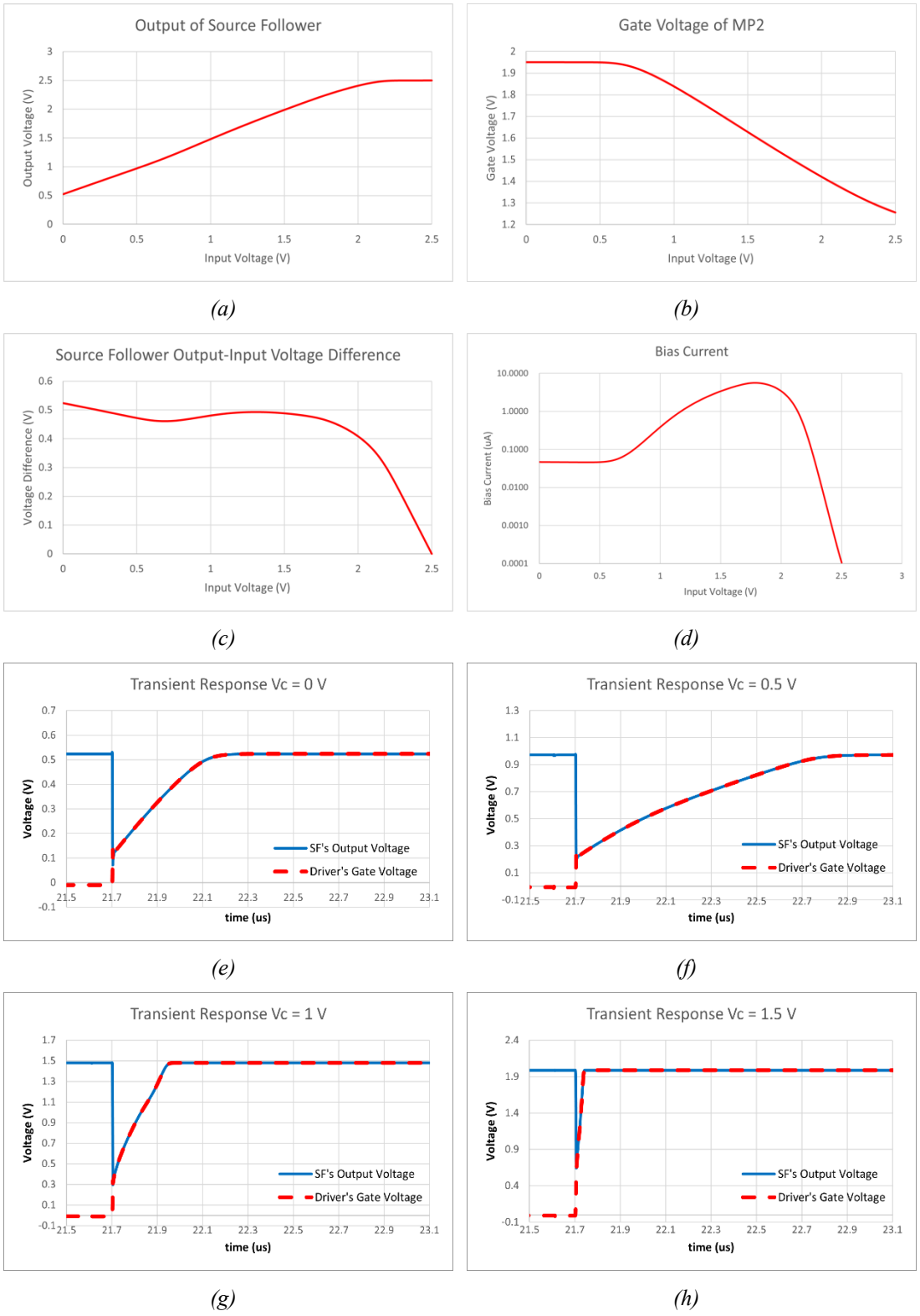


Fig. 4.24 Post-layout simulation result of the gate driver's dynamic biasing source follower

The transition times in Fig. 4.24(g) and (h) are a little more than 200 ns and

less than 50 ns respectively. The results show the dynamic biasing source follower reduces the transition time significantly. The reduction is more for a high input voltage, V_C, because the bias current increases quadratically with V_C while final value only increases linearly.

The implementation of the current source for charging C_{state} is shown in Fig. 4.23(b). M_{N3} is on at the beginning of every new frame to reset C_{state}. M_{S3}, M_{S4N} and M_{S4P} are the switches to control the charging of C_{state}. M_{S3} is turned on for a fixed period every cycle and M_{S4N} and M_{S4P} is controlled by the comparator's output. When these three switches are all on, the source terminals of M_{P4}, M_{P5} and M_{P6} are tied to V_{DD} and a current is drawn from the power supply through M_{P4}, M_{P5} and M_{P6}. M_{N4} conducts a current to balance the leakage current from the PMOSs, preventing C_{state} from being charge to V_{DD}. To reduce the circuit area, only current sources for charging C_{state} are used. Thus, the current of M_{N4} is set slightly higher than the leakage current. When the C_{state} is charged above the target value, M_{N4} can pull down the voltage at C_{state} slowly and passively. Alternatively, it is possible to insert two more switches in series with M_{N4} and controlled by EN and the counter's output, so the circuit is able to discharge C_{state} actively when the temperature at the controlled NPA pixel is too high.

The reason for using three transistors instead of a single transistor to charge C_{state} is to compensate the high non-linearity relationship between the gate voltage of the heater current driver and the temperature at the NPA pixel. The amount of the current flowing through M_{P4} is set by V_{B2} and does not vary with the voltage across C_{state} while the currents flowing through M_{P5} and M_{P6} change dynamically with voltage across C_{state}. In saturation region, the current of the heater current driver increases quadratically with its gate voltage. From the thermal simulation results, the temperature at the NPA pixel increases roughly quadratically with the heater current if the variation of heater's

resistance due to the temperature is neglected. Thus, the temperature at the NPA pixel is fourth power related to the gate voltage of the heater driver. If a constant current source is used to charge C_{state}, the step size is much smaller when the voltage at C_{state} is small. The gate of M_{P5} is set by the voltage at C_{state} after the source follower M_{P7} and M_{P8} and is conducted when V_C is small. M_{P5} is cut off as V_C increases and V_{B3} sets the cutoff voltage.

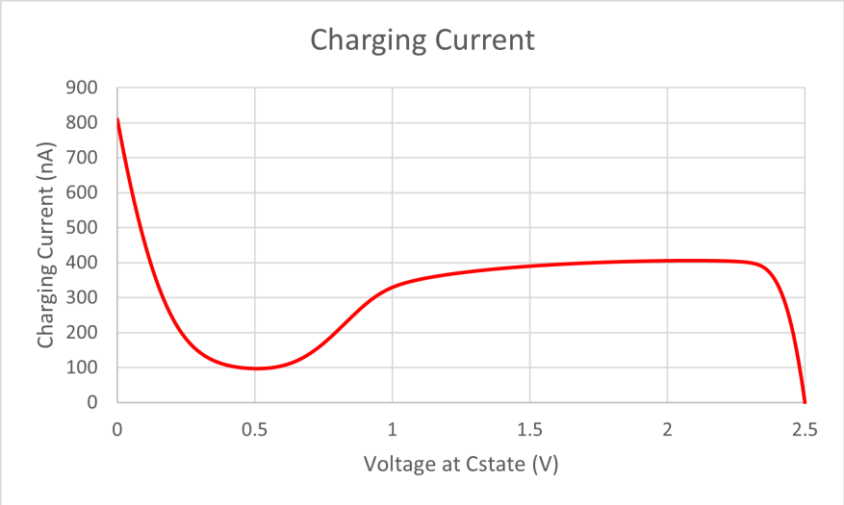
When V_C is too high, the heater current driver enters the triode region and Fig. 4.24 shows the difference between V_{C_SF} and V_C decreases. Thus, in this region, the relation between the gate voltage and current of the heater current driver is flattened. The current for charging C_{state} should be increased to maintain the step size. When V_C rises above the threshold of M_{N6}, M_{N5} and M_{N6} starts to draw a current from M_{P7} and this current is mirrored to M_{P6}. V_{B4} controls the slope of the charging current with respect to V_C and the saturated current value. Thus, more current is available for charging C_{state} when V_C is high.

Post-layout simulation results of the C_{state} charging circuit are shown in Fig. 4.25. Fig. 4.25(a) shows the total charging current with respect to V_C when V_{B3} and V_{B4} are 1.2 V and 0.8 V respectively. The current starts at 800 nA and it drops to 100 nA when V_C is 0.5 V. The current increases gradually when V_C is higher than 0.6 V and it reaches 400 nA when V_C is 2 V. A flat region at the bottom can be extended by inserting a diode-connected NMOS between M_{N5} and M_{N6}, so V_C has to be higher to turn on the current mirror M_{P7} and M_{P6}.

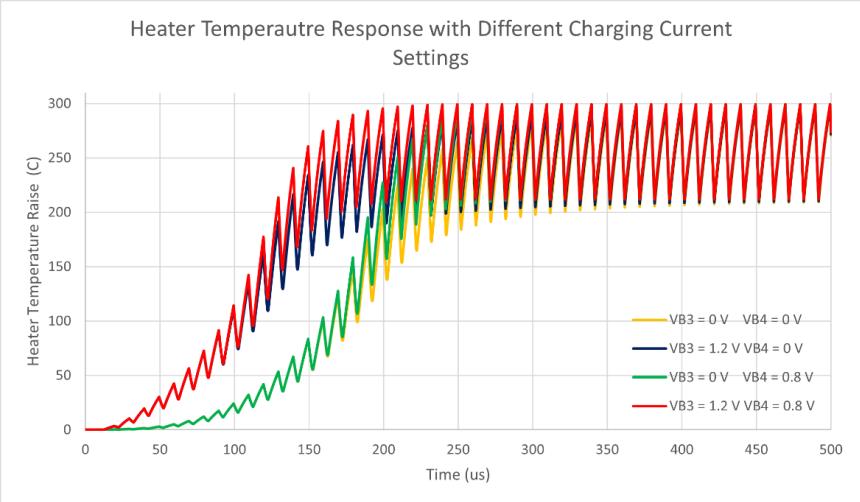
Figure 4.25(b) shows the transient response of the temperature at the heater under different settings of the charging current when the input reference voltage of the comparator is set to fully turn on the heater current driver. C_{state} is charged for 200 ns every cycle. When V_{B3} is 0 V, M_{P5} is off for all V_C. The results show that without M_{P5}

being on, the driver takes more than 6 cycles before the temperature at the heater starts to increase significantly each cycle. With V_{B4} equal to 0.8 V, the temperature at the heater reaches the steady state at 220 μs and 280 μs for V_{B3} equal to 1.2 V and 0 V respectively while with V_{B4} equal to 0 V, the temperature at the heater increases at a very low rate when the temperature is higher than 200 °C and requires more than 400 μs to reach the 99% of the temperature at the steady state.

To sum up, the charging current compensation circuit is essential as it speeds up the transient response and reduces the step size variation at both ends.



(a)



(b)

Fig. 4.25 Post-layout simulation result of the gate driver's charging current compensation circuit

The layout of a 2-by-2 NPA driver group using the analog gate driver is presented in Fig. 4.26. The sense amplifier and the comparator are located at the center, same as the placement of the trial layout in Fig. 4.18. Four heater driver state capacitors are placed at each corner. To minimize the influence of switching noise on the states capacitor's voltage, the capacitance is maximized by placing a MIM capacitor above a M1-M6 MOM capacitor and connecting them in parallel. The nominal capacitance is 340 fF. Instead of placing the heaters drivers in narrow and high regions, 2.8 μm-by-15 μm, the dimension of the new layout is 4 μm-by-5.5 μm. They are located at the center of top and bottom boundary and between the state capacitors. The area of the heater driver is reduced because the multiple lines of unit transistors are replaced by a single interdigitated transistor. The area overhead of vertical spacing between different lines is eliminated.

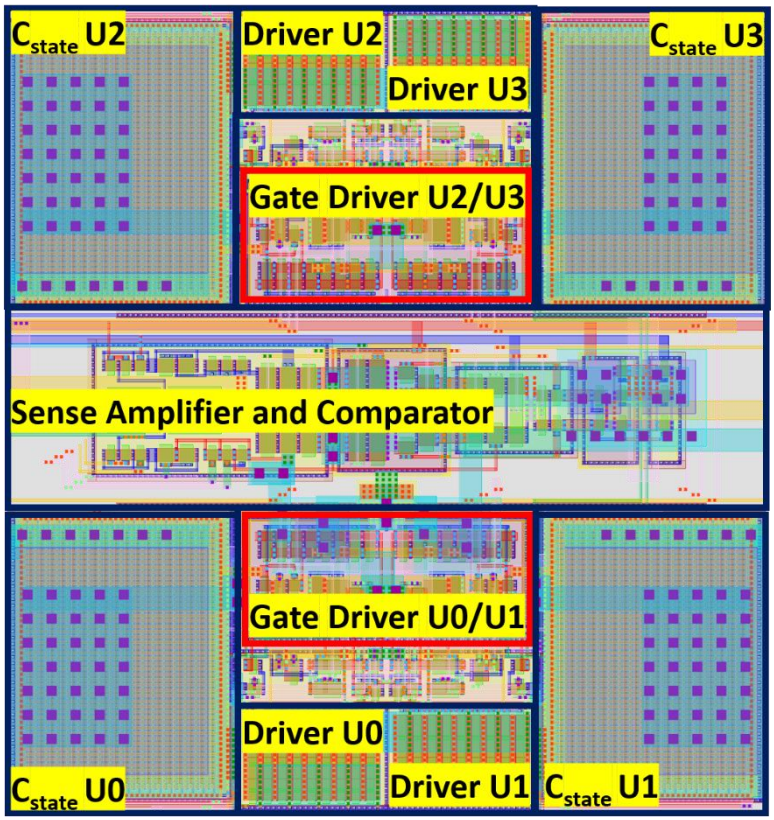


Fig. 4.26 The layout of a 2-by-2 NPA driver group using the analog gate driver

The gate drivers are next to the heater drivers vertically. The NMOSs and PMOSs of two gate drivers are grouped in two clusters respectively. The width of the clusters matches the sum of the width of two heater drivers. The two red rectangles in Fig. 4.26 are the PMOS clusters of the gate drivers. The spaces between the PMOS clusters and heater driver I/O transistors required by the latch-up prevention design rules are filled with the NMOS clusters so no extra area overhead is introduced. All transistors are surrounded by their corresponding guard rings. This design passes all design rules. Comparing to the mixed-signal driver design, the size of the heater driver is smaller, and the design does not need level shifters for transforming core voltage of the counters' outputs to I/O voltage. More space is saved for analog gate drivers and state capacitors.

The state capacitors' voltage and the gate voltage of the heater drivers determines the pixel temperature. It is important to reduce the interference between the switching signals and these two nodes. Since the driver circuit is compact, the circuit area does not allow to reduce the parasitic coupling capacitors between nodes by increasing the spacing between their interconnections. Shielding metal lines and plates are used to eliminate the coupling parasitic capacitors. Fig. 4.27 shows an example of removing the coupling capacitor between the node of the capacitors upper plate and the comparator's outputs.

The comparator's outputs are connected to the gate driver using vertical M3 lines at the center of Fig. 4.27. The upper plate of C_{state} is connected to the source follower in the gate driver by a horizontal M6 line which is running above the M3 lines of comparator's outputs. Two vertical grounded metal lines from M2 to m5 layers are placed between the parallel sections of the M3 lines of the comparator outputs and C_{state} upper plate connection metal where the fringe capacitors formed. Multiple layers are used because parallel metal pieces in close proximity and adjacent layers can also form

fringe capacitors. A grounded M5 plate is inserted between the overlapped region of the comparators' outputs and the horizontal M6 connection line of C_{state}'s upper plate. Parasitic coupling fringe and overlapped capacitors are removed at the expense of having larger parasitic capacitors to ground for both nodes. For nodes with more complex interconnection metal lines, the above two shielding methods may not be able to entirely remove coupling parasitic capacitors.

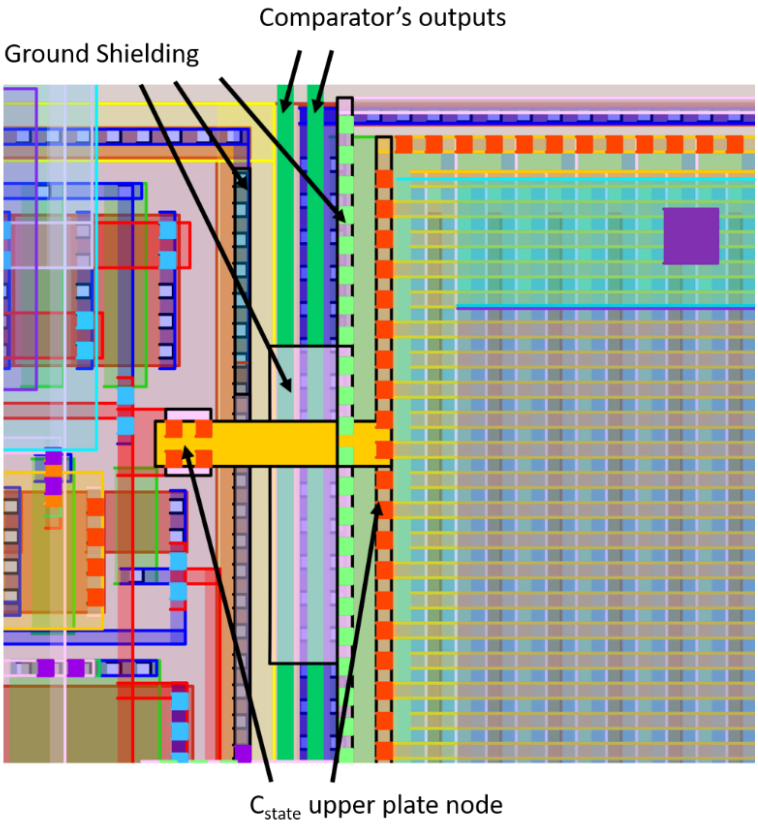


Fig. 4.27 Example of coupling parasitic capacitor elimination

4.3 Layout Implementation with Deep N-well

The layout implementation described above violates many standard recommended layout rules of analog circuit, e.g., placing a MIM capacitor above some noise sensitive nodes or circuits having large output swing, or failing to isolate the analog and digital circuit properly, due to the strict area limitation. To follow these recommended rules, it is inevitable that the area overhead of the circuit will increase. The driver system needs to have more units per group to accommodate the additional area overhead.

For a VR NPA driver with the analog gate driver architecture, including more units in a group does not offer much more space saving because the gate driver, the state capacitor and the heater driver which are required by each individual unit occupy most of the area. In contrast, for a LIDAR NPA driver using a digital counter, only the heater driver is required for each driver unit. Gathering more units in a group allows much more space for implementing the shared components.

Figure 4.28 shows the layout of a 16-unit LIDAR NPA driver circuit. All NMOS clusters are placed in separate deep n-well and all transistor clusters are enclosed by double guard rings for best performance on noise isolation and latch-up prevention. The capacitor in the integrator sense amplifier is moved to a free space so no circuit is beneath it. The total number of the counter digit is 8. The additional circuit required is some buffers because a single level-shifter in Fig. 4.19 is not capable of driving 16 units of the higher bits of the heater driver unit transistor array. Because the output signals to the gates of the heater drivers are digital, the buffer circuit can be implemented by inverters which is extremely simple and space efficient. The circuit area is about 3000 μm^2 which is less than the area of the sum of 16 NPA pixels. Thus, the temperature at a LIDAR NPA can be sampled at least every 16 pixels with a mixed-signal driver implementing in deep n-wells and containing double guard rings.

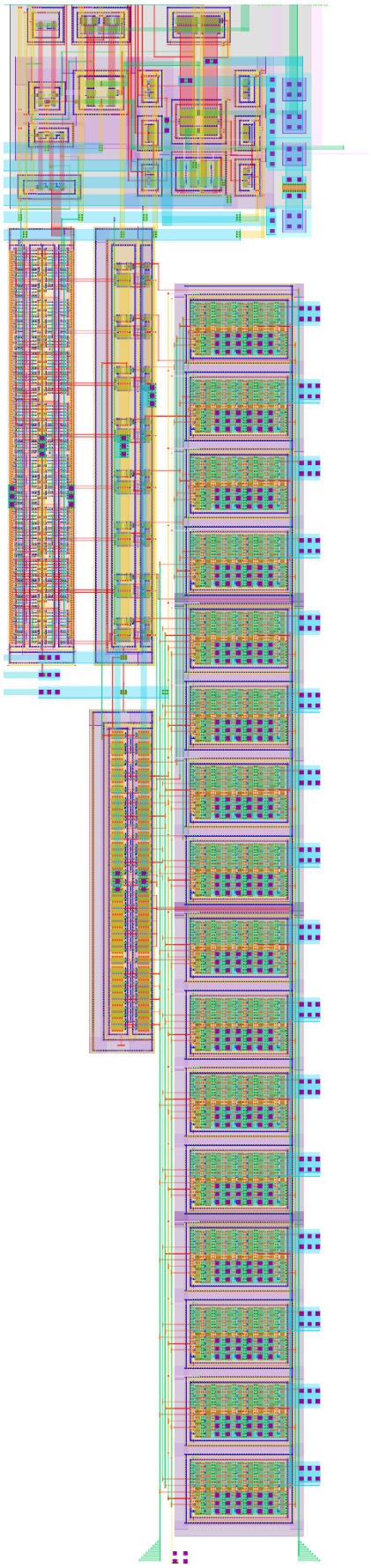


Fig. 4.28 The layout of a 16-unit LIDAR NPA driver circuit in deep n-well

4.4 Chip Level Layout

We used TSMC 65 nm I/O library to create the I/O ring and bonding pads of the testing chip for validating our driver design. The I/O library is not included in the PDK. The library provides power and ground I/O cells for all combinations of digital/analog and core voltage/IO voltage domain. Our design contains many switch signals, a counter operating at core voltage and analog circuit components operating at I/O voltage. Ideally, these signals should be grouped and assigned to I/O pads in different power domains. A power-cut cell is required to isolate two different power domains, and additional power and ground I/O cells should be placed right next to a power-cut cell both sides for robust ESD and latch-up protection. This reduces the available signal pads for probing the internal nodes of the driver circuit. We only implemented an analog I/O voltage power domain in our I/O pad frame to maximize the number of available signal pads. Core voltage is supplied to internal core transistors by connecting a 1 V power source to a standard signal analog I/O pad. A buffer made by I/O transistors is inserted between the core transistors and I/O pad.

The chip level layout is provided in Fig. 4.29. The size of the layout including the seal ring is 1035 μm-by-935 μm. The size of the bonding pads is 50 μm-by-60 μm and are positioned in two rings. The outer ring and the inner ring are staggered so the equivalent pitch size of the bond pads is 50 μm. The total number of the bonding pads for power supply, ground and global ESD ground bus are 16 and the rest of the 36 pads are signal pads for input and output signals, and DC bias points.

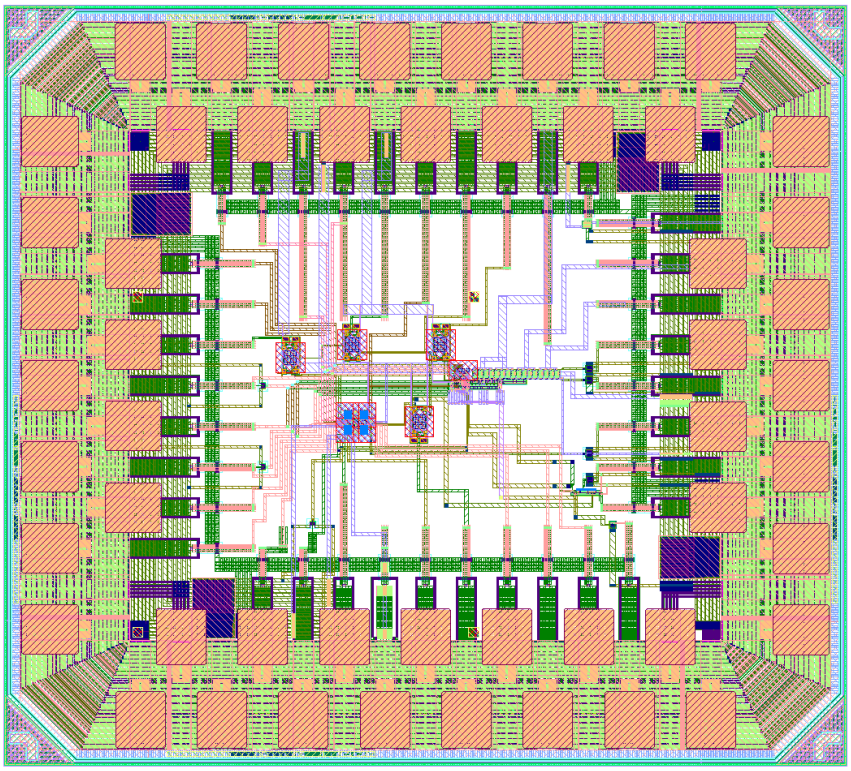


Fig. 4.29 Chip level layout of a VR NPA driver and a LIDAR NPA driver

In conclusion, this chapter provides a thorough description of a novel and compact NPA driver circuit with a broken-loop feedback temperature regulation mechanism. The analog driver circuit meets the area requirement matching the size of a NPA pixel, 15 µm x 15 µm. Thus, this driver can be applied to both VR and LIDAR applications. On the other hand, the mixed-signal driver circuit exceeds the area constraint in 65 nm technology but is still useful for a LIDAR application which does not require independent phase control for every pixel. With more advanced technology such as 28 nm, the gate driver digital counter can be shrunk so this circuit topology may be used for VR applications too. The advantages, disadvantages, challenges, circuit operations and analysis of two design approaches, mixed-signal and pure analog, are discussed in detailed. Circuit layouts and post-layout simulation results of each circuit component for design validation are presented.

Chapter 5: Devices Experimental Measurement and Performance Evaluation

In this chapter, the fabrication steps and characterization results of our thermocouple samples are provided. This is followed by a presentation of the experimental results obtained from characterization of our electrical control circuit fabricated at TSMC on their 65 nm technology product line. I conclude with a demonstration of the proposed broken-loop feedback control method.

Contributions described in this chapter include:

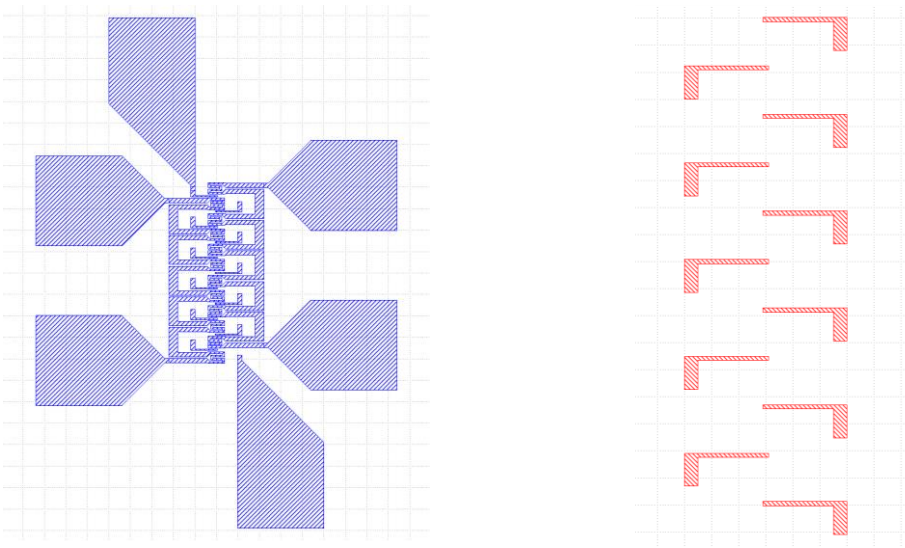
- Characterization of the thermo-electrical property of a nano-scale thin film heater/thermocouple sample by fitting the measuring data with simulation results.
- Electrical measurements illustrating the functionality of each component in the broken-loop feedback driver circuit
- Verification of the broken-loop feedback control method using a sense resistor to simulate the temperature feedback signal. The thermally related phase error was simulated by varying the load resistance ±20% and the sensing resistor ±10%. The maximum and average phase error is reduced to less than 3.6% and 1%, respectively.
- A presentation of the comprehensive phase error analysis. We demonstrate that the broken-loop feedback reduces the thermally introduced phase error (both proximity effect and temperature gradient) effectively. The control method on average increases the SSIM from 0.5 to 0.9 in image applications and maintains the beam power and pointing accuracy in beam steering.

5.1 Thermocouple Sample Experiment

5.1.1 Fabrication Steps Used to Make the Thermocouple Sample

The thermocouple samples are made by two steps of metal deposition: 80 nm of gold and 100 nm of nickel. The metal patterns of each step are shown in Fig. 5.1(a) and (b). The first gold layer includes most of the sample real estate: the gold segments of the thermocouple, two serpentine heaters and six 200 μm-by-200 μm bonding pads. The second layer only includes the nickel segments of the thermocouple. The length of the overlaps of gold and nickel layers creating the thermocouple junctions along the center vertical line is 4.

The metal layers were deposited on a 10 μm thick layer of thermal oxide coating a silicon wafer. The thick silicon dioxide layer between the metal layers and the substrate provides thermal insulation and reduces the power requirement for raising the junction temperature. Both metal layers were patterned by electron beam (e-beam) lithography to achieve a sub-micron scale, ultra-fine structure.



(a) *(b)*

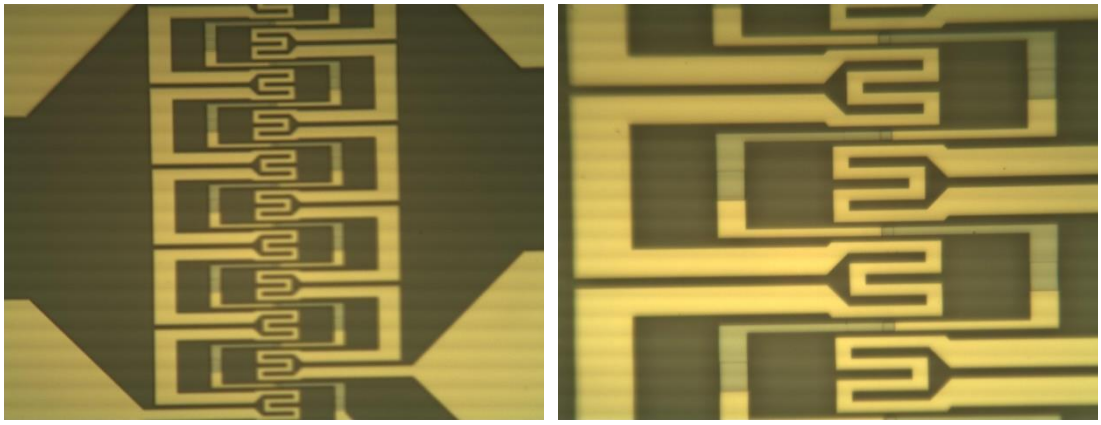
Fig. 5.1 Thermocouple sample metal layer patterns (a) gold layer and (b) nickel layer

Two layers of photoresist were coated for the e-beam lithography. The first layer of the photoresist was ZEP520 and was spun at 4,500 rpm for 60 seconds. Next, the wafer was baked at 180 °C for 300 seconds before the coating of the second layer. The second layer of the photoresist was aquaSave and was spun at 4,500 rpm for 45 seconds. AquaSave creates a conductive layer to prevent electrons from accumulating as a charge sheet on the surface during beam writing, destroying e-beam resolution.

After e-beam lithography, the photoresists were developed in 4 steps. First, the sample was soaked in water to remove aquaSave from the surface. Next, the sample was developed in ZED-N50 for 60 seconds and was immersed into 1:3 MIBK-to-IPA solution immediately thereafter 30 seconds to stop the photoresist development. Finally, IPA was used to wash out the development stopper residue on the sample surface.

Both metal layers were deposited using an Angstrom brand e-beam evaporator. A thickness of 5 nm chromium layer was added beneath the gold layer as adhesive layer. The deposition rate was kept at 0.1 nm per second. The sample was soaked in remover PG to lift off the photoresists. The lift-off process was expedited by using an ultrasonic bath.

Two light-microscope images of the thermocouple sample are shown in Fig. 5.2. The images show that the two layers of metal are well-aligned and the thermocouple junctions are well-formed.



(a) *(b)*

Fig. 5.2 The thermocouple sample under microscope

In order to probe and connect the devices, Au/Sn submounts were glued on the edges of the sample. 200 μm-by-200 μm pads were connected to the submounts by gold wire bonding. Wires were soldered on the submounts so the sample could be hooked to a power supply or the control circuit. Fig. 5.3 shows the finished thermocouple sample.

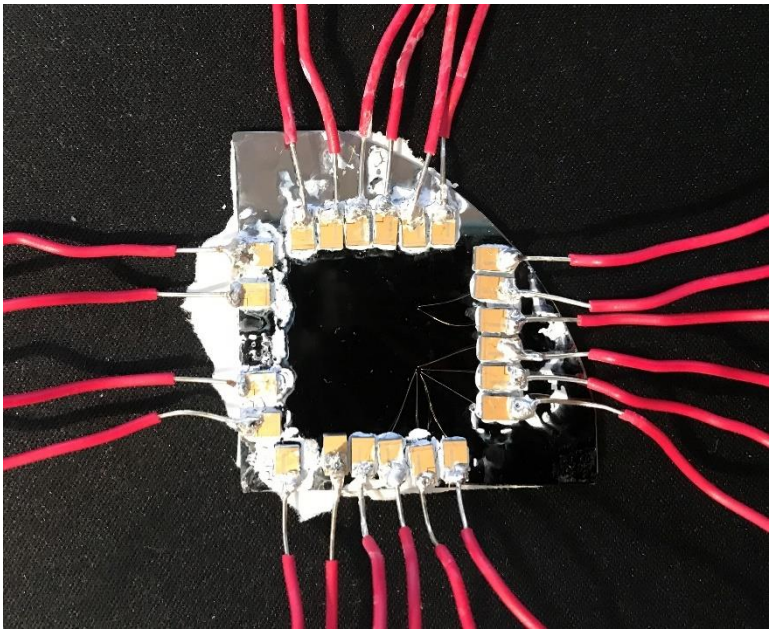


Fig. 5.3 A photo of the finished thermocouple sample

5.1.2 Thermocouple Sample Measurement Results

The main goal of this experience is to measure the thermocouple voltage response

and to experimentally determine the Seebeck coefficient for the material used. As stated before, the challenge is that the temperature at the thermocouple junctions cannot be accurately measured directly with a commercially available IR camera. This is because the size of the junctions is much smaller than the camera's resolution. Two methods were used to estimate the sample's Seebeck coefficient by relating the sample thermocouple's output voltage to: (i) the temperature reading from an IR camera directly and (ii) the average temperature of a 20 μm x 300 μm from the COMSOL thermal simulation as shown in Fig. 3.24(b). The comparison of the results from these two approaches will be presented.

First, the resistances of the two heaters and the thermocouple at room temperature were measured which are 143 Ω for both heaters and 350 Ω for the thermocouple. Next, the sample were placed on a hot plate and the temperature at the hot plate surface was varied from 30 °C to 100 °C. Since the sample is thin and small compared to the hot plate surface, the sample temperature is assumed to be uniform and the same as the temperature at the hot plate surface. The result is shown in Fig. 5.4. The resistance of the heater increases to 160 Ω at 100 °C. Because the resistivity of chromium is higher than that of gold and the chromium layer of the heaters is much thinner than the gold layer, the conductivity contributed by the chromium layer can be neglected. Using the FEM simulation data and the heater resistance measurement result, the resistivity of 80 nm of gold in our sample can be approximated as,

$$(0.00533*T+2.2208) \times 10^{-8}\ \Omega\cdot m \qquad (5.1)$$

where T is the absolute temperature. At 300 K, the resistivity is 3.82×10^{-8} Ω·m which is higher than the value in bulk form, 2.44×10^{-8} Ω·m, as expected.

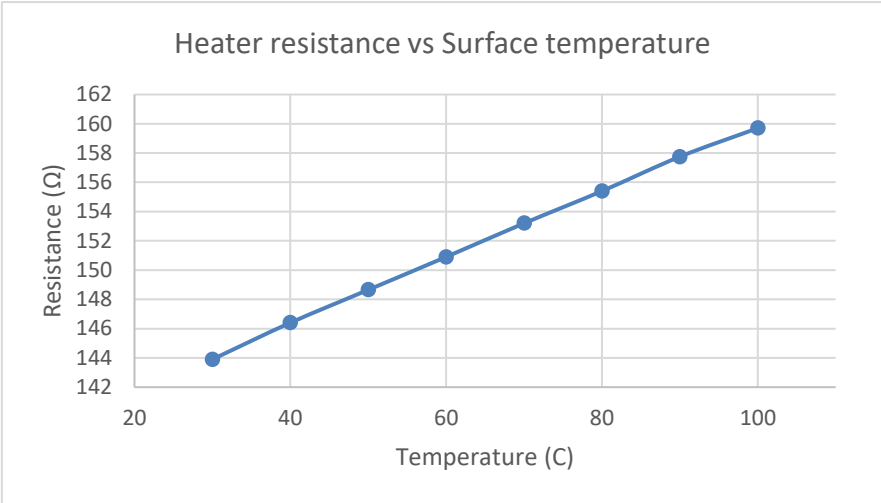
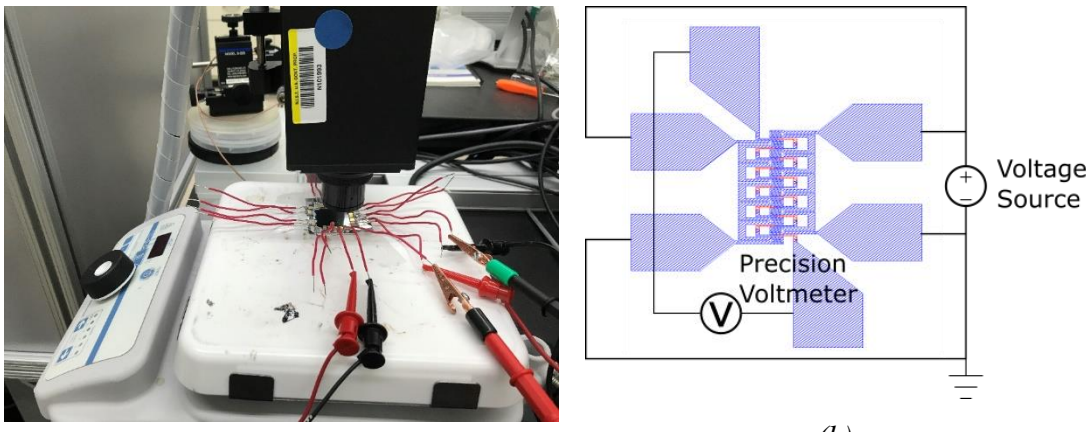


Fig. 5.4 Temperature dependence of the heater resistance

The experimental setup for the Seebeck coefficient measurement is shown in Fig. 5.5. The sample was again placed on a hot plate which was used for emissivity calibration and an IR camera was mounted over the sample. The two heaters were connected in parallel to an external voltage source and the thermocouple was connected to a precision voltmeter to read the thermoelectric response.

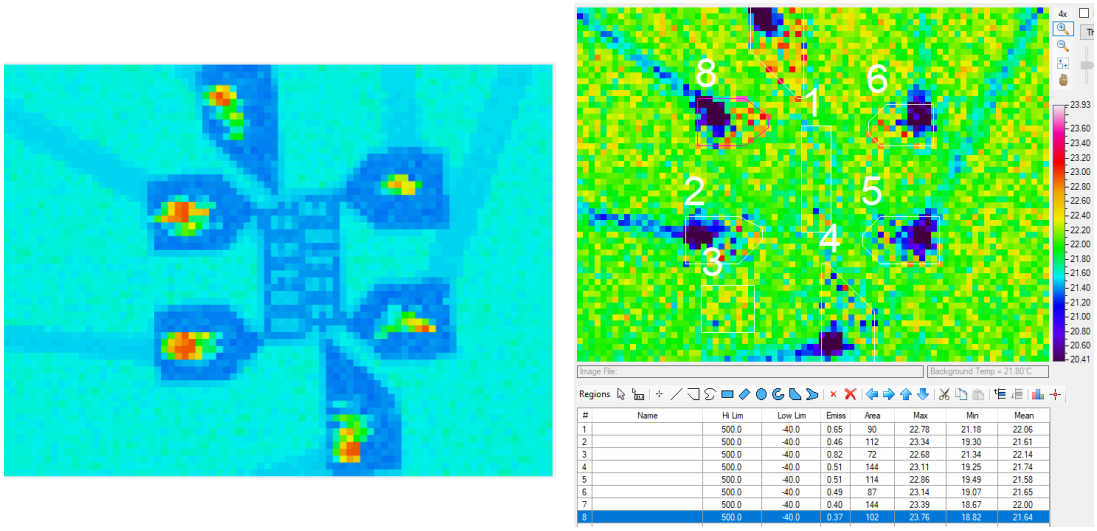


(a) *(b)*

Fig. 5.5 Experiment setup of Seebeck coefficient measurement (a) a photo of the devices setup (b) a schematic of connection between the sample and the instruments

To have accurate temperature readings from the IR camera, emissivity calibration for the sample was necessary. This was performed as follow. The temperature at the hot plate surface was set to a high temperature first. At high temperature, the contrast of IR

camera images due to the emissivity difference of different materials on the sample at uniform temperature distribution is large, so the sample images from the IR camera are clearer at high temperature than at ambient temperature. Fig. 5.6(a) shows a sample IR camera image at 60 °C without emissivity calibration. Eight regions on the image were created, one at the center which enclosed all the thermocouple junctions along the center line, six over the bonding pads and one at an empty space closed to the devices, as shown in Fig. 5.6(b). Except the first region, the rest of the regions contain only single material. After the sample was cooled down to ambient temperature, the emissivities at these regions were adjusted so the temperature readings from the IR camera at these regions coincide with the ambient temperature. Fig. 5.6(b) is an IR camera image of the sample after emissivity calibration. The structures of the thermocouple and heaters can be barely seen in the image comparing to Fig. 5.6(a).



(a) *(b)*

Fig. 5.6 Sample IR camera images (a) at 60 °C without emissivity calibration and (b) at ambient temperature with emissivity calibration

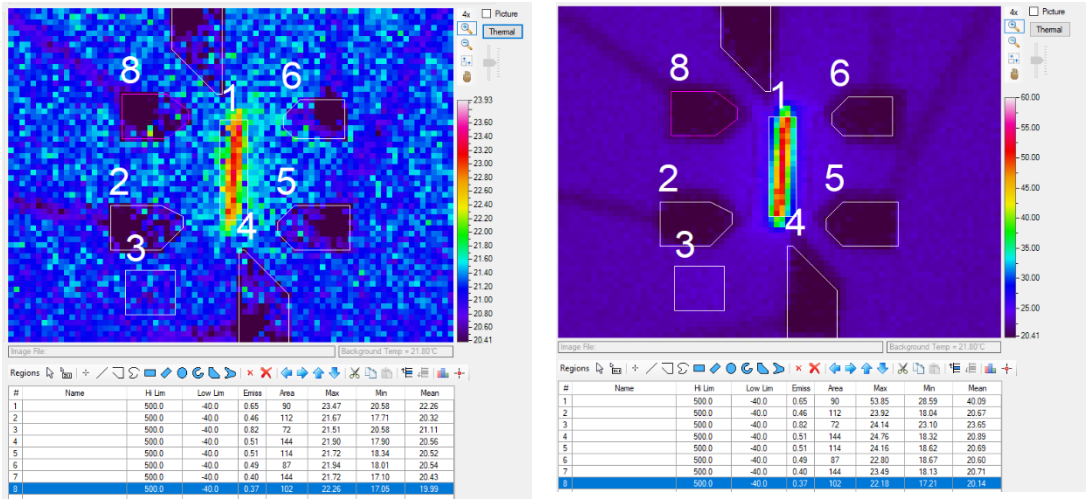
After the calibration, the voltage applied to the heaters was varied from 0 to 5 V. The voltage across the thermocouple, the input current to the heaters and the

temperature readings from the IR camera were recorded at each step. The heaters are designed to raise the temperature at the region around the center line near uniform so the temperature variation can be observed from the IR camera. The thermal images of the sample with 1.4 V and 5 V across the heaters are illustrated in Fig. 5.7. The images show that the heat is concentrated along the center line as designed.

The measurement results are summarized in Fig. 5.8. Fig. 5.8(a) shows the voltage across the thermocouple with respect to the input voltage applied to the heaters. The quadratically increasing curve proves that the output voltage of the thermocouple results from the thermoelectric response. The thermocouple's output is expected to be linearly proportional to the difference between the temperature at the junctions and the ambient temperature which is linearly related to the power dissipated by the heaters. The output voltage of the thermocouple reaches 11.5 mV at 5 V input.

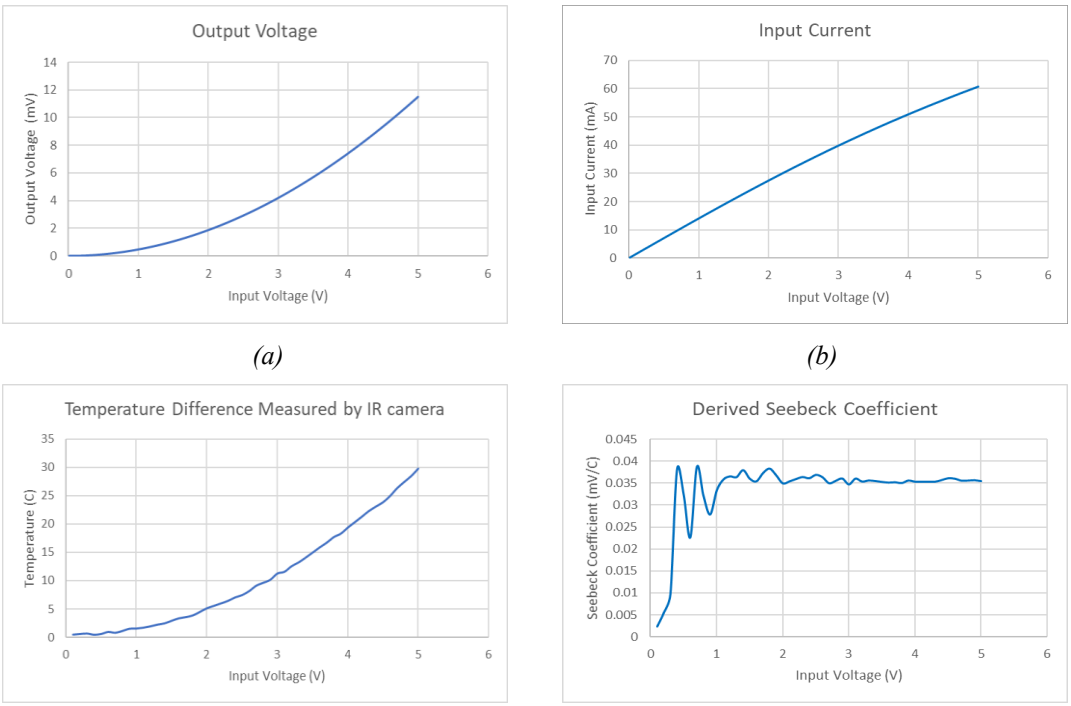
Figure 5.8(b) shows the total current flowing through the two heaters. The current decreases with the input voltage because the heaters' resistance increases with the temperature. The resistance increases from 142 Ω to 164.5 Ω for an input voltage varying from 0 V to 5 V.

Figure 5.8(c) shows the amount of the temperature that contributes to the thermoelectric effect with respect to the heaters' input voltage. The value is assessed by the difference between the maximum temperature in region 1 and the ambient temperature which is the average temperature in region 3.



(a) *(b)*

Fig. 5.7 Sample thermal images with (a) 1.4 V input and (b) 5 V input



(a) *(b)*

(c) *(d)*

Fig. 5.8 Sample measurement results (a)thermocouple output voltage, (b) heater input current, (c) sample temperature recorded by IR camera and (d) derived Seebeck Coefficient from the IR camera results

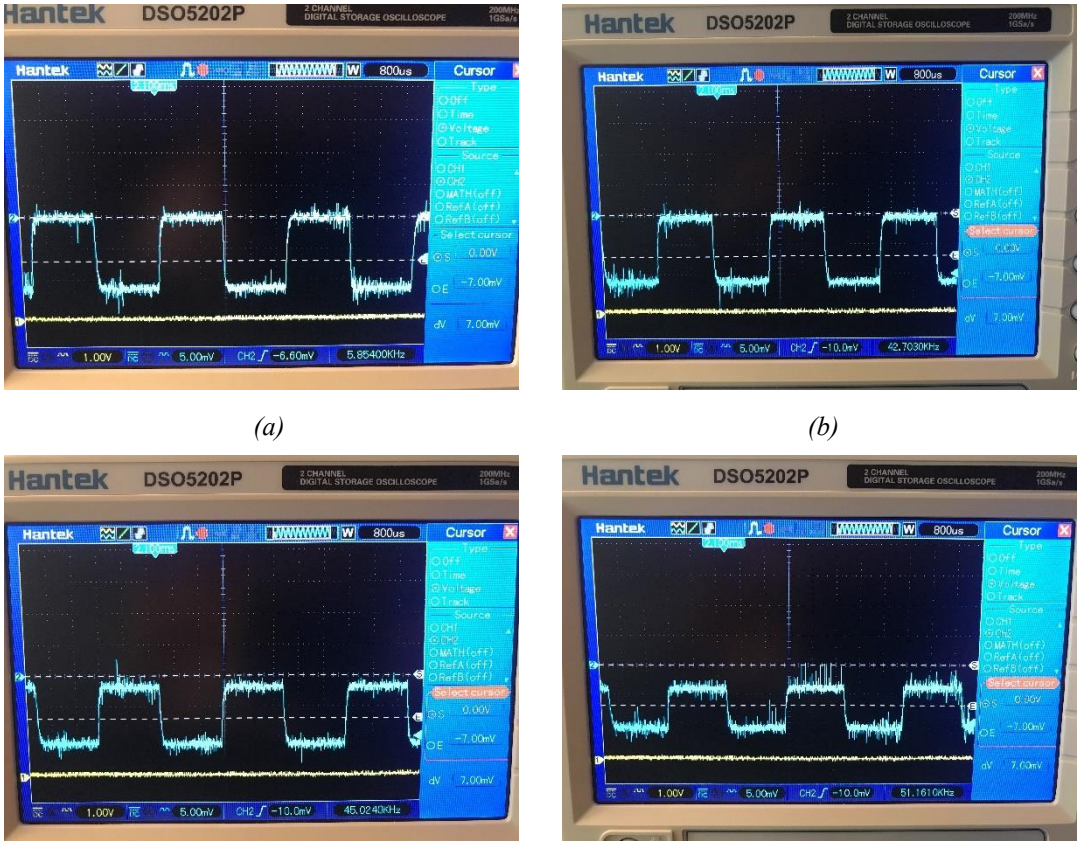
The curve in Fig. 5.8(d) is obtained by dividing the data in Fig. 5.8(a) by the corresponding data in Fig. 5.8(c) and it represents the derived Seebeck coefficient. The curve contains large spikes when the input voltage is small. The power dissipated by

the heaters is extremely low at this region so as the temperature raise at the thermocouple junction. Hence the measured temperature values are highly susceptible to the background noise. As the input voltage increases, the curve reaches a constant value at 35 μV/°C. This derived value is 3 times more than the value reported in [76]. This is mostly because the temperature reading from the IR camera underestimates the actual thermocouple junction temperature.

From Fig. 3.24(b), we see that the estimated junction temperature is about 54 °C above room temperature (the average temperature of a 20 μm x 300 μm area at the center of the sample above room temperature, 70 °C, divided by the ratio of the input Seebeck coefficient, 10 μV/K, to the extracted Seebeck coefficient from the average temperature, 7.7 μV/K) with 5 V being applied to the heaters. Using this number, the estimated Seebeck coefficient of the sample is 19 μV/°C which is closer to the value reported in [76]. A further calibration of the temperature reading with simulation result may be needed to obtain a more accurate estimation of the Seebeck coefficient.

Next, square-wave voltage signals were applied to the heaters to measure the transient response of the thermocouple. The peak-to-peak voltage of the input square waves ranged from 5 V to 2 V and the maximum voltages were kept at 5 V. The thermocouple output waveforms are shown in Fig. 5.9(a) to (d). Because the probe was connected to the thermocouple with reversed polarity, the displayed thermocouple output voltage was negative. As the peak-to-peak of input voltage decreases, the power dissipated by the heaters at the idle state increases. Therefore, the thermocouple output voltage at the idle state approaches its lowest value. The lowest outputs were all the same, about -11 mV, for four different heater input peak-to-peak values. This indicates that the steady temperature of the thermocouple depends mostly on the absolute applied heating power and does not vary with its initial temperature state. That is, the

relationship between the temperature of the thermocouple and the applied power is a simple one-to-one function.



(a) *(b)*

(c) *(d)*

Fig. 5.9 Thermocouple output waveforms with square signals applying to the heaters.

5.2 Experiment Results of the Control Circuit

In this section, I provide the measurement results of individual components in the broken-loop feedback system followed by the validation of the broken-loop feedback control mechanism. The circuit chip contains 52 I/O bonding pads. The chip was assembled in a 48-pin dual-in-line package (DIP) by having two power VDD pads to be assigned to a same pin at each corner of the I/O pad ring. For testing flexibility, the circuit system was built on a breadboard so the connections could be changed easily. A FPGA was used to generate the control signals for the chip, e.g., the integration and

reset signals of the integrator sense amplifier. The frequency of the FPGA's internal clock is 200 MHz. The internal clock is fast enough for deriving all the control signals for our system. The voltage level of the FPGA's outputs is 1.8 V. Inverters were used to boost the output level to 2.5 V, matching the I/O voltage level of our chip. Fig. 5.10 shows the experiment setup of the control circuit chip.

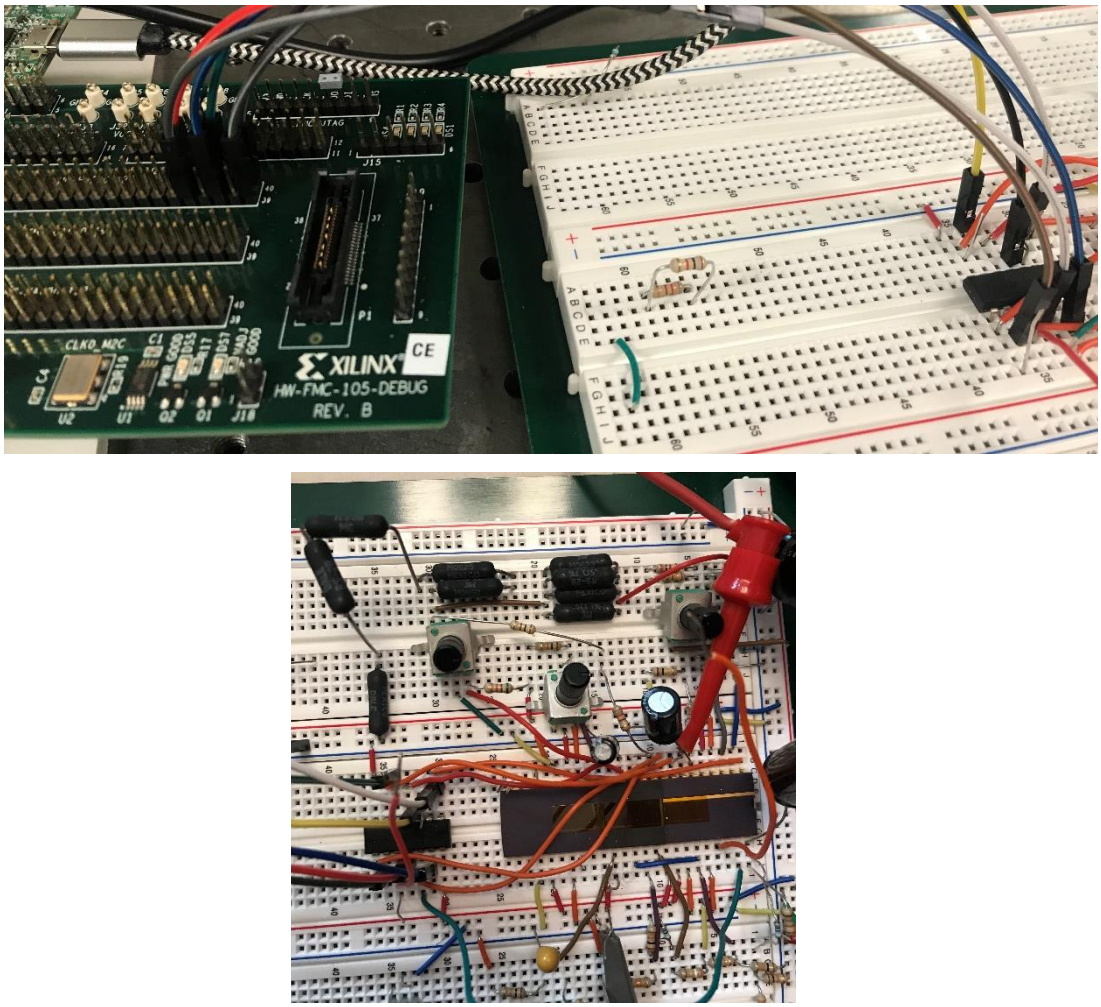


Fig. 5.10 Experiment setup of the control circuit

5.2.1 Integrator Sense Amplifier

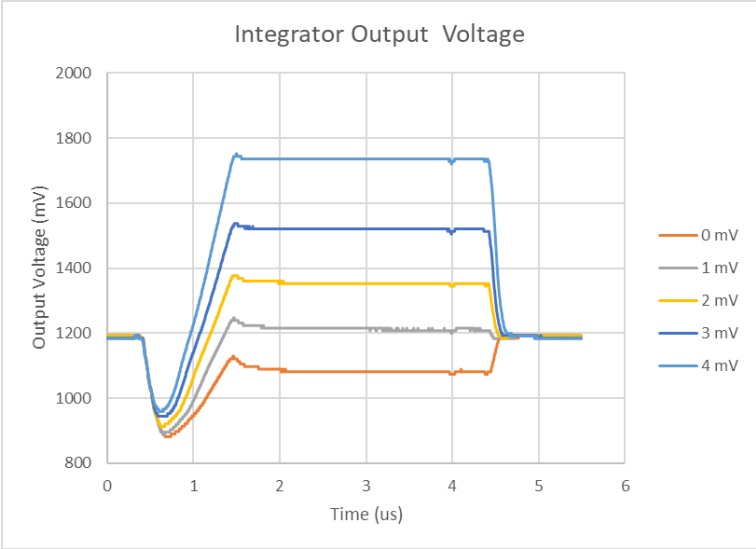
The first stage of the control loop is an integrator whose main purpose is to amplify the feedback signal from the thermocouple. The thermoelectrical response of a single gold-nickel junction is on the order of 10 μV/°C. The temperature at the junction at 2π phase shift is 400°C, so the expected output swing of the integrated

heater/thermocouple device is about only 4 mV. The integrator increases the dynamic range of the feedback signal so the product of the step size and the number of the step of the control reference input can be larger than 4 mV.

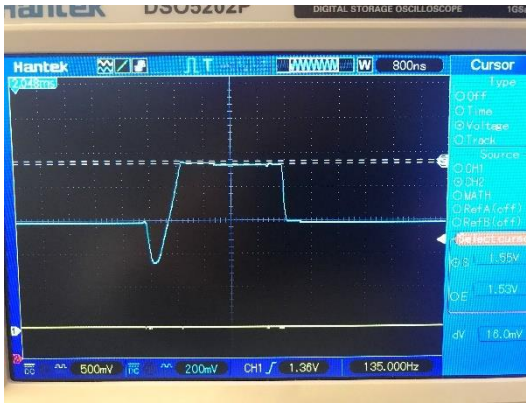
Ideally, an integrator sense amplifier with rail-to-rail output swing is desired. However, the intrinsic gain of transistors in submicron process technologies is small. It is unrealistic to pursue large output swing, high open loop gain, and small circuit area at the same time. In our design, we sacrificed the output swing for the circuit area and open loop gain. The output range of the integrator was designed to be more than 400 mV. For 1.5 mV step size, the precision of the input reference voltage can be 8-bit. This corresponds to 360°/256, 1.4°, step size of phase shift, which is less than 0.01π. On the other hand, the linearity between the integrator's input and output is not a major concerned, because it is possible to compensate the nonlinearity by adjusting the reference input to match the response of the integrator.

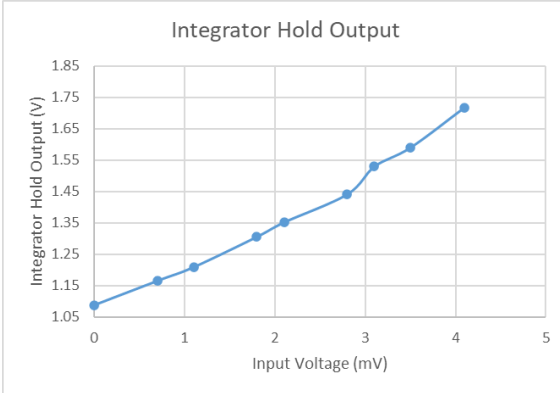
The measurement results of the integrator sense amplifier are summarized in Fig. 5.11. The integration time was 1 μs and the hold time was 3 μs in the testing. The hold time in the testing is longer than value in the normal operation, which is only 1 μs. The extended hold time help verify the speed of the integrator. A constant input was applied to the integration resistor which is connected to the negative terminal of the core amplifier. The input was varied between 0 to 4 mV below the DC bias operation as shown in Fig. 5.11(a). The data were exported from an oscilloscope and plotted in Excel. Due to the process variations, an offset of about 20 mV existed between the differential input pair. The DC bias points of the differential terminals were adjusted to cancel the intrinsic input offset. This offset is expected because all the transistors were minimized to meet the area constraint. Fortunately, the input offset can be compensated by tuning the DC bias at the positive terminal. This can be performed along with the output non-

linearity calibration at phase zero after the fabrication.



(a)





(b) *(c)*

Fig. 5.11 Integrator Sense Amplifier Output Measurement Result

The integrator output experiences a 200 ns undershoot during the integration because of the clock feedthrough from the control switches. The shortest integration time should be more than 200 ns so the output is not dominated by the switching noise. The speed of the circuit is fast enough for a 1 μs integration time.

When the integrator enters the hold mode, the output experiences a small overshoot from the clock feedthrough again and stabilize within 1 μs. The reset is fast and can be done within 100 ns. Together these results demonstrate that the

thermocouple feedback signal can be sensed and amplified within 2.5 μs. Therefore, the integrator meets our design requirement. The integrator can be shared among four phased array pixels and operates at a 100 kHz sample rate.

Figure 5.11(b) shows a photo of an oscilloscope output waveform with 3 mV input. The input and steady output voltage in hold state curve is illustrated in Fig. 5.11(c). Th output range was more than 400 mV to meet the requirement. The slope of the curve increases slightly with the input voltage because the equivalent resistance of the integration resistor, a NMOS operating in deep triode region, drops as the input voltage increases.

In summary, the testing results show the integrator provides enough gain for a 1 μs integration time and the transient response is fast enough to be switched between four phased array units at 100 kHz feedback signal sampling rate. Thus, the circuit meets the requirement of the feedback control loop design.

5.2.2 Intermediate Stage Comparator

To maintain a specific temperature at the phase shifter, the feedback loop adjusts the output current of the heater driver automatically so that the difference between the amplified feedback signal and the control reference voltage is small. This function is performed by the comparator in the feedback loop. The comparator outputs a digital signal to change the state of gate driver based on that result. The primary requirement of the comparator is the ability to respond to the difference between its input signals within the range of the amplified feedback signal accurately.

Because the number of the bonding pads is limited, only one of the input terminals, control reference, and the output of the comparator are connected to the pads. The other input terminal which is fed by the output of the integrator in the previous stage is an internal node. This node is followed by a unity gain buffer for measuring its voltage

without loading the output of the integrator. External voltage sources cannot be connected to this node to set a DC voltage for testing the output transfer characteristic of the comparator.

The circuit connection for measuring the output response of the comparator is as follow. A constant reference voltage was applied externally to the input terminal of the comparator which is linked to a bonding pad. A sawtooth signal (the blue waveform in Fig. 5.12) was applied to the other terminal of the comparator internally by adding a square wave to the input of the integrator. As this signal exceeds or falls below the reference voltage by a certain amount, the output of the comparator (the yellow waveform in Fig. 5.12) switches state. The instantaneous values of this signal when the state transitions started were recorded as the output characteristic of the comparator.

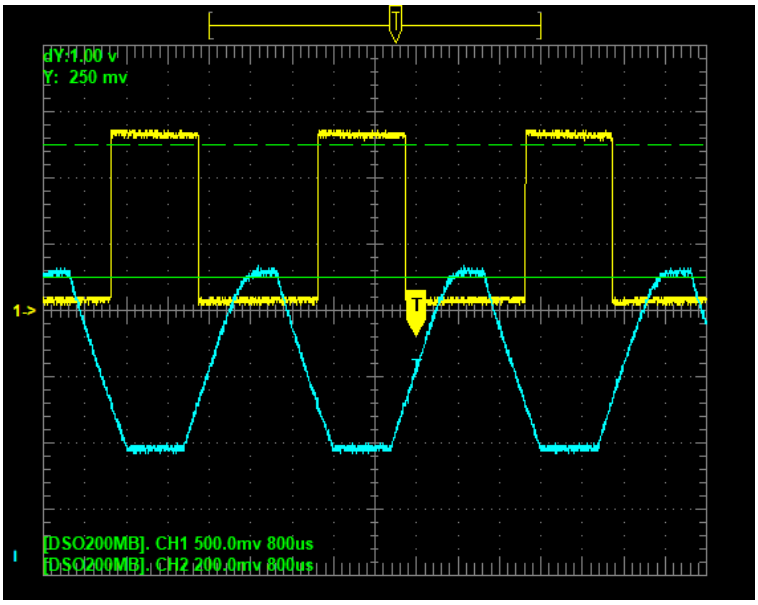


Fig. 5.12 The input and output waveform of the comparator under testing

The oscilloscope used in the experiment could not export the waveforms of its two channels simultaneously. An uncertain time delay existed between the two channels of exported data, so the actual value of the sawtooth like signal that triggers the comparator to swing could not be identified accurately from the exported data. Therefore, the

waveforms were measured on the screen of the oscilloscope directly by using the built-in cursors. The accuracy is limited by the step size of the cursors which is 4 mV in our case.

The measurement results are summarized in Fig. 5.13. Fig. 5.13(a) shows the threshold voltage for triggering the comparator to swing at different levels of reference voltage. Two curves are plotted because the comparator was designed with hysteresis to enhance the stability. The orange curve represents the threshold voltage for the comparator switching from 0 to 1 while the blue curve represents the opposite. The difference between the two curves is plotted in Fig. 5.13(b). The amount of the hysteresis is almost constant between 1.1 V to 1.7 V reference voltage. The value is 68 mV and is about 7 times more than the simulation result. This may result from some layout dependent effects which were not modelled accurately in the simulation.

The difference between the thresholds of two opposite transitions and the reference voltage are shown in Fig. 5.13(c) and (d). Ideally, the difference should be constant and be ±34 mV for the entire operational range. However, similar to the case in the integrator, offset exists at the input differential pairs of the comparator due to process variations. The voltage differences are 8 and -60 mV instead. The range of the difference is 12 mV for the reference voltage between 1.1 to 1.7V. That is, the comparator introduces a maximum 2% of error.

In conclusion, ***the comparator functions correctly***. The amount of the hysteresis is higher than the designed value, but this does not affect the function of the control loop. An input offset exists, but it can be negated by calibration. However, the maximum comparison error is 2% which is higher than the expected value. This will reduce the precision of the feedback control.

It is worth noting that large signal coupling between the output of the integrator

and the output of the comparator was observed. At some operating points, the coupling voltage exceeds the 70 mV hysteresis and it causes oscillation at these two nodes. Since the control system was wired and tested on a breadboard, it is highly possible that the signal coupling results from the parasitic elements of the breadboard instead of the internal chip interconnections. A carefully designed PCB with minimized signal coupling through the external connections on board is needed to distinguish the source of oscillation. Fortunately, because of the unique property of the broken-loop feedback, the loop control is not disabled by the oscillation. The control loop is not actually closed at any moment, so the oscillation at comparator does not lead to oscillations over the entire loop.

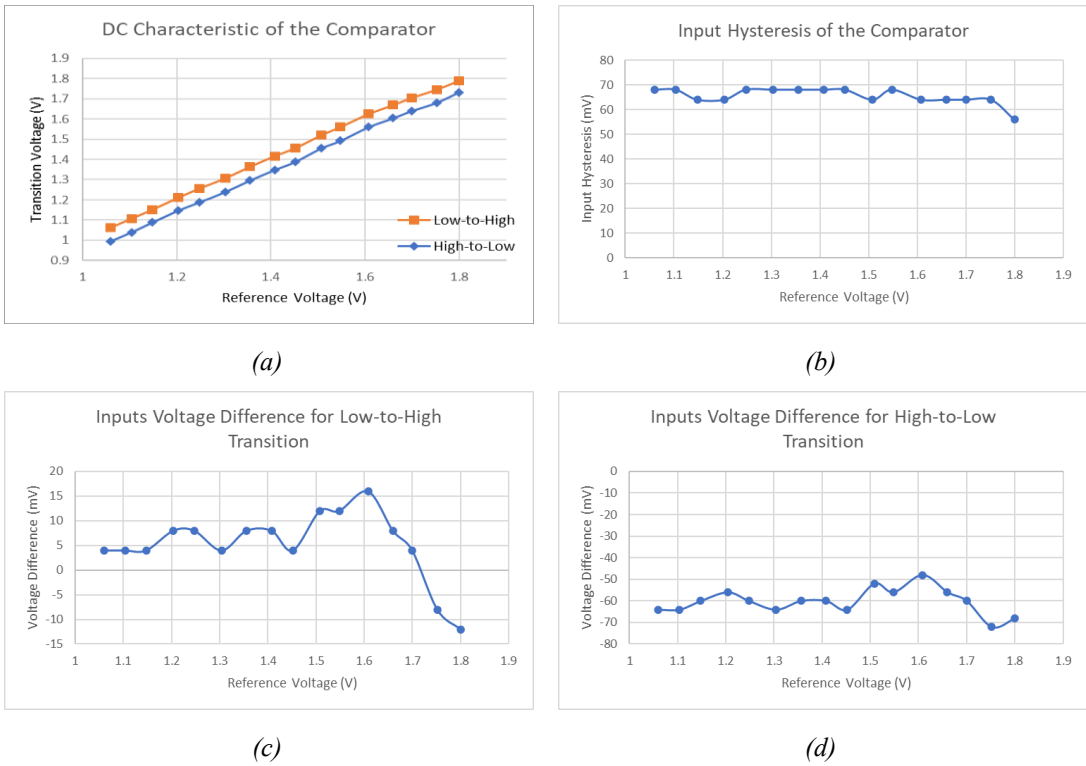


Fig. 5.13 Measurement results of the comparator

5.2.3 Final Stage Gate Driver and Heater Driver

The gate driver of the heater element receives commands from the comparator. These commands provide the proper gate bias voltage to the NMOS array heater driver.

The heater driver delivers a heating current to the integrated heater/thermocouple connected to its drain terminal. The most important requirement for the heater driver is that it must be able to supply a current for a 2π phase shift. The power for a 2π phase shift from the simulation results is 4.8 mW. Hence, if the target V_{DS_MIN} is 0.2 V, the output current of the NMOS array heater driver should be higher than 2.09 mA with a 1.1 KΩ load.

To reduce the temperature ripple at the heater as described in the previous chapter, the changes of the gate bias voltage between each cycle have to be limited. To achieve this, the gate driver needs to contain a memory element for storing the bias state because the gate voltage is reset every cycle for sensing the feedback signal from the thermocouple. We explored two gate driver implementations, an analog driver and digital counter. The measurement results of the analog gate driver are provided first, followed by the results of the digital counter approach.

5.2.3.1 Analog Gate Driver

Our analog gate driver uses a capacitor to store the state of the heater driver. The testing is aimed to verify the output current capability of the heater driver and the functionality of the gate driver. We will show that the gate driver can update the voltage of the state capacitor gradually as designed and the dynamic charging current source increases the step size near light and full load to compensate for the nonlinearity between the state capacitor voltage and the heater power.

The gate of the heater driver is an internal node so we cannot measure its current characteristic by applying a voltage to it directly. Instead, the gate of the heater driver can only be controlled by adjusting the V_{B1} in Fig. 4.23(a) and V_{B2} and V_{leak} in Fig. 4.23(b). A unity gain buffer is added to isolate both the output of the source follower buffer in Fig. 4.23(a), and the bonding pads for the external connection. In this way, the

circuit operation is not affected by an external probe. The unity gain buffer does not provide rail-to-rail output. The lowest output voltage is capped at 810 mV. The actual gate voltage of the heater is not readable when the output current is small. This defect only hinders the testing and does not affect the circuit operation. The unite gain buffer will be redesigned so the gate voltage of the heater within entire operation range can be measured in the next chip fabrication cycle.

The load resistance connected to the heater driver in the experiments for testing the gate driver is 1.1 KΩ. A 900 Ω external resistor is in series with a 220 Ω internal resistor, thus providing ESD protection. Because of the ambiguous description in the TSMC I/O library manual, this ESD resistor was added as a precaution against ESD device damage. In the same chip, we had a stand-alone MOSFET array whose drain terminal is connected to the bonding pad directly without a ESD resistor. We next confirmed the ESD resistor can be removed safely with this stand-alone MOSFET array. In the next fabrication cycle, the load of the heater driver can be entirely made using a 1.1 KΩ heater.

The first test was to measure the drain current of the heater driver by sweeping the V_{B1} in Fig. 4.23(a) without any charge in the state capacitor. The main function of the V_{B1} is to set the DC bias point of the source follower buffer so the heater driver is in the sub-threshold region in the initial state. In this way, the heater driver does not "get stuck" in the cut-off region for several cycles after the state capacitor being reset. The heater driver current with respect to the V_{B1} in Fig. 4.23(a) is shown in Fig. 5.14. The heater driver starts to turn on when the V_{B1} exceeds 0.5 V and the current reaches 370 μA when the V_{B1} is 2 V. The V_{B1} is set at 0.8 V, the same voltage level as the one of the bias voltages of the integrator sense amplifier. The heater driver current is about 20 μA so the quiescent heater power is less than 0.01 % of the power at 2π phase shift. In this

test, the heater driver was only turned on partially, so the actual output of the source follower buffer could not be measured through the unity gain buffer.

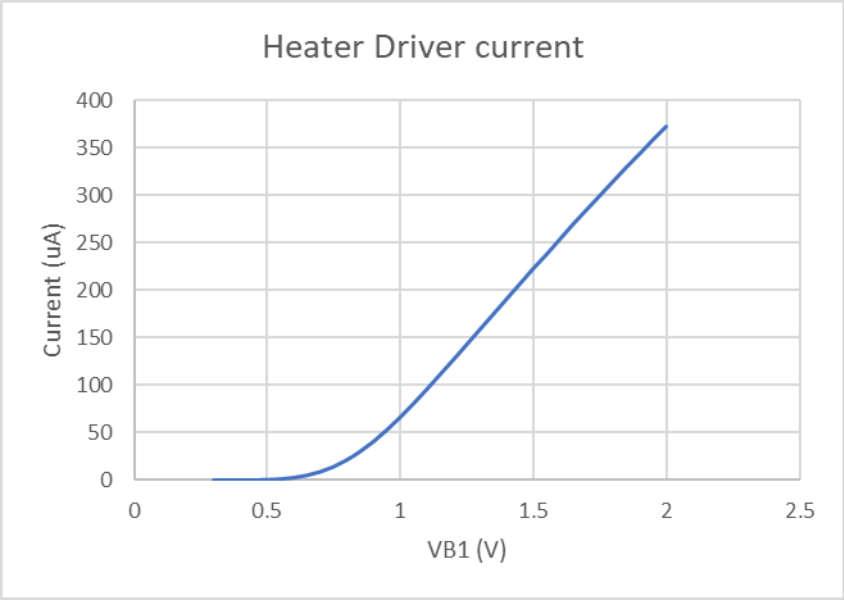


Fig. 5.14 Heater driver current vs V_{B1}

Next, to test the maximum current capability of the heater driver, we managed to set a DC voltage across the state capacitor by varying the V_{B2} and V_{leak} in Fig. 4.23(b). The EN and UP were tied to GND and V_{DD}, so the source of the M_{P4} was in effect connected to V_{DD}. The state capacitor reset transistor M_{N3} was disabled. Two dynamic adjusted current sources, MP_5 and M_{P6}, were shut off by connecting the V_{B3} and V_{B4} to ground. Therefore, the state capacitor was equivalently connected to the drain terminal of one NMOS and one PMOS in series, M_{N4} and M_{P4}, whose source terminals were tied to GND and V_{DD} respectively. By tweaking the gate voltages of M_{N4} and M_{P4}, the equilibrium voltage at their drain voltage was adjusted to realize the correct bias condition of the heater driver.

The curve of the heater driver current vs the output voltage of the unity gain buffer is shown in Fig. 5.15. The curve is vertical at the beginning because the lowest output voltage of the buffer is limited at 810 mV. Once the gate voltage at the heater driver

exceeds this lower limit, the actual gate voltage can be read from the unity gain buffer. The heater driver current is 2 mA when the gate voltage is about 1.2 V gate. After that, the current increases slowly to its maximum value, 2.15 mA. The result validates that the size of heater driver is large enough to provide adequate power to its heater.

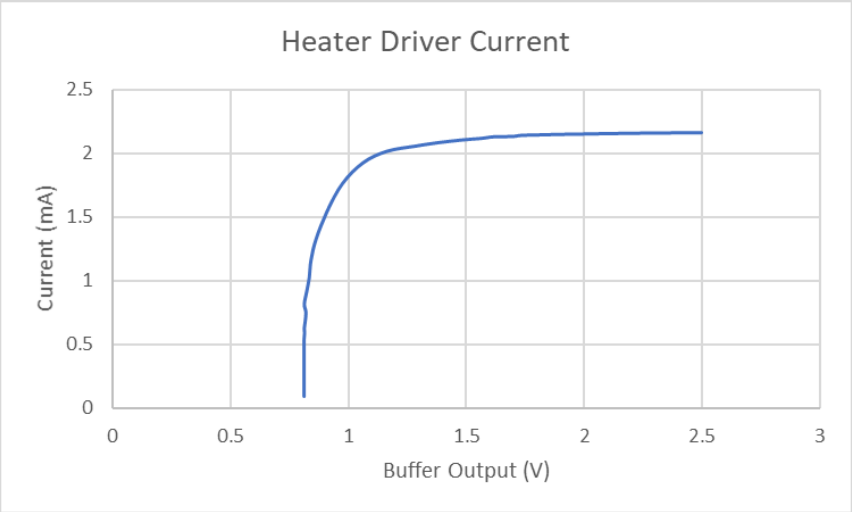


Fig. 5.15 Heater driver current vs unity gain buffer output voltage

Finally, we tested the function of the dynamic current source for charging the state capacitor. The testing procedure was as follow. The state capacitor was fully discharged first. Next, we input a few 1 MHz pulses (using a 50 % duty cycle) to the gate of M_{S3}. This is the EN signal, as shown in Fig. 4.23(b). The EN was set to high after these pules. The voltage capacitor stayed constant for a short amount of time given the leakage was small and the heater driver current during this period was measured.

An example waveform of the output voltage of the unity gain buffer with 30 consecutive pulses at the node EN is shown in Fig. 5.16. The state capacitor was reset at 4 μs, so the unity gain buffer output dropped around 810 mV. Again, the output of the unity gain buffer did not change at the beginning of the pulses because the gate voltage of the heater was less than 810 mV during this period. When the gate voltage exceeded 810 mV, a staircase like waveform was over an interval from 22 μs to 35 μs.

In each step, the output voltage increased for 500 ns and held constant for another 500 ns. After the 30 pulses, the output voltage was constant because the state capacitor was floating. We measured the heater driver current during this stage. ***This waveform demonstrated that the mechanism for charging the state capacitor and increasing the heater driver current worked as designed.***

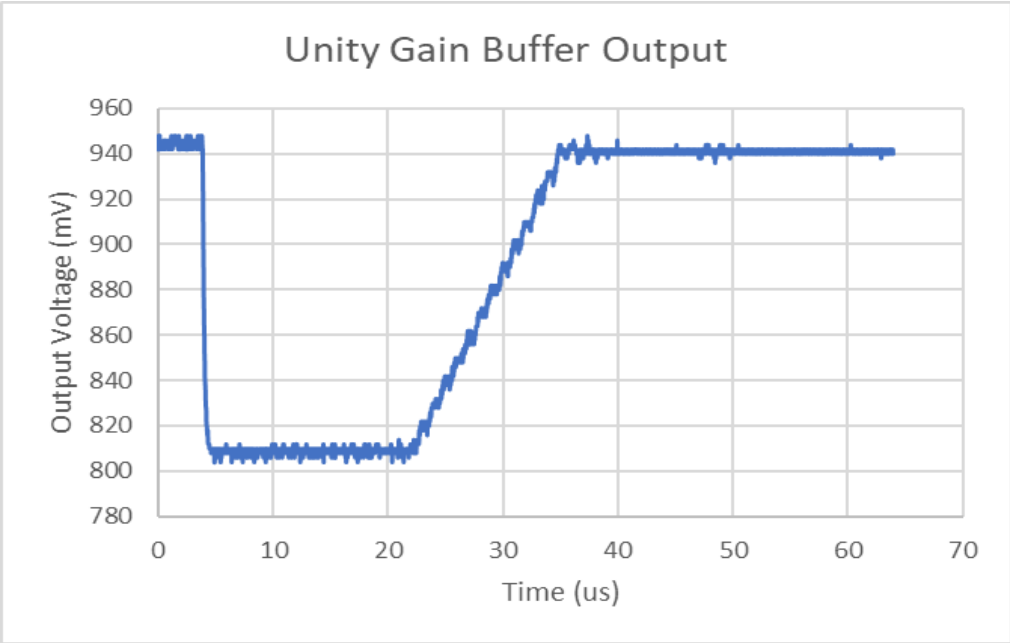


Fig. 5.16 The output voltage of the unity gain buffer with 30 state capacitor charging pulses

The number of the pulses was varied with/without the dynamic current sources, M_{P5} and M_{P6}, being on. The resulting heater driver currents for different bias configurations in the hold state are shown in Fig. 5.17. The curve of V_{B3} and V_{B4} connecting to ground indicates that the state capacitor is charged by a constant current source, M_{P4}, only. This curve grows quadratically so the step size of the heater power would be too small at the beginning cycles. With V_{B3} being 0.8 V, the M_{P5} is on moderately when the voltage across the state capacitor is low so the step size of the heater driver current is increased in this region. As the voltage across the state capacitor keeps increasing, M_{P5} is gradually turned off. Hence, the slope of the curve of V_{B3}/V_{B4}

being 0.8/0 V is close to that of the first curve. In contrast, if V_{B3} is increased to 1.2 V, the M_{P5} is on heavily, and it is not completely off as the voltage across the state capacitor increases. The step size of the heater driver current becomes larger than that of the previous two cases in the entire operating range. Finally, with V_{B4} being 0.6 V, the current source M_{P6} is only on when the heater driver current is high. Therefore, the last two curves – with V_{B3}/V_{B4} being 1.2/0 V and 1.2/0.6 V – overlap when the current is less than 1.2 mA. Together, these results verify the function of a dynamically adjusted current source which is capable of adjusting the step size of the heater driver current. Thus, by changing the bias voltages, V_{B2}, V_{B3}, and V_{B4}, the step size of the gate driver circuit can be manipulated to match the current-temperature response of the heater in an actual phased array pixel. In that way, a more uniform step size of the temperature change can be achieved.

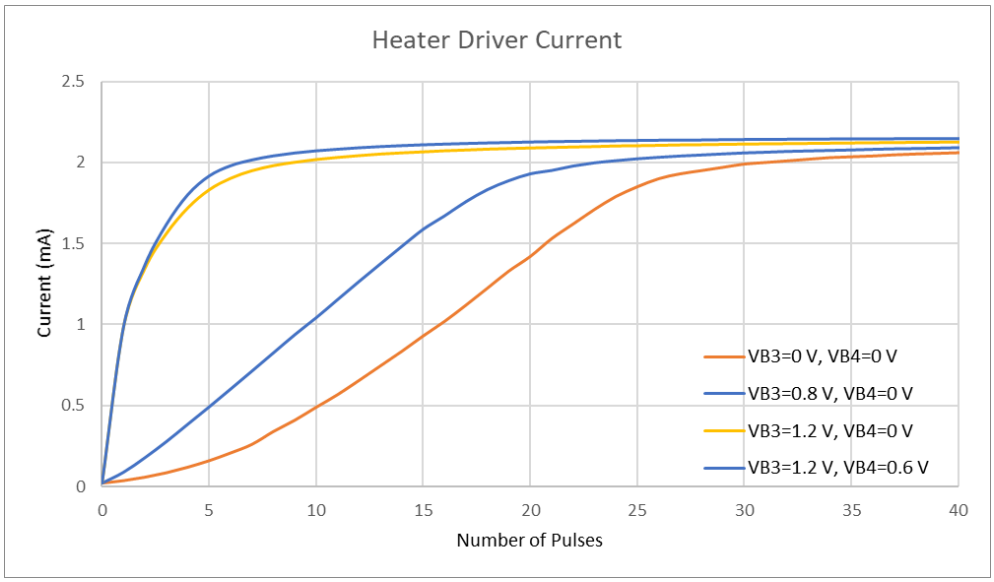


Fig. 5.17 Heater driver current vs number of pulses at EN with various bias conditions

In summary, the experiments described in this section showed that the heater driver array can supply current for 2π phase shift and verified that the gate driver circuit functioned correctly. The dynamic adjusted current source setting the state of the heater

driver was proved to be able to compensate for the non-linearity of the current-temperature response of the heater.

5.2.3.2 Digital Gate Driver

The digital gate driver is much simpler than the analog counterpart. An 8-bit counter is used to represent the state of heater driver. The heater driver array was divided into groups using binary weighted code. All the gates of the transistors in the same group are connected to the corresponding output bit of the 8-bit counter. This control method is straightforward, but the major drawback is the mixed usage of core and I/O transistors in the design. The extra area overhead due to the ESD and latch-up protection makes this approach less favorable for our stringent space constraint design.

The experiment described here focused on testing the functional correctness of the 8-bit counter. The counter was not implemented by logic cells in the standard library. The layout of the counter was customized to minimize the space consumption. It is important to verify the full and correct functionality of this component. Moreover, because the heater driver array used in the digital gate driver design is different than the one used in the analog method, it is necessary to test the current capability of the heater driver again here. Because of the limited number of the bonding pads, not all the outputs of the 8-bit counter are connected to the pads. The counter cannot be tested directly. To overcome this issue, a single pulse was applied to the clock input of the counter step-by-step. The heater driver current was measured after each pulse. The testing results can prove the counter was operated properly if the current increased after each pulse and 256 states existed. Thus, the function of the counter and the current capability of the heater driver could be tested at the same time.

The heater driver current at different output values of the counter is plotted in Fig. 5.18. The current wrapped around when the number of the pulses exceeded 255 and this

indicates the counter was working correctly. The maximum current was more than 2.5 mA, so the heater driver can provide a power for 2π phase shift without problem. The maximum current can be lowered by reducing the number or increasing the channel length of the unit transistor in the heater driver array. Both methods need extra circuit space. The former method requires us to change the mapping between the unit transistors and the outputs of the 8-bit counter. The latter method increases the size of the unit transistor itself. The current grows near linearly with the state of the counter at the beginning because the unit transistors turning on in this early-stage act like ideal current sources. The curve is flattened because the drain voltage at the heater driver array drops. The channel length modulation or the case in which the transistors enter the triode region decreases the current of each unit transistor.

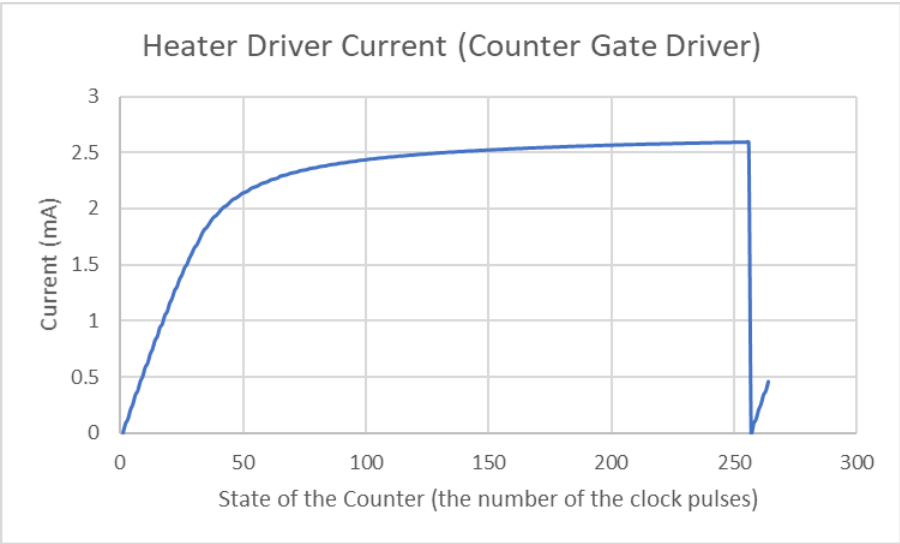


Fig. 5.18 Digital gate driver, the heater driver current vs outputs of the counter

Unlike the analog driver approach, the heater power is digitized by using a counter to control the state of the heater driver. The control precision is limited by the quantization error of the heater power. Fig. 5.19 shows the ratio of the heater power step size to the maximum heater power versus the normalized heater power at each step,

from 0 to 254, assuming a constant heater resistance. The curve contains several swings because of the unit transistor mismatch due to layout dependent effects. This causes the non-uniform heater current increments with the counter counting up. The heater power is not quantized uniformly. The step sizes are between 1.3 % and 2.6 % of the maximum heater power for heater powers between 20 % and 50 % of the maximum heater power (corresponding to 0.4π to π phase shifts). When the heater power is more than 80 % of its maximum value, the step sizes reduce significantly and are less than 0.5 % of maximum value. The maximum error is about 0.05 π when the target phase shift is π. To reduce the quantization error, one can increase the number of bits in the counter. It is also possible to adjust the bit encoding scheme, so the step size is distributed more evenly. However, either way needs additional circuitry and the area constraint may not be met.

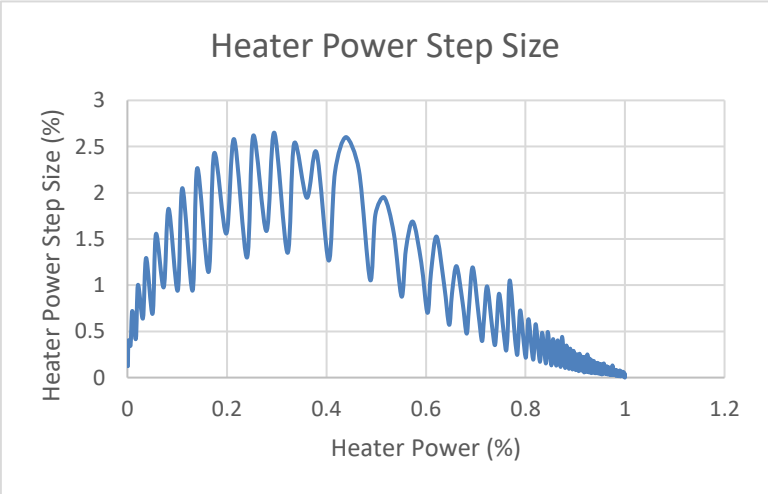


Fig. 5.19 The step size of the heater power increment vs heater load conditions

In conclusion, ***the experiment results reported in this section prove that the counter gate driver is operated correctly and the heater driver is capable enough to deliver power for 2π phase shift***. Because of the simple binary encoding of the heater driver array, the step sizes of the output heater power are distributed non-uniformly and

the maximum phase error is estimated to be 0.05 π. More sophisticated encoding method to reduce the phase error due to this quantization error will be explored in the future.

5.3 Control loop test

Ideally, a phased array unit with the exact structure in the COMSOL simulation described above is preferred as the load of our driver circuit for testing the broken-loop feedback control. This means the temperature at the load should be raised to 400 °C while dissipating only 4.8 mW. This is only valid when the heater only heats up an area of 1 μm x 13 μm and thermal barriers such as air trench are introduced to increase the power efficiency. During the course of this thesis, I did not have access to the optical part of our phased array system. To overcome this difficulty, we designed two experiments to prove the concept of the broken-loop feedback indirectly.

For the first experiment, we substitute the heater load with two resistors in series as shown in Fig. 5.20. The resistor R_2 contributes most of the total load resistance while the resistor R_1 is small and convert the heater driver current to a small feedback signal which simulates the thermocouple output. Unlike a real thermocouple which can generate a voltage signal from temperature gradient, the resistor R_1 needs the heater driver to source a current continuously to generate a feedback signal. Therefore, the major difference of the circuit operation between the testing method and the actual broken-loop feedback control was that the heater driver was not shut off for sensing a feedback signal. Nevertheless, the feedback signal was only measured at particular intervals and the state of gate driver was updated a fixed amount of time after each feedback signal sampling. In essence, the control loop in testing was not actually closed and was divided in two phases, just as in the actual proposed broken-loop control. This experiment intends to verify whether the broken-loop feedback control can adjust the

heater driver current so that the feedback signal from the resistor R_1 passes the control reference voltage to the comparator after sampling and amplifying by the integrator. This would simulate the actual control loop in the optical array.

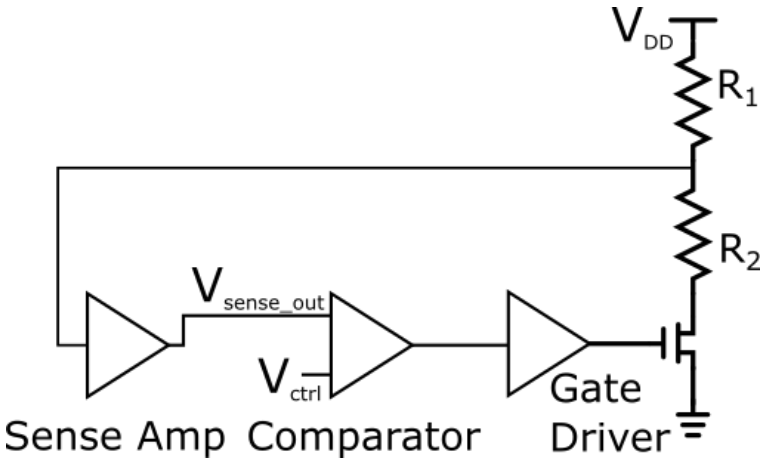


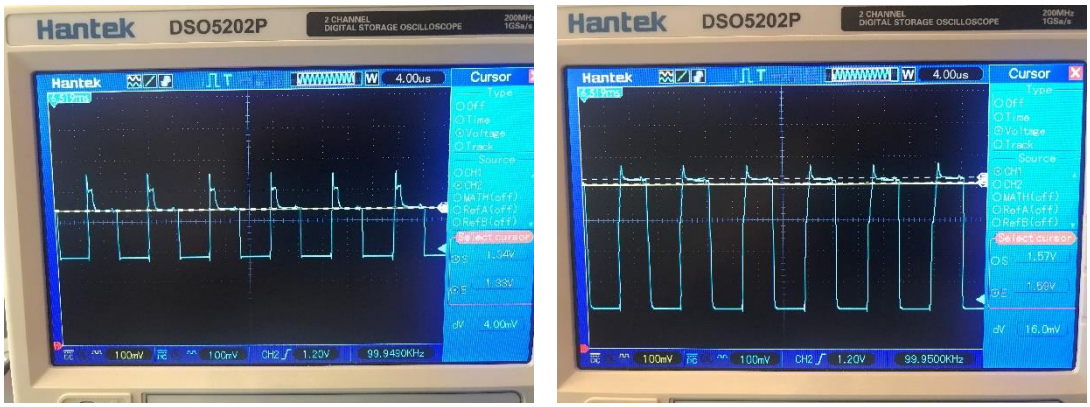
Fig. 5.20 Schematic of testing broken-loop feedback control

The integration time was 1 μs and the state of the gate driver was updated every 10 μs in the experiment setup. The control reference voltage was swept from 1.08 V to 1.7 V with 10 m V step size. To test the regulation performance, the resistance of R_1 and R_2 were varied intentionally to simulate phase errors due to device mismatch or proximity effect. The resistance values of R_1 and R_2 used in the experiment were 2.527/819 Ω, 2.304/644 Ω, and 2.078/478 Ω. These represent ±10 % and ±20% variations of R_1 and the total load resistance respectively with respect to the second combination, 2.304/644 Ω. The variation of R_1 changes the magnitude of the feedback signal given the heater driver current is unchanged. Thus, the control loop should adjust the bias point of the heater driver to reduce the voltage difference at the output of the integrator and the control reference under difference resistance values of R_1.

The gain of the integrator was tuned for the second R_1/R_2 combination so the heater driver could be fully turned on within the control voltage range. For example, if the gain is too large, the heater driver only needs to be on partially for the amplified

feedback signal being close to the control input. The circuit bias points were kept unchanged for different combinations of R_1/R_2. This made the output voltage range of the integrator vary with the change of R_1/R_2 because the maximum heater current and feedback ratio were different.

Figure 5.21(a) and (b) illustrate the steady waveform of the output of the integrator (blue curve) and control input (yellow curve, overlapped with the dash line cursor) for control input being 1.336 V and 1.576 V. The waveforms demonstrate that the integrator was operated correctly again, and the broken-loop feedback brought the two signals close to one another.



(a) *(b)*

Fig. 5.21 Stabilized waveforms of integrator output (blue) and control reference voltage (yellow) (a) 1.336 V control reference voltage, 4 mV difference between two signals, (b) 1.576 V control reference voltage, 16 mV difference

The error, voltage difference between the control input reference and the output of the integrator, for three combinations of R_1/R_2 within the entire input sweep range are plotted in Fig. 5.22. The absolute error is more than 20 mV for a reference voltage lower than 1.15 V. This is because the reference voltage is lower than the minimum output voltage of the integrator. The minimum voltage is limited by the heater driver operating in an off state. The output of the integrator starts to track the input reference voltage

closely as the reference exceeds 1.17 V.

Fig. 5.22 Voltage error versus the control input reference voltage

For the typical case, R_1/R_2 equal to 2.304/644 Ω, the maximum error is 20 mV for the reference between 1.18 V and 1.63 V. This corresponds to 4.4% of the total range of the integrator output, 450 mV. This results from the control mechanism of the broken-loop feedback. The gate driver keeps increasing the heater driver current until the output voltage of the integrator is larger than the reference voltage. After then, the heater driver current decreases slowly due to the leakage current at the state capacitor so as the output voltage of the integrator. The gate driver charges the state capacitor again once the output voltage of the integrator is lower than the reference value and this process repeats for the remaining time of the frame. If the step size is too big, the output voltage of the integrator may stay higher than the reference voltage for several cycles before it becomes lower than the target value. Therefore, the average output voltage of the integrator is a bit higher than the target, as shown in Fig. 5.21(b). This error can be reduced by making the step size smaller.

The curve presents the unique characteristic function of the reference input to the output of the integrator under this particular load condition (e.g., R_1/R_2 equal to 2.304

Ω /644 Ω). We can calibrate the control loop and offset the reference input for the target value at the output of the integrator. This means there is no error for this load. Next, the load is varied as the other two R1/R2 combinations, and we examine the amount of the variation at the output of the integrator as shown in Fig. 5.23. The figure illustrates that the broken-loop feedback indeed regulates the output against these variations. In Fig. 5.22, the maximum error is 16 mV, 3.6%, and the error is less than 4 mV, 0.89%, in most of the input region. The region of the blue curve where the input reference climbs above 1.57 V is neglected because the heater driver is already fully on when the input reference reaches 1.57 V. Above that point, the feedback loop cannot raise the output of the integrator anymore. The results show that the broken-loop feedback is capable of reducing the error below 3.6 % when the variation of the load resistance is as large as ±20%.

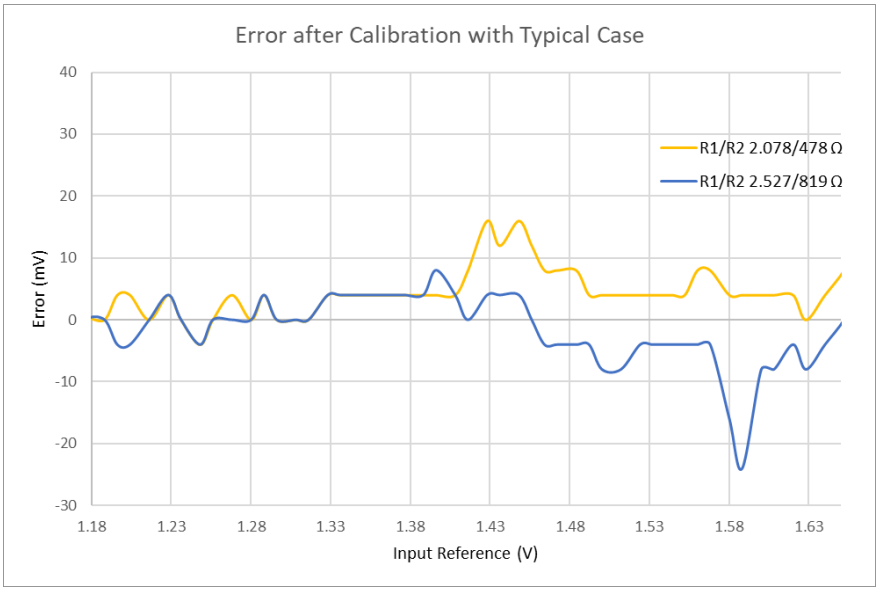


Fig. 5.23 Voltage error versus the control input reference voltage. The difference between the typical curve and the two variation curves in Fig. 5.22

Finally, the heater driver currents at different values of input reference for three R_1/R_2 combinations are shown in Fig. 5.24. The heater current decreases with the increment of the feedback coefficient, R_1, to keep the same feedback signal unchanged

as expected. This proves that the control circuit can adjust the output current under different circuit operating conditions, e.g., temperature offset as the result of the heat leaking from the neighboring pixels or device mismatch, to meet an equal target temperature. ***Thus, the first part of the experiment demonstrates the operation of the broken-loop feedback concept.***

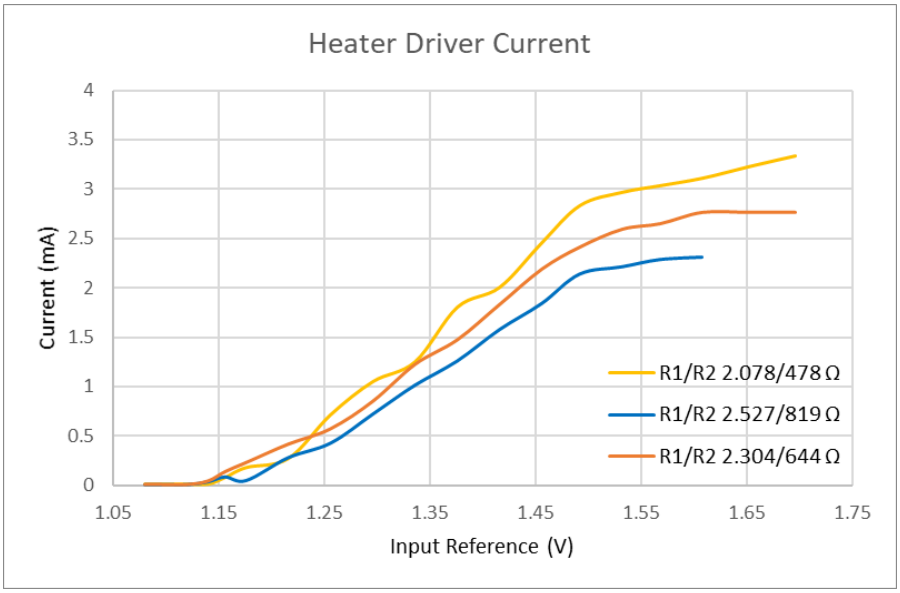


Fig. 5.24 Heater driver current versus the control input reference voltage

In the second of my experiments, I measured the transient response of the thermocouple. The main reason for conducting this test is that we assumed that when the heater driver MOSFET array is on, the voltage across the heater/thermocouple device is dominated by the IR voltage drop. The thermo-electric response is affected by the conducting current. When the heater driver is switched off, the heater/thermocouple device needs time to reach an equilibrium state to generate a voltage output representing its temperature. If this assumption is valid, we need to check the response time and ensure the circuit can sense an output from the thermocouple before the temperature drops to the ambient temperature.

The experimental setup is shown in Fig. 5.25. We connected the thermocouple

sample to our control circuit, as described previously. The heaters next to the thermocouple were connected to an external voltage source. The control circuit switched the heater driver array on and off with a pre-defined fixed gate voltage without processing any feedback signal. The external voltage source was either off or supplied a 5 V voltage. The transient voltage across the thermocouple with/without the external voltage source being on was compared.

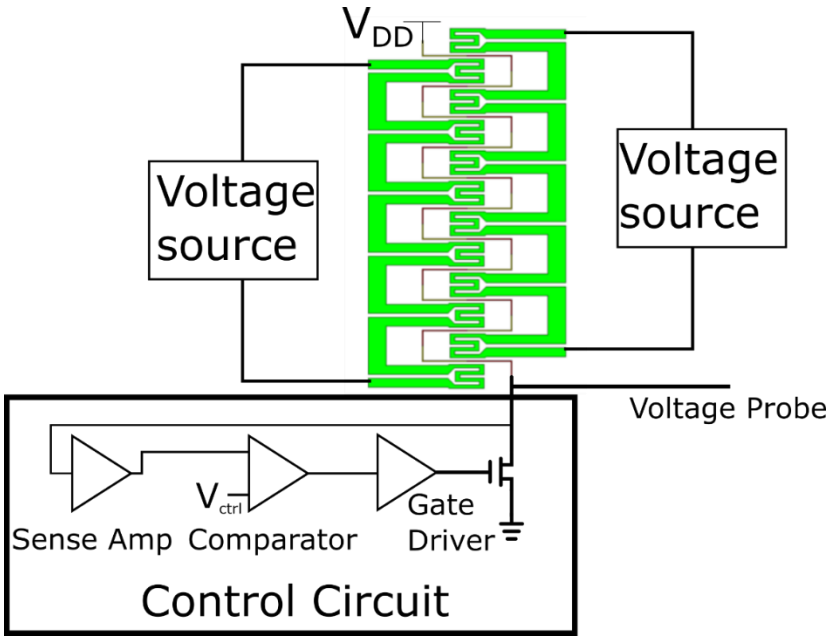


Fig. 5.25 Experiment setup for testing transient response of the thermocouple

The expected waveform of the voltage across the thermocouple is illustrated in Fig. 5.26. When the heater driver array is on, the voltage across the thermocouple is somewhat larger if the external voltage source is on. This small voltage difference results from the temperature dependence of the resistance of the thermocouple. The temperature of the thermocouple sample is higher when the external voltage is on. Thus, the IR voltage drop is a little bit bigger. When the heater driver array is off, no current passes through the thermocouple. The voltage jumps back to V_{DD} (2.5 V) when the heater driver array is on. If the external voltage source is 5 V, the expected output

voltage would be V_{DD} minus the thermocouple response which is about 12 mV according to the thermocouple measurement result described above. The time interval between t_1 and t_2 is the assumed response time for the thermocouple to reach its equilibrium state. This experiment is aimed to quantify this value.

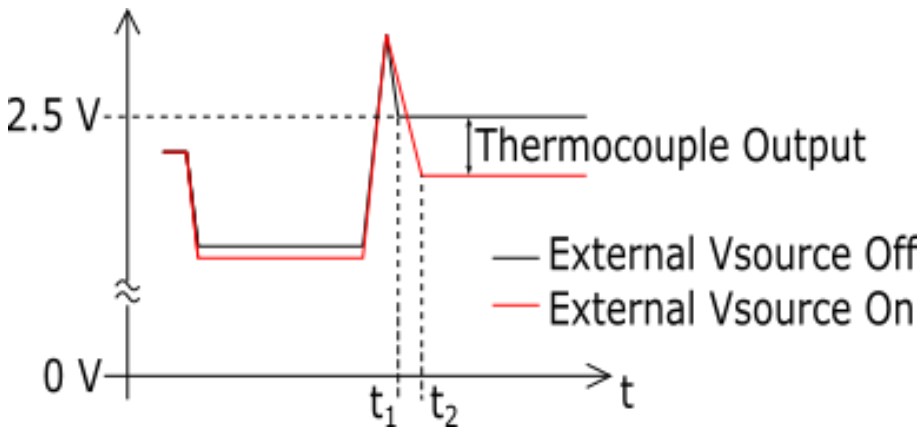
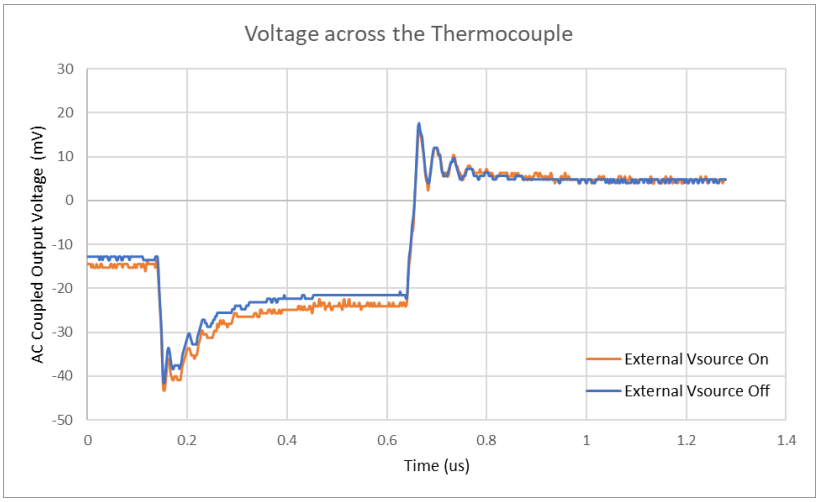


Fig. 5.26 Expected transient voltage across the thermocouple of the circuit in Fig. 5.24 with heater driver array being on and off

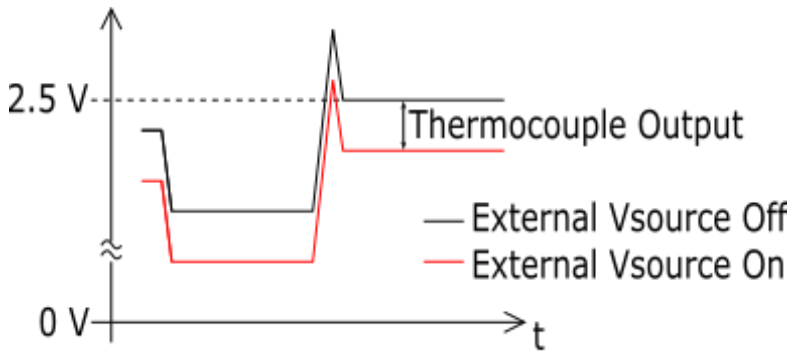
The measured transient waveforms of the voltage across the thermocouple can be seen in Fig. 5.27. In Fig. 5.27(a), the heater driver array was on between 0.4 μs and 0.64 μs. It conducts a small current flowing through the thermocouple, and the voltage drops below V_{DD}. The heater driver array was off after 0.64 μs so the voltage went back to V_{DD}. Somewhat surprisingly, the waveforms of heaters on and off cycles almost overlap perfectly, except the voltage is a little bit lower when the external voltage is on.

Figure 5.27(b) is the reconstructed DC coupled waveform of the output voltage from Fig. 5.27(a). It is different from the predicted waveforms in Fig. 5.26. This means our assumption of the thermocouple response is incorrect. The thermo-electric response is independent of the IR voltage drop. The thermocouple generates an output voltage related to the temperature difference between its junction and the ambient which is not affected by the current flowing through it. Superposition can be applied here. The

voltage across the thermocouple is the IR voltage drop plus the thermo-electric response. ***Therefore, the thermocouple respond time is not an issue in the broken-loop feedback control.***



(a)



(b)

Fig. 5.27 Transient voltage across the thermocouple with heater driver array being on and off (a) ac coupled data exported from an oscilloscope (b) dc coupled output waveform reconstructed from (a)

In conclusion, this chapter provides the results of my experimental evaluation of our NPA driver circuit fabricated in TSMC 65 nm technology. This evaluation is done at both component and system levels followed by the proof of the broken-loop feedback control methodology. The results show that the maximum error of the feedback loop is less than 3.6 %, 0.07 π phase error, and the average error is about 1 %, 0.02π phase

error, for a ±20 % variation at the load resistance. The result demonstrates the feasibility of the broken-loop feedback control for pixel level temperature regulation using the integrated TF H/TCs.

5.4 Feedback Control Performance Evaluation

In the previous section, the experimental results show the regulation error for a given range of load variations. This, however, does not explain how the broken-loop feedback control is essential to a VR or LIDAR NPA system. To demonstrate this, we compare the performance of VR and LIDAR NPA systems using a naïve direct control method with that using our unique broken-loop feedback control mechanism. To evaluate the performance degradation of a direct control, the causes of phase error in a NPA system must be identified. Thus, in this section, we carefully examine various sources of phase errors in an integrated NPA system.

For thermo-optic based phase shifters, the phases errors mainly result from failing to maintain the temperature at the exact target value. These errors cause the far field interference pattern to deviate from the ideal target. In the VR application, the image quality is degraded, and the background noise is increased by the phase errors. On the other hand, in the LIDAR application, the target steering angle is disturbed and the contrast between the main beam and the side lobes is reduced. Here, I quantify the performance degradation resulting from different types of phase errors for both VR and LIDAR applications. The results conclude that a feedback system is essential to suppress these errors. We prove that our broken loop feedback control system effectively enhances the image quality and beam forming accuracy. This circuit topology is a cornerstone of realizing large scale high performance integrated NPA systems.

5.4.1 Sources of Phase Error

For a highly integrated NPA and driver circuit system, we consider three major effects resulting in unwanted output phase deviations. The first effect is the device mismatch between the elements of the array unit due to process variations, e.g., the resistance of the heaters for raising the temperature at the phase shifters or the heater driver current source arrays. It is obviously that if some slight differences between the device parameters of two array units exist, the phase of the output beams of the two units will not be the same for a same input command. Given the number of the devices in each array unit and the complexity of the system, small process variations are expected with most current process technologies.

The second source of the phase error is the thermal crosstalk or the proximity effect. Because the array units are not completely thermal isolated from each other, the heat energy applied to one pixel for introducing a phase delay inevitably leaks to the adjacent pixels and perturbs the surrounding temperature. In chapter 3, we discussed a way for improving the heating efficiency by including thermal insulation trenches or free-standing structures inside the array pixel. Here, we show that these thermal insulation structures help suppress the proximity effect.

The proximity effect matrices of NPAs with 2 μm deep isolation air trenches and without any thermal barriers from the thermal simulation results are provided in Table 5.1. Table 5.1(a) shows the "spill-over" for trench-isolated cells, normalized Table 3.4 with respect to the center cell; and (b) shows the result without trench isolation. These tables show the unwanted heat added to pixels in the vicinity of the target pixel (the center element with the value 1). In both matrices the unity value represents the target cell, while the rest of the elements are the cells neighboring of the target cell. These values represent the percentage of the temperature shift relative to value of the target cell. The total phase shifts with the deviations of each element in a NPA can be

computed by convoluting the original phase assignments and the proximity effect matrix.

Table 5.1 Proximity effect matrices (a) with 2 μm deep isolation air trench (b) without any thermal barriers. The table shows the thermal crosstalk between a center cell and its neighboring cells.

Array unit position in the y direction

3	0.000578	0.000775	0.000982	0.001096	0.001018	0.000804	0.000641
2	0.000784	0.001224	0.001983	0.002683	0.00232	0.001393	0.000896
1	0.001069	0.002195	0.005864	0.014769	0.009723	0.002782	0.001242
0	0.001263	0.003317	0.016235	1	0.015714	0.003276	0.001341
-1	0.00115	0.00276	0.009569	0.014311	0.005746	0.002166	0.001126
-2	0.00084	0.001389	0.002305	0.002668	0.001983	0.001228	0.00084
-3	0.000657	0.000876	0.001119	0.001188	0.001058	0.000833	0.000641
	-3	-2	-1	0	1	2	3

Array unit position in the x direction

(a)

Array unit position in the y direction

2	0.017425	0.024317	0.032797	0.025469	0.016027
1	0.019776	0.043691	0.10195	0.052689	0.01872
0	0.023621	0.090783	1	0.092788	0.021828
-1	0.019947	0.050856	0.100225	0.044311	0.017431
-2	0.013996	0.02026	0.026553	0.01917	0.011998
	-2	-1	0	1	2

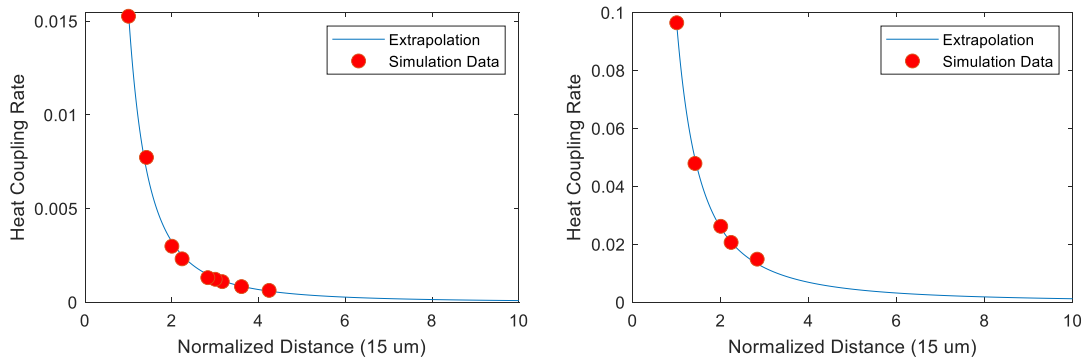
Array unit position in the x direction

(b)

From Table 5.1, one can find that, without any thermal barriers, the pixels in direct contact with the target cells receive an average of 10% of the heat applied to the target while with trench structures, the leakage is reduced to 1.6%. This is about a 6x reduction. The suppression of the heat coupling is even improved as the pixels move further away from the center target cell.

The size of the matrices is limited to 7 x 7 because of simulation computation resource restrictions. To obtain a larger proximity effect matrix for a more accurate phase error analysis, the thermal coupling between the target unit and the array units beyond the matrices in Table 5.1 Proximity effect matrices (a) with 2 μm deep isolation air trench (b) without any thermal barriers is estimated by extrapolating with a power function. The extrapolation results are shown in Fig. 5.28. The absolute values of the

exponents of both extrapolated curves are close to 2 which match the intuition of heat spreading being inversely proportional to the square of the distance. A slightly bigger exponent for the isolation trench verifies again that the heat generated at the targe cell is more confined.



$y = 0.01546x^{-2.261}$ $y = 0.09589x^{-1.902}$

(a) *(b)*

Fig. 5.28 Extrapolation results of the proximity effect (a) 2 μm deep isolation trench (b) without thermal barriers

In addition, because the driver circuit chip is closely integrated with the optical chip, the temperature at the driver circuit is susceptible to the heat dissipated in the optical chip. This heat may not be uniformly distributed. As the current-sourcing characteristics of a transistor is temperature dependent, the thermal coupling between the driver circuit and the optical chip introduces additional phase errors. This thermal coupling effect is much more complicated than the proximity effect within the optical chip. So, for simplicity, we only consider the local proximity effect and exclude this cross-domain coupling effect in our phase error analysis. Fortunately, this simplification does not undermine our conclusion. Including more phase errors only degrades the system performance and emphasizes the importance of feedback temperature regulation further.

The last source of the phase error considered in my analysis is the temperature gradient across the entire optical chip because of the non-uniform thermal resistance

between the heat sink and each array element. Fig. 5.29(a) shows the simulation result of the temperature distribution at the phase shifters of the butterfly image in Fig. 3.19(b) in 128 x 128 resolution. Here, I used a simplified heat source assignment same as the array size study discussed in the previous chapter. The highest temperature is at the array center and decays outwardly to the edge of the array. Fig. 5.29(b) shows the temperature distribution of the ideal phase assignment altered by the proximity effect only from MATLAB.

The obvious difference of the temperature distribution pattern can be observed from Fig. 5.29(a) and (b). No apparent temperature gradient is visible in Fig. 5.29(b). With thermal insulation structures, for a target pixel, the heat energy generated at this pixel is well-confined. The curve of the proximity effect in Fig. 5.28(a) decays faster than the number of pixels outside this target pixel. Beyond a certain distance from this target pixel, the sum of temperature elevations due to thermal crosstalk from its neighbor pixels becomes zero. In addition, when the array size is large, the ideal phase shifts are distributed fairly uniformly across the entire array and the average phase shift is close to π. Combining these two reasons, the proximity effect in Fig. 5.28(a) raises the system temperature near uniform level instead of leading to high temperature gradients on the system scale.

The temperature distribution in Fig. 5.29(a) indicates a hot spot at the array center. According to equation (3.5), a temperature gradient must exist to create a heat flux so the heat from the pixels at the array center can flow outward. This explains the observed temperature gradient pattern in Fig. 5.29(a). In addition, the heat dissipated at the driver circuit, which is closely integrated with the OC in our architecture, contributes to the system-level thermal gradient which is not modelled in the proximity effect in Fig. 5.28(a).

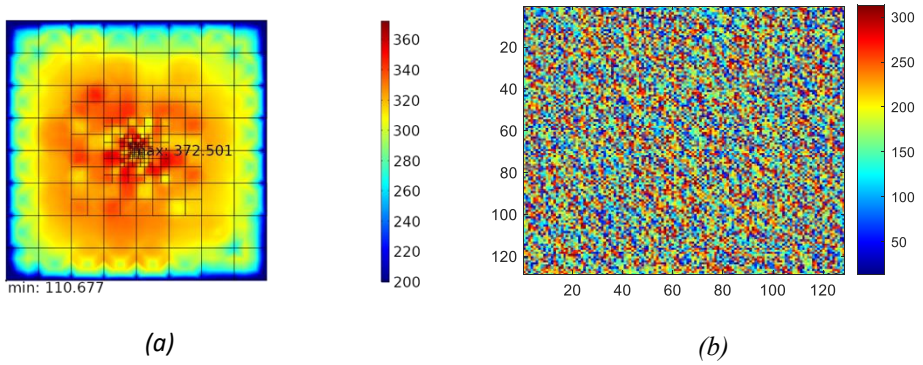


Fig. 5.29 Temperature distribution at the phase shifters of the 128x128 resolution butterfly image (a) thermal simulation with simplified heat source assignment (COMSOL simulation) (b) ideal phase assignment convoluted with the proximity effect matrix (Matlab simulation)

In the simplified model, the thermal crosstalk between pixels is large because heat confinement structures are omitted, and heat sources are applied to all the pixels uniformly. To verify that the chip-level temperature gradient still exists when the thermal crosstalk is minimized, the simulation model is modified as shown in Fig. 5.30. Two 8 x 8 sub-arrays of pixels with air thermal insulation structures are included, as are metal connections between the OC and the MB. There are disposed as follows: one at the center and one at the corner. In this way, the thermal proximity effect at these locations in this new set of simulations responds similarly to the curve in Fig. 5.28 (a). Temperature distributions of only the two target arrays and the entire system being as heated by the butterfly image input are compared.

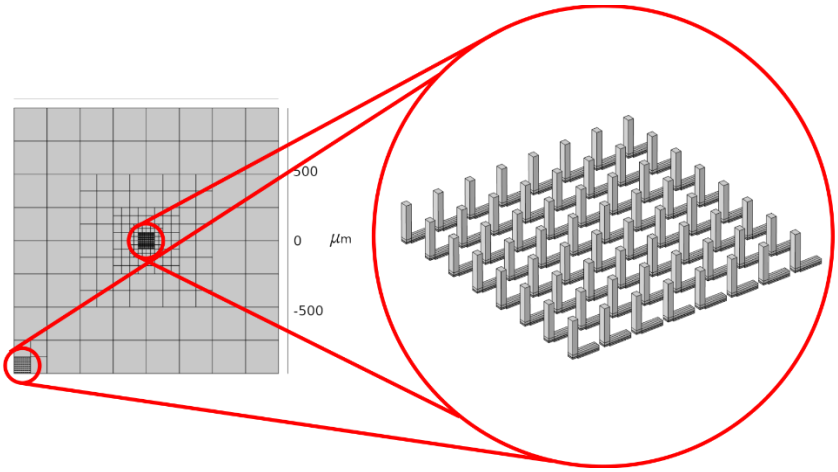


Fig. 5.30 Thermal simulation model for verifying chip-level temperature gradient

The simulation result for a single fully on heater is shown in Fig. 5.31(a). This shows that the heat is well-confined in the target pixel. The thermal crosstalk between the phase shifters in the target sub-arrays and the rest of chip is limited by the air. The comparison of the proximity effect between the simplified model and the original model in Fig. 5.31(b) shows that the simplified model has less thermal crosstalk between pixels. Thus, this model is suitable for checking the actual chip-level temperature profile, for either the uniformly distributed case or for the case that the hottest at the center (when the thermal crosstalk is minimized).

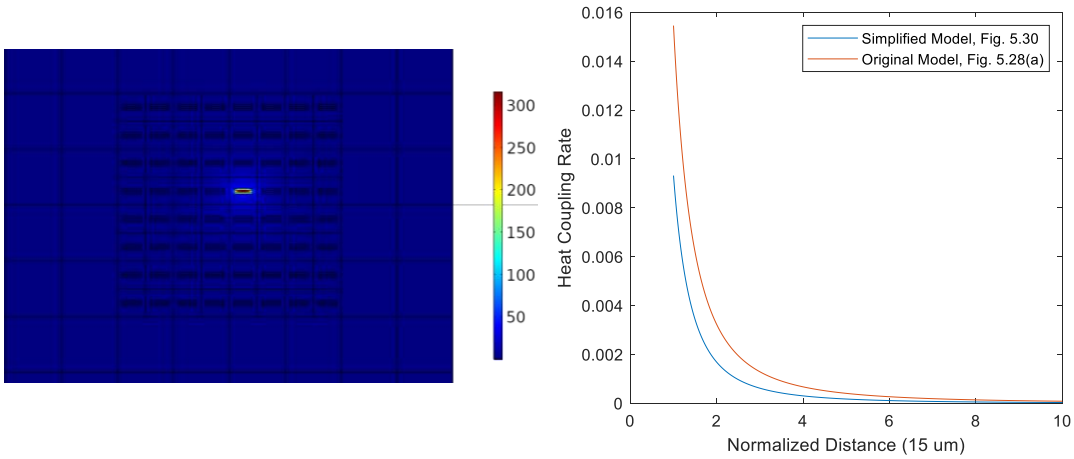


Fig. 5.31 Simulation result with single heater fully on (a) temperature distribution and (b) Proximity effect comparison between the original model, the curve in Fig. 5.28(a), and the simplified model

Next, heat sources for generating the butterfly image are applied to the two target sub-arrays only and to the entire system respectively. The temperature distributions at the center and at the corner sub-arrays are shown in Fig. 5.32. Comparing the results in Fig. 5.32(a) with Fig. 5.32(c) and Fig. 5.32(b) with Fig. 5.32(d), the local temperature relationships between the pixels are preserved with or without heat sources being applied to the entire system. This means the proximity effect relationship extracted from partial heat source assignments can be used to estimate the local relative temperature variations for the complete simulation. However, the average temperatures in Fig. 5.32(c) and (d) are higher than the average temperatures in Fig. 5.32(a) and (b). The

system temperature has to be higher as more heat is drawn into the heat sink.

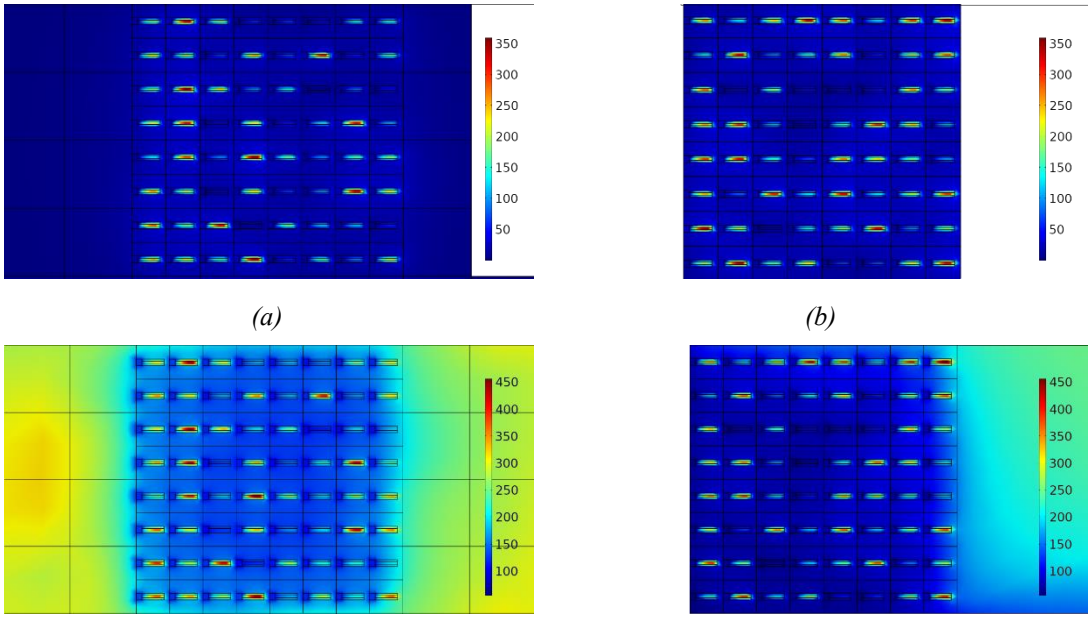


Fig. 5.32 Temperature distributions at (a) center sub-array and (b) corner sub-array when heat sources for rendering the butterfly image are applied to target sub-arrays only, and temperature distributions at (c) center sub-array and (d) corner sub-array when heat are applied to the entire system.

Table 5.2 summarizes the average temperatures at two target sub-arrays computed under different assumptions. The ideal average temperature at the corner sub-array is about 20 °C above the average temperature at the center sub-array. This relationship still holds for both results from as they are ascertained by convoluting the ideal temperature distribution with the proximity effect matrix of the original model and the simplified model. The difference between the average temperature at the corner and the center is smaller in the former case because the proximity effect ratio is higher in the original model and the pixels at array center are affected more by the thermal crosstalk.

The result of COMSOL thermal simulation having heat sources only at the sub-arrays is close to the convolution result. It validates again the convolution method for estimating the proximity effect. For the last case, heat sources are applied to the entire

array in thermal simulation. This system-level simulation would give a more accurate temperature distribution estimation than the previous methods. The average temperature at the center becomes 29 °C higher than the temperature at the corner which reverses the ideal relative temperature differences. At system level when the array size is large and the total amount of heat stresses the heat sink, the proximity effect matrix alone does not faithfully reflect the actual temperature distribution. Therefore, an extra term representing the phase errors caused by the temperature gradients should be superimposed on the phase distribution altered by the proximity effect to estimate the total phase errors accurately.

Table 5.2 Comparison of the computed average temperatures at the target sub-arrays with different methods and conditions

	Average Temperature at the Center Sub-Array (°C)	**Average Temperature at the Corner Sub-array (°C)**
Ideal	119.25	140.92
MATLAB – Convoluted by the Proximity Matrix from the Original Model	156.43	167.31
MATLAB - Convoluted by the Proximity Matrix from the Simplified Model	136.31	154.06
COMSOL - Heat Sources Applied to the Target Sub-Arrays only	130.44	153.22
COMSOL - Heat Sources Applied to All Pixels	239.48	210.68

The rest of this section provides simulation results quantifying the impact of phase error on image forming and beam steering. The sources of the phase error discussed above were added to the ideal desired phase output step-by-step. The image degradation

and the steering angle deviation were quantified at each step. The expected output image and beam angle of a NPA using the feedback control driver was computed using the circuit measurement data mentioned above. The results are compared to demonstrate the importance of a feedback control by in-situ temperature sensing in a large scale NPA system for phase error minimization.

5.4.2 Error Analysis for VR Imaging Forming

A set of test images were used to examine the image quality losses by the phase errors. Because our NPA does not modulate the amplitude of the output beams, the ideal output phases cannot be obtained by applying inverse Fourier transform to the test images directly. The ideal output phases with constant amplitude for rendering the test images are obtained using the Gerchberg-Saxton (GS) algorithm [95].

The algorithm is an iterative process which performs a series of conversions between the phased array output and the far field pattern, going back and forth via Fourier and inverse Fourier transform pairs. During each iteration, the phase information is kept and the amplitudes of the output and far field are changed to the constraint values (the constant and the absolute value of the target image, respectively.) The initial condition of the output phases can be randomly assigned but most often we use the phases of the inverse Fourier transform of the target image. The entire process stops until the difference between forward and backward transformation of the NPA output and far field pattern is smaller than a predefined tolerance.

The output phases obtained by GS algorithm were used as the reference for the error analysis. The phase errors discussed above were added to this reference step-by-step. To quantify the image with phase errors objectively, we used the Structural Similarity Index (SSIM) [112], a widely adopted metric for comparing the similarity between a reference and a distorted image. The value is range from 0 to 1 and 1 means

the two comparing images are identical.

The simulation results of a butterfly image are shown in Fig. 5.33. The resolution of the test image is 512 x 512. The reference image obtained from GS algorithm is in Fig. 5.33(a). Fig. 5.33(b) shows the image affected by the device mismatch. The assumed error is ±2.5% (a value generally considered reasonable by process engineers). The actual phase of the output beam of the array element (i, j) can be expressed as:

$$\varphi_{\text{Fig.5.33}(b)_{i,j}} = \varphi_{\text{Fig.5.33}(a)_{i,j}}(1 + 0.025 \cdot U[-1, 1]) \tag{5.2}$$

where $\varphi_{\text{Fig.5.33}(a)}$ is the output phase for forming the reference image and U[-1, 1] is a uniform distributed random variable between -1 and 1. It is important to note that the amount of the phase deviation is proportional to the original reference output phase. It is not randomly distributed within a certain constant range, such as ±π/8. The SSIM index is reduced to 0.9018.

Figure 5.33(c) and (d) are the results with phase errors due to the device mismatch plus limited proximity effect. This proximity effect is captured in the 7x7 matrix in Table 5.1(a) and full array extrapolated proximity effect respectively. The output phases can be express as:

$$\varphi_{\text{Fig.5.33}(c)} = \varphi_{\text{Fig.5.33}(b)} * M_{PE\_7X7} \tag{5.3}$$

and

$$\varphi_{\text{Fig.5.33}(d)} = \varphi_{\text{Fig.5.33}(b)} * M_{PE\_ext} \tag{5.4}$$

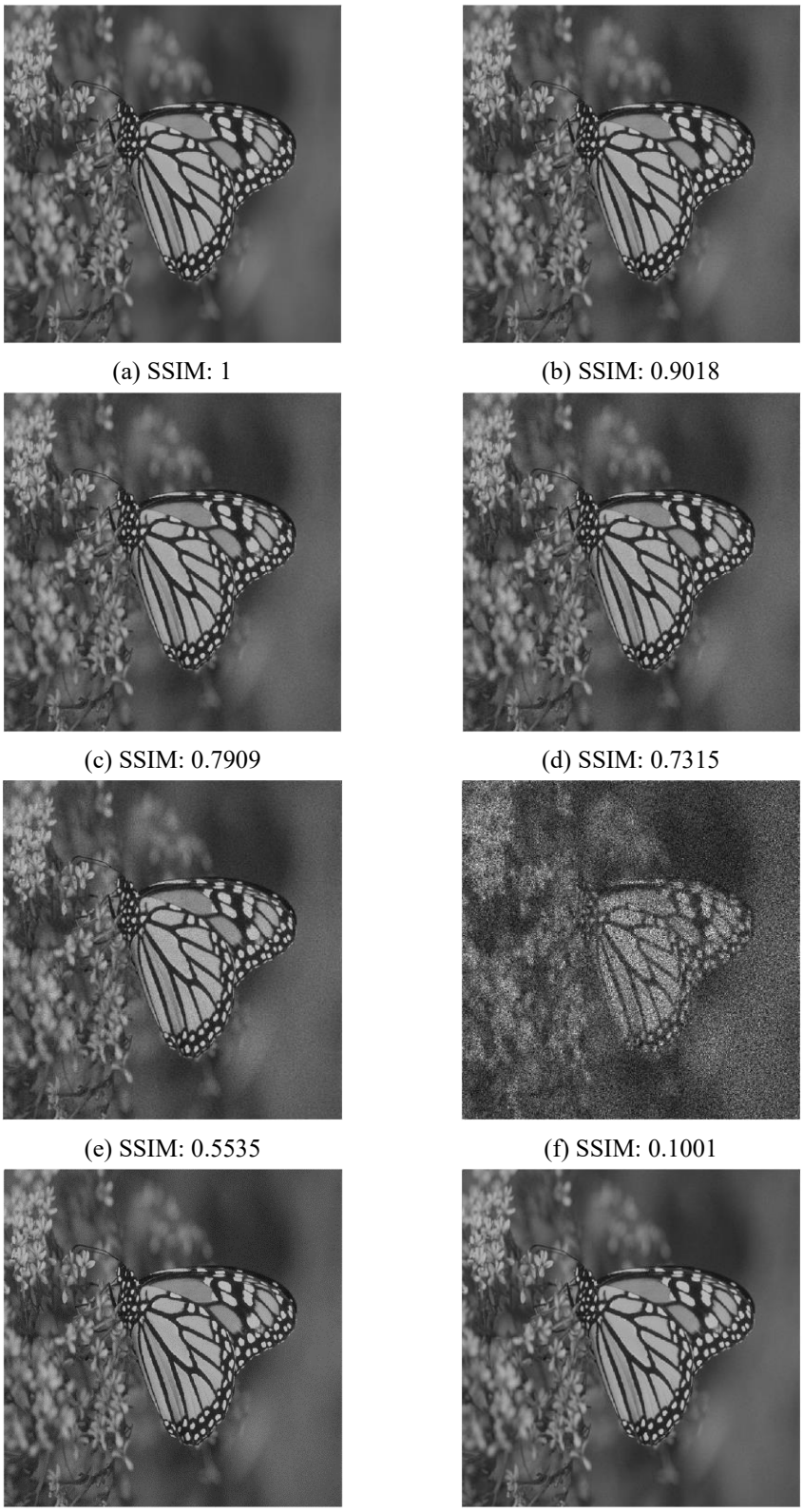
where the * denotes the convolution operation and M_{PE_7X7} and M_{PE_ext} are the local and extrapolated proximity effect matrix. The background noise is increased and the SSIM index is further decreased as expected.

Figure 5.33(e) gives the result image with all sources of phase error including the one caused by the temperature gradient. The mathematical expression for the output phases is:

$$\varphi_{\text{Fig.5.33}(e)} = \varphi_{\text{Fig.5.33}(d)} + M_{T\_gradient} \tag{5.5}$$

where $M_{T\_gradient}$ is the matrix introducing the temperature gradient due to local heat concentration effects. Finally, Fig. 5.33(f) shows the worst-case scenario. All phase errors are included as Fig. 5.33(e) and the proximity effect is replaced by the one without trench isolation. The image is extremely blurred and the SSIM index is only 0.1001.

By introducing different sources of phase error step-by-step, we can clearly understand how the image quality is lowered by each term of phase error. If only the device mismatch is considered, the image is barely affected. A NPA can tolerate some moderate random phase errors because these errors are averaged out across the entire array. However, when we include the proximity effect and heat aggregation at the center of the array which is critical for a compact and highly integrated electro-optical system, the image quality is severely deteriorated. Although the butterfly in the figure can still be distinguished, high background noise is observed.



(a) SSIM: 1 (b) SSIM: 0.9018

(c) SSIM: 0.7909 (d) SSIM: 0.7315

(e) SSIM: 0.5535 (f) SSIM: 0.1001

(g) SSIM: 0.8948 (h) SSIM: 0.9174

Fig. 5.33 Phase error analysis results of image forming (a) reference, (b) +/- 2.5% random error, (c) limited proximity effect (7x7 matrix), (d) extrapolated proximity effect, (e) with additional temperature gradient error, (f) proximity effect without thermal barriers, (g) constant worst-case regulation error (feedback control), and (h) output dependent regulation error (feedback control)

Finally, Fig. 5.33(g) and (h) show the estimated image using the broken-loop driver circuit measurement results provided in the previous section. For Fig. 5.33(g), we assume a worst case where a 3.6 % of phase error is applied to all the output phases. This can be express as:

$$\varphi_{\text{Fig.5.33}(g)_{i,j}} = 1.036\ \varphi_{\text{Fig.5.33}(a)_{i,j}} \tag{5.6}$$

The image quality of Fig. 5.33(g) is about the same as Fig. 3(b) using our error measure criteria in which only the device mismatch error is included. For Fig. 5.33(h), the amount of phase error is derived from Fig. 5.23. The regulation error is a function of the desired output phase so the total phase shift is:

$$\varphi_{\text{Fig.5.33}(h)_{i,j}} = f(\varphi_{\text{Fig.5.33}(a)_{i,j}}) \cdot \varphi_{\text{Fig.5.33}(a)_{i,j}} \tag{5.7}$$

where $f(\varphi_{\text{Fig.5.33}(a)_{i,j}})$ is given in Fig. 5.34. The highest phase error is 3.6 %. This is the worst-case error manifest in Fig. 5.33(g). But this error only exists in two intervals of the desired output phase. The regulation error is less than 1 % for most output phase. The SSIM index with this assumed error pattern is 0.9651. These results demonstrate that the control circuit can restore the image quality and is an effective way to offset the phase error.

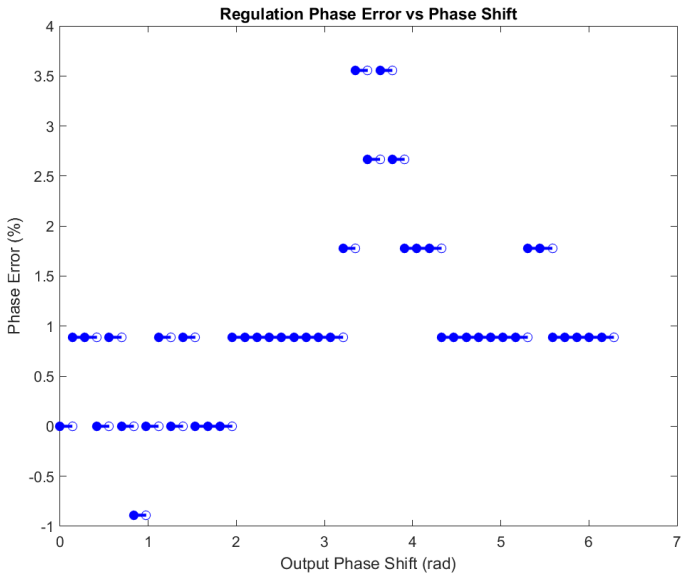
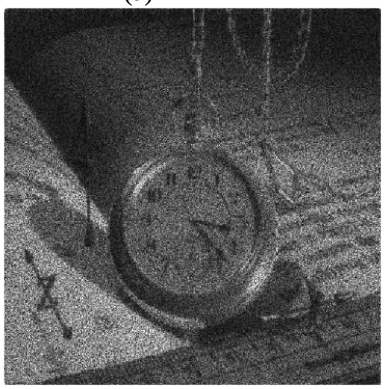


Fig. 5.34 Broken-loop feedback phase regulation error as a function of the output phase shift

More results using different test images are given in Fig. 5.35. Two reference images without any phase errors are shown in Fig. 5.35 (a) and (e). Fig. 5.35 (b) and (f) are images degraded by all three phase error terms. The proximity effect in these two cases is limited by the air insulation structures. Despite this issue, the images are still recognizable, although the background noise does increase. The SSIM index are 0.57 and 0.38 respectively. Fig. 5.35 (c) and (g) show the images with high proximity effect without air insulation structures. The image qualities of both images are extremely low. The items in the images are barely recognizable. Fig. 5.35 (d) and (h) show that for both cases, the image quality is restored by the broken-loop control mechanism.

(a) SSIM: 1

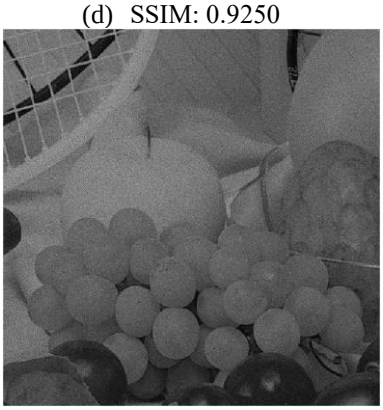
(b) SSIM: 0.5775



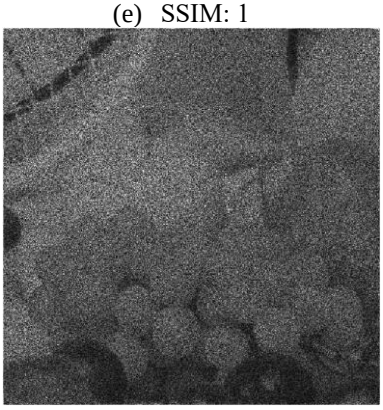
(c) SSIM: 0.1372

(d) SSIM: 0.9250

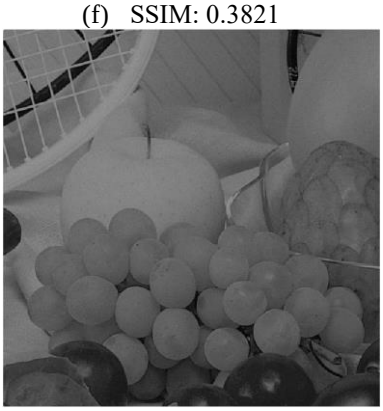
(e) SSIM: 1



(f) SSIM: 0.3821



(g) SSIM: 0.0615



(h) SSIM: 0.8955

Fig. 5.35 More simulation results using different test images

5.4.3 Error Analysis for Beam Steering

To demonstrate the effectiveness of our driver circuit in restoring the steering performance degradation due to random and thermal related phase errors, we simulated the far field interference intensity of a 512-by-512 uniformly spaced phased array. The pixel pitch of the phased array is 15 μm, the same as the value in the VR phase error study, but the wavelength of the light is 1550 nm. Because the pixel pitch is much larger than half wavelength, the field of view is limited to ±3° by the grating lobes.

The intensity at any point in the far field can be expressed as:

$$I(\theta,\rho) = \left| \sum_{N=1}^{512} \sum_{M=1}^{512} e^{i\,(2\pi d \frac{N\sin(\theta)\cos(\rho)+M sin(\theta)\sin(\rho)}{\lambda} + \varphi(N,M))} \right|^2 \tag{5.8}$$

where N and M are the array indices, θ and ρ are the far field position in spherical coordinate, d is the pixel pitch, λ is the wavelength and $\varphi(N,M)$ is the phase shift introduced by the phase shifter including the phase error of the element (N, M). The simulation sweeps through θ and ρ. The values of θ and ρ that give a peak intensity is the steering angle coordinate for a given input phase shift.

The performance degradation is evaluated in two ways. The first is the deviation of the steering angle. The phase error changes the value of θ and ρ that produce a peak intensity. The angle between the reference beam without phase error and the deviated beam is quantized. The second performance metric is the beam contrast. With phase errors, the phase differences between adjacent pixels are not uniform. This leads to optical energy redistribution between the main lobe and side lobes: the energy at the main lobe decreases while the energy at the side lobes increases. The intensity of the main beam is reduced, and the background noise becomes larger. The maximum intensity of the beam without phase errors is used as a reference to be compared with the reduced beam intensity with phase errors.

Two possible phase shifter placements of a beam stearing NPA are shown in Fig. 5.36. The phase shifters can be either located along the optical light bus, Fig. 5.36(a) or

after the individual coupler of each pixel, Fig. 5.36(b). Again, the phase differences between the adjacent pixels are all the same in beam steering. For the placement in Fig. 5.36(a), because the light pass successively through all the phase shifters, the phase delay introduced by each phase shifter is the same and the phase control can be simplified.

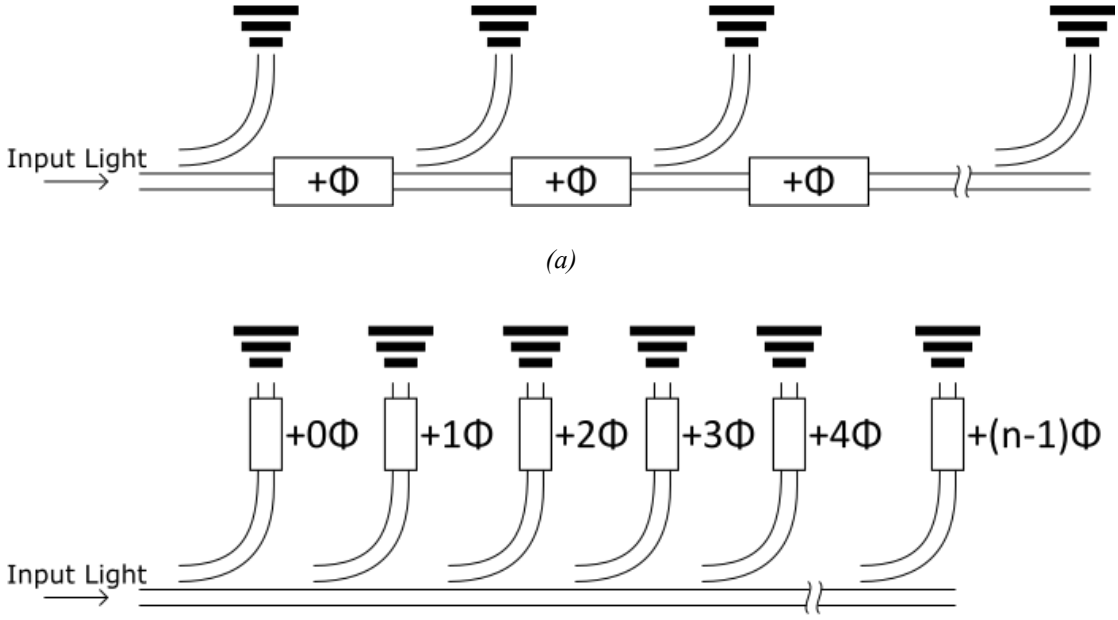


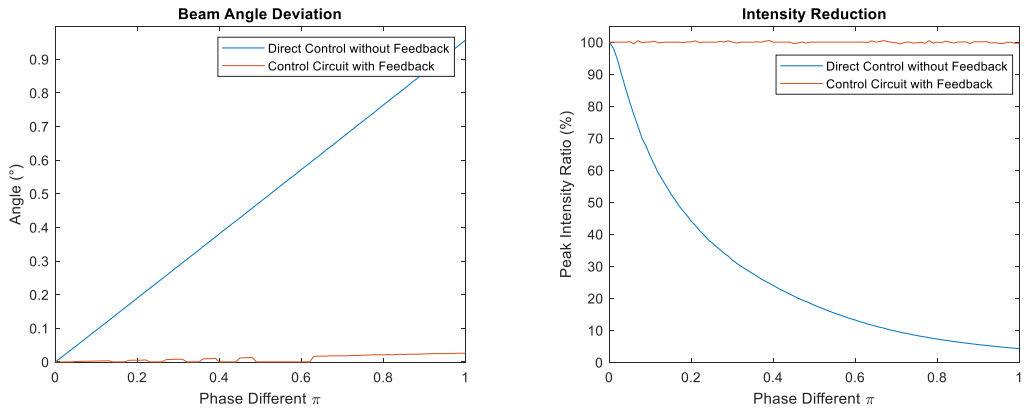
Fig. 5.36 Possible phase shifters placements of a beam steering NPA (a) along the optical light bus (b) after the couplers of each pixel

In constrast, each phase shifter Fig. 5.36(b) has to create an extra φ delay that the phase delay of its predecessor. The phase delay can be wrapped around if it exceeds 2π to minimize the power dissipation. This phase shifter layout arrangement requires individual control to all the phase shifters. Besides the difference of phase control requirement between the two designs in Fig. 5.36, the required power for steering a beam is different. For Fig. 5.36(a), it is straightfoward that the input power is proportional to the steering angle. When the steering angle is small and pointing in positive direction, the phase shift, φ, is small. This means the temperature increment at the phase shifters and the heater power are also small. As the steering angle increases

or even shifts to the negative direction (phase shift φ is greater than π), the heater power increases and more heat will be concentrated at the center of NPA. That is, the phase errors are related to the steering angle. On the other hand, for Fig. 5.36(b), when n is large, the average phase shift of all phase shifters is around π for most steering angle. The phase errors are less dependent on the steering angle.

As mentioned before, our digital driver circuit cannot provide independent phase control to each pixel due to the area constraint. The NPA topology of the simulation is based on the layout in Fig. 5.36(a).

The simulation result is shown in Fig. 5.37. The phase difference between adjacent pixels along one direction was swept from 0 to π and the other direction was fixed at 0.25π. Without feedback control, the beam deviation increases and the peak intensity decreases with the steering as expected because of the reason explained above. When the phase difference is π, the angle deviation is 0.95° and the intensity is below 5% of the intensity without phase error.



(a) *(b)*

Fig. 5.37 Beam steering phase error simulation result (a) angle deviation and (b) peak intensity reduction

On the other hand, the maximum main-beam pointing angle deviation is only 0.026° and the peak intensity is almost unaffected using our feedback control mechanism. For any input phase delay, the regulation error was assumed to be fixed

which corresponds to the load resistance being varied by 20%. The added phase error to each pixel was the same. The phase differences between all adajecnt pixels were constant except for their being deviated from the ideal value. Thus, only the beam angle deviation was seen in the simulation for the control case. In reality, the regulation errors will be varied among the pixel because of different operating conditions. The intensity is expected to be close but slighty lower than the ideal value. As the beam is steered to negative direction, the angle deivation and intensity reduction are expected to be much greater.

To sum up, this analysis shows that our driver circuit, using broken-loop feedback control, effectively offsets the performance degradation in both image forming and beam steering applications. In the image forming application, the background noise is surpressed and the SSIM is improved to around 0.9. In the beam forming application, the angle deviation is kept below 0.026° and the intensity is maintained.

We demonstrate the first NPA driver circuit with in-situ temperature feedback control to minimize the phase error. The control circuit is compact and only occupies 15 μm x 15 μm per pixel matching the pixel size of the target NPA.

Chapter 6: Conclusions and Future Work

6.1 Conclusions

This thesis demonstrates the design and realization of a novel and compact driver circuit for thermo-optic phased arrays in TSMC 65 nm. The unique circuit topology and control scheme together achieve a scalable integrated phased array system by matching the size of a driver circuit unit with the size of a phased array pixel. An integrated thermocouple/heater device was designed to overcome the difficulty of realizing two functions with a single electrical contact between each driver unit and its corresponding phased array pixel. The broken loop feedback control method realizes the first real time pixel-level temperature regulation system. I summarize the efforts performed in completion of this thesis:

1. A comprehensive simulation-based thermal study for a scalable integrated NPA system was conducted. This includes a detailed single pixel electrical-thermal cross domain simulation for guiding the driver circuit design, simulations of a small array to explore the thermal crosstalk, time-domain simulations for characterizing transient response, and simplified system-level simulations for investigating system thermal properties and limitations. The transient simulation results show the pixel can reach a target temperature within 50 μs. The model simplification methodology significantly reduces the required computation resources and simulation time.
2. A thin film gold-nickel thermocouple device was designed and fabricated. The experimental measurements for characterizing the thermoelectric properties of the device were provided. The results proved the concept of achieving dual functions, sourcing current and measuring temperature,

using a single integrated device and a single electrical contact.

3. Design, implementation, and experimental results of the control driver circuit for realizing the unique broken-loop feedback temperature regulation method were described. I provided two design approaches and compared them to one another. In 65 nm technology, the analog driver unit on average meets the main design constraints of 15 m -by- 15 m per pixel and sourcing enough current for a 2π phase shift. The possibility of porting the digital driver design to a more advanced technology node (28 nm) to meet the area constraint was explored.
4. A novel broken-loop feedback control method was demonstrated independent of the optical part of the completed system. The temperature regulation errors were quantified by intentionally varying the resistance of load and sensing resistors which simulated a thermocouple/heater device. The maximum and average error is 3.6% and 1% respectively with ±20% and ±10% variations of the load and sensing resistors respectively.
5. A complete phase error of an NPA system study was conducted. Multiple sources of phase errors were carefully identified with the aid of COMSOL thermal simulation. The performance degradations of image forming and beam steering resulting from phase errors of direct control method were quantified. The results demonstrated the broken-loop feedback control method effectively improves the image quality (on average increasing the SSIM index from 0.5 to 0.9) and steering accuracy (at π phase shift, a 0.026° of angle deviation and near 100% intensity as compared to a 0.95° of angle deviation and only 10% intensity of the ideal intensity without feedback control). Thus, this control mechanism is critical for achieving a high

performance integrated NPA system.

6.2 Future Work

Much work can be done to improve the performance of the control driver circuit described in this thesis and to complete the integrated NPA system. Important aspects to be considered in the future work on the control circuit and the NPA system. Such tasks would include:

1. The control error can be reduced by increasing the sensitivity of the thermocouple temperature sensor to achieve a higher signal-to-noise ratio. Possible approaches include using different materials or designing a more sophisticated process flow to fabricate multiple thermocouple junctions in series using multiple metal layers in a pixel.
2. Reducing the power requirement for 2π phase shift so a larger phase array can be realized. This must be done together with improving the sensitivity of the temperature sensor if the power reduction is achieved by lowering the required temperature raise. Lowering the 2π temperature also reduces the maximum feedback signal.
3. The actual process flow of flip-chip bonding the optical chip and the motherboard to complete the whole system needs to be investigated.
4. Different realizations of the components in the control loop can be explored to achieve higher performance such as lowering the offset of the sense amplifier. Investigation of implementing the control circuit in more advanced technologies is worth conducting, so the size of the control circuit can be shrunk with the size of a NPA pixel to obtain a larger field of view.
5. This work focuses on the design and implementation of the driver circuit unit itself. The design of other peripheral circuits, e.g., circuits for

generating the switching signals and for distributing control signals to individual driver unit/group, and the integration between them and the driver unit must be done to complete the entire electrical system.

Appendix A: Heater Verilog-A Model

```
// VerilogA for heater_model, veriloga

`include "constants.vams"
`include "disciplines.vams"

module heater_model_new(Iin, Iout, Vtemp, Vthermo, gnd, probe_1, probe_2, probe_3);

	inout Iin, Iout, gnd;
	output Vtemp, Vthermo, probe_1, probe_2, probe_3;

	electrical Iin, Iout, gnd, Vtemp, Vthermo;
	electrical probe_1, probe_2, probe_3;
	electrical res_mid_p, res_mid_n;
	electrical n_Vtemp_src;

	branch(Iin, res_mid_p) b_pos_half;		//1st seg of heater
	branch(res_mid_p, res_mid_n) current_probe;
	branch(res_mid_n, Iout) b_neg_half;		//2nd seg of heater

	branch (n_Vtemp_src, Vtemp) b_Vtemp_res;
	branch (n_Vtemp_src, gnd) b_Vtemp_src;
	branch (Vtemp, gnd) b_Vtemp_cap;

	branch (Vthermo, gnd) b_Vthermo;

	// Probes for debug
	branch (probe_1, gnd) b_probe_1;
	branch (probe_2, gnd) b_probe_2;
	branch (probe_3, gnd) b_probe_3;

	//Can be changed in the schematic editor to match different thermal response
	parameter real cap_coeff = 2f;
	parameter real res_coeff = 1G;
	parameter real thermocouple_gain = 10u;
```

```
    real power;
    real power_to_temp;
    real power_coeff;

    parameter real R0 = 517;
    parameter alpha = 0.003012;
    parameter real power_gain = 78125;

    analog begin

        V(b_pos_half) <+ (R0/2)*(1+alpha*V(b_Vtemp_cap))*I(b_pos_half);
        V(b_neg_half) <+ (R0/2)*(1+alpha*V(b_Vtemp_cap))*I(b_pos_half);
        V(current_probe) <+ 0;

        power = (V(b_pos_half)+V(b_neg_half)) * I(current_probe);
        power_to_temp =  power_gain*(V(b_pos_half)+V(b_neg_half)) * I(current_probe);

        V(b_Vtemp_src) <+ power_to_temp;
        V(b_Vtemp_res) <+ res_coeff * I(b_Vtemp_res);
        I(b_Vtemp_cap) <+ cap_coeff * ddt(V(b_Vtemp_cap));

        V(b_Vthermo) <+ 2.5-thermocouple_gain* V(b_Vtemp_cap);

        V(b_probe_1) <+ 0;
        V(b_probe_2) <+ power_to_temp;
        V(b_probe_3) <+ power;
    end
endmodule
```

Bibliography

[1] K. F. Braun, "Electrical oscillations and wireless telegraphy," *Nobel Lecture,* pp. 226-245, 11 Dec. 1909.

[2] D. Parker and D. Zimmermann, "Phased arrays—Part 1: Theory and architectures," *IEEE Trans. Microw. Theory Techn.,* vol. 50, no. 3, p. 678–687, 2002.

[3] D. Parker and D. Zimmermann, "Phased arrays—Part 2: Implementations, applications, and future trends," *IEEE Trans. Microw. Theory Techn.,* vol. 50, no. 3, p. 688–698, 2002.

[4] C. L. Dolf, "A current distribution for broadside arrays which optimizes the relationship between beamwidth and sidelobe levels," *Proc. IRE,* vol. 34, pp. 335-348, 1946.

[5] C. A. Balanis, Antenna Theory: Analysis and Design, 4th ed, Hoboken, NJ: Wiley, 2016.

[6] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag,* vol. 5, no. 2, p. 4–24, Apr. 1988.

[7] J. S. Herd and M. D. Conway, "The evolution to modern phased array," *Proc. IEEE,* vol. 104, no. 3, pp. 519-529, Mar. 2016.

[8] R. L. Haupt and Y. Rahmat-Samii, "Antenna array developments: A perspective on the past, present and future," *IEEE Antennas Propag. Mag,* vol. 57, no. 1, pp. 86-96, 2015.

[9] G. M. Rebeiz and J. B. Muldavin, "RF-MEMS Switches and Switch Circuits," *IEEE Microw. Mag.,* vol. 2, no. 4, p. 59–71, Dec. 2001.

[10] D. F. Sievenpiper, J. H. Schaffner, H. J. Song, R. Y. Loo and G. Tangonan, "Two-dimensional beam steering using an electrically tunable impedance surface," *IEEE Trans. Antennas Propag.,* vol. 51, no. 10, p. 2713–2722, Oct. 2003.

[11] E. Topak, J. Hasch, C. Wagner and T. Zwick, "A novel millimeter-wave dual-fed phased array for beam steering," *IEEE Trans. Microw. Theory Tech.,* vol. 61, no. 8, p. 3140–3147, Aug. 2013.

[12] S. V. Hum and J. Perruisseau-Carrier, "Reconfigurable reflectarrays and array lenses for dynamic antenna beam control: A review," *IEEE Trans. Antennas Propag.,* vol. 62, no. 1, p. 183–198, Jan. 2014.

[13] L. Y. Ji, Y. J. Guo, P. Y. Qin, S. X. Gong and R. Mittra, "A reconfigurable partially reflective surface (PRS) antenna for beam steering," *IEEE Trans. Antennas Propag.,* vol. 63, no. 6, p. 2387–2395, Jun. 2015.

[14] M. I. Skolnik, Introduction to Radar Systems, 3rd ed., New York, NY, USA: McGraw-Hill, 2001.

[15] R. J. Mailloux, Phased Array Antenna Handbook, Norwood, MA, USA: Artech House, 2018.

[16] E. Brookner, "Phased-array and radar breakthroughs," in *Proc. IEEE Radar Conf.*, Boston, MA, 2007.

[17] A. J. Fenn, D. H. Temme, W. P. Delaney and W. E. Courtney, "The development of phased array radar technology," *Lincoln Lab. J.,* vol. 12, no. 2, p. 321–340, 2000.

[18] E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini and R. Valenzuela, "MIMO radar: An idea whose time has come," in *Proc. IEEE Radar Conf*, Honolulu, HI, Apr. 2004.

[19] D. S. Zrnic, J. F. Kimpel, D. E. Forsyth, A. Shapiro, G. Crain, R. Ferek, J. Heimmer, W. Benner, T. J. McNellis and R. J. Vogt, "Agile-beam phased array radar for weather observations," *Bull. Amer. Meteorol. Soc.,* vol. 88, no. 11, p. 1753–1766, Nov. 2007.

[20] C. Fulton, J. L. Salazar, Y. Zhang, G. Zhang, R. Kelly, J. Meier, M. McCord, D. Schmidt, A. D. Byrd, L. M. Bhowmik, S. Karimkashi, D. S. Zrnic, R. J. Doviak, A. Zahrai, M. Yeary and R. D. Palmer, "Cylindrical polarimetric phased array radar: Beamforming and calibration for weather applications," *IEEE Trans. Geosci. Remote Sens.,* vol. 55, no. 5, pp. 2827-2841, May 2017.

[21] E. Yoshikawa, T. Ushio, Z. Kawasaki, S. Yoshida, T. Morimoto, F. Mizutani and M. Wada, "MMSE beam forming on fast-scanning phased array weather radar," *IEEE Trans. Geosci. Remote Sens.,* vol. 51, no. 5, p. 3077–3088, May 2013.

[22] L. C. Godara, "Applications of antenna arrays to mobile communications. I. Performance improvement, feasibility, and system considerations," *Proc. IEEE,* vol. 85, no. 7, pp. 1031-1060, July 1997.

[23] L. C. Godara, "Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations," *Proc. IEEE,* vol. 85, no. 8, pp. 1195-1245, Aug. 1997.

[24] D. Tse and P. Viswanath, Fundamentals of Wireless Communications, Cambridge, U.K.: Cambridge Univ. Press, 2005.

[25] C. G. Christodoulou, Y. Tawk, S. A. Lane and S. R. Erwin, "Reconfigurable antennas for wireless and space applications," *Proc. IEEE,* vol. 100, no. 7, p. 2250–2261, July 2012.

[26] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas and A. Ghosh, "Millimeter Wave Beamforming for Wireless Backhaul and Access in Small Cell Networks," *IEEE Trans. Commun.,* vol. 61, no. 10, pp. 4391-4403, Oct. 2013.

[27] S. Kutty and D. Sen, "Beamforming for millimeter wave communications: An inclusive survey," *EEE Commun. Surveys Tuts.,* vol. 18, no. 2, p. 949–973, 2016.

[28] P. F. McManamon, P. J. Bos, M. J. Escuti, J. Heikenfeld, S. Serati, H. K. Xie and E. A. Watson, "A review of phased array steering for narrow-band electrooptical systems," *Proc. IEEE,* vol. 97, no. 6, pp. 1078-1096, Jun. 2009.

[29] J. K. Doylend, M. J. R. Heck, J. T. Bovington, J. D. Peters, L. A. Coldren and J. E. Bowers, "Two-dimensional free-space beam steering with an optical phased array on silicon-on-insulator," *Opt. Soc. Amer. Opt. Exp.,* vol. 19, no. 22, p. 21595–21604, 2011.

[30] K. V. Acoleyen, K. Komorowska, W. Bogaerts and R. Baets, "One-dimensional off-chip beam steering and shaping using optical phased arrays on silicon-on-insulator," *J. Lightw. Technol.,* vol. 29, no. 22, pp. 3500-3505, Dec. 2011.

[31] M. J. R. Heck, J. F. Bauters, M. Davenport, J. K. Doylend, S. Jain, G. Kurczveil, S. Srinivasan, Y. Tang and J. E. Bowers, "Heterogeneous silicon photonic integrated circuit technology," *IEEE J. Sel. Topics Quantum Electron.,* vol. 19, no. 4, pp. 6100117-6100117, Jul./Aug. 2013.

[32] A. Yaacobi, J. Sun, M. Moresco, G. Leake, D. Coolbaugh and M. Watts, "Integrated phased array for wide-angle beam steering," *Opt. Lett.,* vol. 39, pp. 4575-4578, 2014.

[33] C. V. Poulton, A. Yaacobi, D. B. Cole, M. J. Byrd, M. Raval, D. Vermeulen and M. R. Watts, "Coherent solid-state lidar with silicon photonic optical phased arrays," *Opt. Lett.,* vol. 42, no. 20, pp. 4091-4094, 2017.

[34] C. V. Poulton, M. J. Byrd, P. Russo, E. Timurdogan, M. Khandaker, D. Vermeulen and M. R. Watts, "Long-Range LiDAR and Free-Space Data Communication With High-Performance Optical Phased Arrays," *IEEE J. Sel. Topics Quantum Electron.,* vol. 25, no. 5, pp. 1-8, 2019.

[35] W. M. Neubert, K. H. Kudielka, W. R. Leeb and A. L. Scholtz, "Experimental demonstration of an optical phased array antenna for laser space communications," *Appl. Opt.,* vol. 33, pp. 3820-3830, 1994.

[36] A. Polishuk and A. Shlomi, "Communication performance analysis of microsatellites using an optical phased array antenna," *Opt. Eng.,* vol. 42, no. 7, pp. 2015-2024, July 2003.

[37] W. S. Rabinovich, P. G. Goetz, M. Pruessner, R. Mahon, M. S. Ferraro, D. Park, E. Fleet and M. J. DePrenger, "Free space optical communication link using a silicon photonic optical phased array," *Proc. SPIE,* vol. 9354, no. 93540B, 2015.

[38] D. Kedar and S. Arnon, "Urban optical wireless communication networks: The main challenges and possible solutions," *IEEE Commun. Mag.,* vol. 42, no. 5, pp. S2-S7, May 2004.

[39] W. S. Rabinovich, C. I. Moore, R. Mahon, P. G. Goetz, H. R. Burris, M. S. Ferraro, J. L. Murphy, L. M. Thomas, G. C. Gilbreath, M. Vilcheck and M. R. Suite, "Free-space optical communications research and demonstrations at the U.S. Naval Research Laboratory," *Appl. Opt.,* vol. 54, pp. F189-F200, 2015.

[40] T. Koonen, "Indoor optical wireless systems: Technology, trends, and applications," *J. Lightw. Technol.,* vol. 36, no. 8, p. 1459–1467, Apr. 2018.

[41] K. Wang, A. Nirmalathas, C. Lim, E. Wong, K. Alameh, H. Li and E. Skafidas, "High-speed indoor optical wireless communication system employing a silicon integrated photonic circuit," *Opt. Lett.,* vol. 43, pp. 3132-3135, 2018.

[42] M. J. Byrd, C. V. Poulton, M. Khandaker, E. Timurdogan, D. Vermeulen and M. R. Watts, "Free-space Communication Links with Transmitting and Receiving Integrated Optical Phased Arrays," in *Proc. Frontiers Opt./Laser Sci.*, 2018, Paper FTu4E.1.

[43] J. Sun, E. Timurdogan, A. Yaacobil, E. S. Hosseini and M. R. Watts, "Large-scale nanophotonic

phased array," *Nature,* vol. 493, no. 7431, p. 195–199, Jan. 2013.

[44] Y. Zheng, H. S. Lee, S. Meyne, A. Y. Petrov and M. Eich, "Waveguide-Based Optical Phased Array," *IEEE Photon. Technol. Lett.,* vol. 25, no. 18, pp. 1826-1828, Sept. 2013.

[45] J. Zhou, J. Sun, A. Yaacobi, C. V. Poulton and M. Watts, "Design of 3D hologram emitting optical phased arrays," *Advanced Photonics 2015, OSA Technical Digest (online) (Optical Society of America, 2015),* paper IT4A.7.

[46] F. Aflatouni, B. Abiri, A. Rekhi and A. Hajimiri, "Nanophotonic projection system," *Opt. Exp.,* vol. 23, no. 16, p. 21012–21022, Aug. 2015.

[47] M. Raval, A. Yaacobi, D. Coleman, N. M. Fahrenkopf, C. Baiocco, T. N. A. G. Leake, D. Coolbaugh and M. R. Watts, "Nanophotonic phased array for visible light image projection," in *Photonics Conference (IPC)*, 2016, pp. 206-207.

[48] C. V. Poulton, M. J. Byrd, M. Raval, Z. Su, N. Li, E. Timurdogan, D. Coolbaugh, D. Vermeulen and M. R. Watts, "Large-scale silicon nitride nanophotonic phased arrays at infrared and visible wavelengths," *Opt. Lett.,* vol. 42, no. 1, pp. 21-24, Jan. 2017.

[49] P. F. McManamon, T. A. Dorschner, D. C. Corkum, L. J. Friedman, D. S. Hobbs, M. K. O. Holz, S. Liberman, H. Nguyen, D. P. Resler, R. C. Sharp and E. A. Watson, "Optical phased array technology," *Proc. IEEE,* vol. 84, no. 2, pp. 268-298, 1996.

[50] M. Raval, A. Yaacobi and M. R. Watts, "Integrated visible light phased array system for autostereoscopic image projection," *Opt. Lett.,* vol. 43, pp. 3678-3681, 2018.

[51] J. Notaros, M. Raval, M. Notaros and M. R. Watts, "Integrated-Phased-Array-Based Visible-Light Near-Eye Holographic Projector," *2019 Conference on Lasers and Electro-Optics (CLEO),* pp. 1-2, 2019.

[52] B. Yoo, M. Megens, T. Chan, T. Sun, W. Yang, C. J. Chang-Hasnain, D. A. Horsley and M. C. Wu, "Optical phased array using high contrast gratings for two dimensional beamforming and beamsteering," *Opt. Express,* vol. 21, pp. 12238-12248, 2013.

[53] B. Abiri, F. Aflatouni, A. Rekhi and A. Hajimiri, "Electronic two-dimensional beam steering for integrated optical phased arrays," *OFC 2014,* pp. 1-3, 2014.

[54] K. V. Acoleyen, W. Bogaerts, J. Jágerská, N. L. Thomas, R. Houdré and R. Baets, "Off-chip beam steering with a one-dimensional optical phased array on silicon-on-insulator," *Opt. Lett.,* vol. 34, pp. 1477-1479, 2009.

[55] T. Kim, P. Bhargava, C. V. Poulton, J. Notaros, A. Yaacobi, E. Timurdogan, C. Baiocco, N. Fahrenkopf, S. Kruger, T. Ngai, Y. Timalsina and M. R. Watts, "A single-chip optical phased array in a wafer-scale silicon photonics/CMOS 3D-integration platform," *IEEE J. Solid-State Circuits,* vol. 54, no. 11, p. 3061–3074, Nov. 2019.

[56] S. Chung, H. Abediasl and H. Hashemi, "A Monolithically Integrated Large-Scale Optical Phased Array in Silicon-on-Insulator CMOS," *IEEE Journal of Solid-State Circuits,* vol. 53, no. 1, pp. 275-296, 2018.

[57] R. Fatemi, A. Khachaturian and A. Hajimiri, "A Nonuniform Sparse 2-D Large-FOV Optical Phased Array With a Low-Power PWM Drive," *IEEE Journal of Solid-State Circuits,* vol. 54, no. 5, pp. 1200-1215, May 2019.

[58] B. E. A. Saleh and M. C. Teich, Fundamentals of Photonics 3rd Edition, Wiley, 2019.

[59] X. Wang, B. Wang, P. J. Bos, P. F. McManamon, J. J. Pouch, F. A. Miranda and J. E. Anderson, "Modeling and design of an optimized liquidcrystal optical phased array," *J. Appl. Phys,* vol. 98, no. 7, p. 073101, Oct. 2005.

[60] C. F. Hsieh, R. P. Pan, T. T. Tang, H. L. Chen and C.-L. Pan, "Voltage-controlled liquid-crystal terahertz phase shifter and quarter-wave plate," *Opt. Lett.,* vol. 31, pp. 1112-1114, 2006.

[61] B. Wang, G. Zhang, A. Glushchenko, J. L. West, P. J. Bos and P. F. McManamon, "Stressed liquid-crystal optical phased array for fast tip-tilt wavefront correction," *Appl. Opt.,* vol. 44, pp. 7754-7759, 2005.

[62] E. G. van Putten, I. M. Vellekoop and A. P. Mosk, "Spatial amplitude and phase modulation using commercial twisted nematic LCDs," *Appl. Opt.,* vol. 47, pp. 2076-2081, 2008.

[63] F. Yaraş, H. Kang and L. Onural, "State of the Art in Holographic Displays: A Survey," *J. Display Technol.,* vol. 6, no. 10, pp. 443-454, Oct. 2010.

[64] R. C. Alferness, "Waveguide electro-optic modulators," *IEEE Trans. Microwave Theory Tech.,*

Vols. MTT-30, p. 1121–1137, Aug. 1982.

[65] X. Xiao, H. Xu, X. Li, Z. Li, T. Chu, Y. Yu and J. Yu, "High-speed, low-loss silicon Mach–Zehnder modulators with doping optimization," *Opt. Express,* vol. 21, pp. 4116-4125, 2013.

[66] G. Kang et al., "Silicon-based optical phased array using electrooptic pin phase shifters," *IEEE Photon. Technol. Lett.,* vol. 31, no. 21, pp. 1685-1688, Nov. 2019.

[67] B.-W. Yoo, M. Megens, T. Chan, T. Sun, W. Yang, C. J. Chang-Hasnain, D. A. Horsley and M. C. Wu, "Optical phased array using high contrast gratings for two dimensional beamforming and beamsteering," *Opt. Express,* vol. 21, pp. 12238-12248, 2013.

[68] L. Ye, G. Zhang, Z. You and C. Zhang, "A 2D resonant MEMS scanner with an ultra-compact wedge-like multiplied angle amplification for miniature LIDAR application," *2016 IEEE SENSORS,* pp. 1-3, 2016.

[69] Y. Wang and M. C. Wu, "Micromirror based optical phased array for wide-angle beamsteering," in *Proc. IEEE Int. Conf. Micro Electro Mech. Syst.*, Jan. 2017, pp. 897–900.

[70] Y. Wang, G. Zhou, X. Zhang, K. Kwon, P.-A. Blanche, N. Triesault, K. Yu and M. C. Wu, "2D broadband beamsteering with large-scale MEMS optical phased array," *Optica,* vol. 6, pp. 557-562, 2019.

[71] T. Baba, "Slow light in photonic crystals," *Nature Photon 2,* p. 465–473, 2008.

[72] T. Zhu, Y. Hu, P. Gatkine, S. Veilleux, J. Bland-Hawthorn and M. Dagenais, "Arbitrary on-chip optical filter using complex waveguide Bragg gratings," *Appl. Phys. Lett.,* vol. 108, no. 101, 2016, Art. no. 101104.

[73] X. Sun et al., "Correcting the Proximity Effect in Nanophotonic Phased Arrays," *IEEE Trans. Visual. Comput. Graphics,* vol. 26, no. 12, pp. 3503-3513, Dec. 2020.

[74] X. Sun, Y. Zhang, P.-C. Huang, N. Acharjee, M. Dagenais, M. Peckerar and A. Varshney, "Proximity Effect Correction for Fresnel Holograms on Nanophotonic Phased Arrays," *2021 IEEE Virtual Reality and 3D User Interfaces (VR),* pp. 353-362, 2021.

[75] M. Peckerar, D. Sander, A. Srivastava, A. Foli and U. Vishkin, *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena,* vol. 25, pp. 2288-2294, 2007.

[76] D. Chu, D. T. Bilir, R. F. W. Pease and K. Goodson, "Submicron thermocouple measurements of electron-beam resist heating," *J. Vac. Sci. Technol. B 20,* pp. 3044-3046, 2002.

[77] J. Komma, C. Schwarz, G. Hofmann, D. Heinert and R. Nawrodt, "Thermo-optic coefficient of silicon at 1550 nm and cryogenic temperatures," *Appl. Phys. Lett.,* vol. 101, p. 041905, 2012.

[78] M. Watts, J. Sun, C. DeRose, D. C. Trotter, R. W. Young and G. N. Nielson, "Adiabatic thermo-optic Mach-Zehnder switch," *Opt. Lett.,* vol. 38, pp. 733-735, 2013.

[79] S. V. Deshpande, E. Gulari, S. W. Brown and S. C. Rand, "Optical properties of silicon nitride films deposited by hot filament chemical vapor deposition," *Journal of Applied Physics,* vol. 77, no. 12, pp. 6534-6541, 1995.

[80] A. R. Zanatt and I. B. G. , "“The thermo optic coefficient of amorphous SiN films in the near-infrared and visible regions and its experimental determination," *Appl. Phys. Exp.,* vol. 6, no. 4, p. Art. no. 042402, 2013.

[81] T. F. Krauss, "Slow light in photonic crystal waveguides," *J. Phys. D: Appl. Phys.,* vol. 40, p. 2666, 2007.

[82] G. Skolianos, A. Arora, M. Bernier and M. Digonnet, "Slow light in fiber Bragg gratings and its applications," *J. Phys. D: Appl. Phys.,* vol. 49, p. 463001, 2016.

[83] "Oculus quest 2: Our most advanced new all-in-one VR headset," Oculus, [Online]. Available: https://www.oculus.com/quest-2/. [Accessed 22 Nov 2021].

[84] "Samsung Gear VR with Controller," The Official Samsung Galaxy Site, [Online]. Available: Available: https://www.samsung.com/global/galaxy/gear-vr/. [Accessed 22 Nov 2021].

[85] "PlayStation VR: Live the game with the PS VR Headset," PlayStation, [Online]. Available: https://www.playstation.com/en-us/ps-vr/. [Accessed 22 Nov 2021].

[86] "Microsoft hololens: Mixed Reality Technology for Business," Microsoft HoloLens | Mixed Reality Technology for Business, [Online]. Available: https://www.microsoft.com/en-us/hololens. [Accessed 22 Nov 2021].

[87] D. M. Hoffman, A. R. Girshick, K. Akeley and M. S. Banks, "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of Vision,* p. 1–30, 2008.

[88] A. O'Connor, "The Claim: 3-D Movies Can Induce Headaches and Sickness," *The New York Times,* Feb. 2010.

[89] M. Wöpking, "Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus," *J. Soc. Inf. Disp.,* p. 101–103, 1995.

[90] M. Mon-Williams, J. P. Wann and S. Rushton, "Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display," *Ophthalmic Physiol Opt.,* vol. 13, no. 4, p. 387–391, Oct. 1993.

[91] J. P. Wann, S. Rushton and M. Mon-Williams, "Natural problems for stereoscopic depth perception in virtual environments," *Vision Res.,* vol. 35, no. 19, p. 2731–2736, 1995.

[92] S. J. Watt, K. Akeley, M. O. Ernst and M. S. Banks, "Focus cues to display distance affect perceived depth from disparity," *J. Vis.,* vol. 5, no. 10, pp. 834-862, Dec. 2005.

[93] G. Mather and D. R. Smith, "Depth cue integration: stereopsis and image blur," *Vision Res.,* vol. 40, no. 25, p. 3501–3506, 2000.

[94] J. W. Goodman, Introduction to Fourier optics, New York: Macmillan Learning, 2017.

[95] R. W. Gerchberg and W. 0. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik,* vol. 35, pp. 237-246, 1972.

[96] T. K. Chan, M. Megens, B.-W. Yoo, J. Wyras, C. J. Chang-Hasnain, M. C. Wu and D. A. Horsley, "Optical beamsteering using an 8 × 8 MEMS phased array with closed-loop interferometric phase control," *Opt. Express,* vol. 21, pp. 2807-2815, 2013.

[97] S. Chung, M. Nakai, S. Idres, Y. Ni and H. Hashemi, "19.1 Optical Phased-Array FMCW LiDAR with On-Chip Calibration," *2021 IEEE International Solid- State Circuits Conference (ISSCC),* pp. 286-288, doi: 10.1109/ISSCC42613.2021.9366004, 2021.

[98] M. R. Lueck, J. D. Reed, C. W. Gregory, A. Huffman, J. M. Lannon and D. S. Temple, "High-density large-area-array interconnects formed by low-temperature Cu/Sn–Cu bonding for three-dimensional integrated circuits," *IEEE Trans. ED,* vol. 59, no. 7, pp. 1941-1947, Jul. 2012.

[99] "Indium-bumping design rule," TLMI Corporation, [Online]. Available: http://tlmicorp.com/services-view/indium-bumping/.

[100] A. Bakker and J. H. Huijsing, "Micropower CMOS temperature sensor with digital output," *Journal of Solid-State Circuits,* vol. 31, no. 7, pp. 933-937, 1996.

[101] G. C. M. Meijer, "Thermal sensors based on transistors," *Sensors and Actuators,* vol. 10, pp. 103-125, 1986.

[102] C. Falconi and M. Fratini, "CMOS microsystems temperature control," *Sensors and Actuators B: Chemical,* vol. 129, no. 1, pp. 59-66, 2008.

[103] S. Zhou and N. Wu, "A novel ultra low power temperature sensor for UHF RFID tag chip," in *2007 IEEE Asian Solid-State Circuits Conference*, Jeju, 2007.

[104] C.-K. Kim, B.-S. Kong, C.-G. Lee and Y.-H. Jun, "CMOS temperature sensor with ring oscillator for mobile DRAM self-refresh control,," *2008 IEEE International Symposium on Circuits and Systems,* pp. 3094-3097, 2008.

[105] R. E. Bentley, Handbook of Temperature Measurement Vol. 3 Theory and Practice of Thermoelectirc Thermometry.

[106] *COMSOL Multiphysics® v. 5.4. www.comsol.com. COMSOL AB, Stockholm, Sweden.*

[107] I. Mudawar, D. Bharathan, K. Kelly and S. Narumanchi, "Two-Phase Spray Cooling of Hybrid Vehicle Electronics," *IEEE Transactions on Components and Packaging Technologies,* vol. 32, no. 2, pp. 501-512, June 2009.

[108] P. Sun and R. M. Reano, "Submilliwatt thermo-optic switches using free-standing silicon-on-insulator strip waveguides," *Opt. Express,* vol. 18, pp. 8406-8411, 2010.

[109] V. Gambin et al., "Impingement cooled embedded diamond multiphysics co-design," *2016 15th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm),* pp. 1518-1529, 2016.

[110] D. Altman, A. Gupta and M. Tyhach, "Development of a diamond microfluidics-based intra-chip

cooling technology for GaN," in *Proc. ASME Internat. Tech. Conf. and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems (InterPACK)*, San Francisco, CA, USA, 2015.

[111] B. Razavi, Design of Analog CMOS Integrated Circuits, Boston, MA: McGraw-Hill, 2017.

[112] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing,* vol. 13, no. 4, pp. 600-612, Apr. 2004.