Modeling, Design and Test of an Integrated Optical Neural Recording Device

by

Brendan Crowley

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2014

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

It has long been a goal of neuroscientists to understand how electrophysiological activity in the nervous system corresponds to, and causes, specific physiological actions. Such knowledge could be used to develop cures for disabilities related to nervous system dysfunction, and to control artificial limbs or restore motion to a paralyzed patient. This has motivated research into technologies, broadly termed brain-machine interfaces, for interfacing with the nervous system. One category of such neural interfaces is implantable neural recording devices, which monitor and record neural signals through a microelectronic device implanted in the body. Typical implantable neural recording devices use a micro-electrode array to record electrical signals simultaneously from a multitude of neurons. Unfortunately, devices employing micro-electrode arrays have several issues from both the biological and circuit design points of view. These issues include tissue damage due to implantation of a micro-electrode array, degradation of recording fidelity over time, limited spatial resolution, the requirement to maintain charge balance in tissue, and the difficulty in implementing low-frequency (large time constant) filter cutoffs with limited chip area. These issues provided the motivation to investigate alternative methods for neural recording - namely optical methods based on fluorescence detection with voltage-sensitive fluorescent proteins. Optical recording methods can alleviate many of the issues with electrical recording, as well as provide other advantages, such as recording targeted to specific neurons/neuron types and higher spatial resolution due to reduced recording site pitch. The major limitations of fully implantable optical recording devices stem from size constraints, the attenuation of light in tissue, which limits imaging depth, and the need for genetically programmed voltage-sensitive fluorescent proteins, which must be introduced to the tissue in the case of chronic recording.

This research began with investigating the feasibility of replacing an electrical neural recording front end with an optical front end - the conclusion being that producing an initial design was worthwhile. Thus, this thesis presents a prototype optical neural recording device for detecting individual spikes in Layer I of the brain. The device is designed for the fully implantable scenario, where space for typical fluorescence imaging optical components is limited, and a high level of integration is required. The thesis describes: 1) Modeling: a general framework for modeling near-field fluorescence detection systems is presented; the model is then extended and applied to the design of the optical neural recording device for detecting individual spikes in Layer I of the brain, taking into account light attenuation in tissue; 2) Design: the design of a high-sensitivity CMOS imaging chip used in the device; 3) Packaging: the packaging of the CMOS imager with LED dies and an excitation filter; and 4) Testing: the experimental results from testing the packaged device with a fluorescent tissue phantom designed to emulate layer I of the brain. Ideas for future work on such devices are discussed.

Acknowledgements

This research work and thesis could not have been completed without the help of numerous people along the way. I have probably missed some people here, and for that I apologize, but I thank you as well.

First, I must thank my supervisor, professor Vincent Gaudet. Vincent was consistently a positive force throughout my graduate career, always providing smart input and excellent guidance, from writing tips to encouring my switch from electrical to optical recording work; providing all of the resources I needed to perform the research, including financial support; providing the opportunities to design, tapeout and test several chips, to join the SMART team, to attend multiple conferences, and even to transfer with him to Waterloo from the University of Alberta, where this degree began. I highly recommend Vincent as a supervisor.

I would like to thank my committee members, professors Peter Levine, Chris Backhouse, Matthijs van der Meer and Sebastian Magierowski for their input during my comprehensive exam and thereafter, and for taking the time to read my thesis and attend my defense. A special thanks goes to Peter Levine for allowing me to investigate optical neural recording as my project in his class; that opportunity really got the ball rolling.

I would like to thank professor David Nairn for sharing much insight into circuit design, as well as career choices.

I must thank all of the students in my research group and outside of my research group over the years who contributed to my learning and terrific graduate experience, including Russell Dodd, Karl Jensen, Joyce Li, Qing Li, Tianchi Ma, Ji Sun, Bahareh Ebrahimi, Adam Bray, Adam Neale, Pierce Chuang, Chris Ceroici, Manpreet Singh and many others. I've been going through

school with Russell Dodd since undergrad, and we both ended up doing our Masters and PhD under the supervision of Vincent. We've worked on many projects and solved many problems together, and provided each other with moral support. Russell certainly made my time in grad studies much easier and more fun as well. Of all my fellow students at Waterloo, Karl Jensen helped me out the most. Whether it was designing the PCB for our OR1 chip, passing on optical modeling knowledge, sharing Matlab ninja skills, or just chatting about all kinds of things over lunch, I benefitted greatly from Karl.

I owe a big thanks to PCB design guru Vincent Di Lello for his generous work on my early PCB designs, and for his tremendous input in later designs once I started doing the designs myself. Vincent was always happy to answer my questions clearly no matter how basic they were.

Professor Anastasia Elias and Theodore Ng generously took time to prepare the tissue phantoms used for fluorescence testing of the OR2 chip.

I must thank the University of Alberta, the University of Waterloo, the federal and provincial governments, including NSERC, AHFMR (through the SMART team) and OGS for all of the funding I have received throughout this degree.

I was very fortunate to be a part of the Sensory Motor Adaptive Rehabilitation Technology (SMART) interdisciplinary team in neuralprostheses, based out of the University of Alberta and led by Vivian Mushahwar. The neural recording work was one of the projects this team collaborated on, and I learned a lot from being a member of the team. My thanks go to all of the team members, who are too numerous to list here.

The ECE computer support staff, and administrative staff, at both universities, made my time

easier. A special thanks goes to Phil Regier for all of his help with CAD tool support.

CMC Microsystems provided essential CAD tools, and chip fabrication and packaging services; a special thanks goes to Jianzeng Xu for his help with unique packaging requirements. Thank you to Optotech and Three-Five Compounds Inc. for providing LED die samples.

Lastly I must thank my friends and especially my family for their continuing love and support. I wouldn't have made it through without you.

This thesis was prepared using LaTeX, Make, the GVIM text editor, Inkscape, Matlab and Python with Matplotlib.

Dedication

To my family and friends.

Table of Contents

List of Tables

List of Figures

Nomenclature

ΔI_{fl}	change in fluorescence intensity
λ_{fl}	fluorophore emission wavelength
μ_a	absorption coefficient of a material
μ_f	total absorption coefficient of tissue
μ_n	absorption coefficient of a neuron
μ_s	scattering coefficient of a material
$\mu_s^{'}$	modified scattering coefficient
μ_t	total coefficient of absorption and scattering
ω	solid angle of emission from a point source [sr]
$\sqrt{N_{be}}$	electrical shot noise due to optical signal
$\sqrt{N_{bs}}$	shot noise of the baseline signal
A_n	surface area of neuron facing the light source [μm^2]
A_{pd}	surface area of a photodetector [μm^2]
c	speed of light
C_{fb}	CTIA feedback capacitance integrating the photocurrent
C_{max}	maximum CTIA feedback capacitance (100fF)
C_{mid}	middle CTIA feedback capacitance (20fF)

C_{min}	minimum CTIA feedback capacitance (4fF)
C_{pd}	photodiode capacitance
d	thickness of a material
d_n	thickness of a neuron [μm]
E_{ph}	photon energy at the fluorophore emission wavelength
g	anisotropy coefficient (cosine of scatter) of a material
h	Planck’s constant
I	light intensity transmitted through a material
I_0	light intensity incident on a material
I_e	LED radiant intensity [W/sr]
$I_{fl,base}$	baseline fluorescence intensity
$I_{fl,peak}$	peak fluorescence intensity
$in\ vitro$	outside of a living organism - ex. tissue sample
$in\ vivo$	inside a living organism
LED	light-emitting diode
N_{bs}	number of baseline photons detected
N_{se}	number of signal electrons detected during T_s
N_s	number of signal photons detected during T_s
NIR	near-infrared
P_0	total power emitted by a point source [W]
P_{pd}	optical power absorbed by photodetector
q	charge of one electron
QE_{pd}	quantum efficiency of photodetector

QY_{fl}	quantum yield of a fluorophore
R	radial distance from a point source [μm]
r_{bs}	number of baseline photons detected per second
R_{nl}	distance between the neuron and light source [μm]
S_{fl}	fluorophore sensitivity
SNR_e	SNR in terms of electrical power
T_s	sampling time
V_{ap}	peak action potential voltage difference
$V_{n,rms}$	RMS noise voltage due to the optical signal
V_s	signal voltage
ADC	analog-to-digital converter
AP	action potential (spike)
API	application programming interface
APS	active pixel sensor
Arch	Archaerhodopsin 3 - a VSFP
BMI	brain-machine interface
BRAM	FPGA block RAM
CMFB	common-mode feedback
CQFP80	80-pin Ceramic Quad Flat Pack chip package
CSD	calcium-sensitive dye
CTIA	Capacitive trans-impedance amplifier
Cy3	Cy3 fluorescent dye
DAC	digital-to-analog converter

DNL	differential non-linearity
DRAM	dynamic random access memory
DRIE	deep reactive ion etching
EEG	electroencephalography
EMCCD	electron-multiplying charge-coupled device
FD	fluorescence detection
FF	fill factor
FOM	figure of merit
FPGA	field-programmable gate array
FPN	fixed pattern noise
FR	frame rate
FWHM	full width at half-maximum (width of spectrum at half of peak intensity)
GECI	genetically encoded calcium indicator
HbO_2	oxygenated hemoglobin
HbR	de-oxygenated hemoglobin
HDL	hardware description language
INL	integral non-linearity
LED	light-emitting diode
LFP	local field potential
LNA	low-noise amplifier
LOD	limit of detection
MEA	micro-electrode array
MUR	multi-unit recording

MUX	multiplexer
PDMS	polydimethylsiloxane
PSD	power spectral density
RAM	random access memory
RMSE	root mean square error
SAR	successive approximation register
SiP	system-in-package
SRAM	static random access memory
SUR	single-unit recording
VSD	voltage-sensitive dye
VSFP	voltage-sensitive fluorescent protein

Chapter 1

Introduction

1.1 Motivation

It has long been a goal of neuroscientists to understand how electrophysiological activity in the nervous system corresponds to, and causes, specific physiological actions. Such knowledge could be used to develop cures for disabilities related to nervous system dysfunction, such as Parkinson's disease [5], and to control artificial limbs or restore motion to a paralyzed patient. This has motivated research into technologies, broadly termed brain-machine interfaces (BMIs), for interfacing with the nervous system. A BMI is a device, either invasive (implantable) or non-invasive, that provides access to neural signals, either through recording or stimulation of neural activity. In theory, functions of the nervous system could be restored artificially if access to the necessary signaling was available.

Past experiments have shown that rats [13], monkeys [98], and paralyzed humans [46], can control computer cursors or robotic arms through neuroprosthetic devices that monitor multiple neural signals. Such experimental results suggest that interfacing with the nervous system is a possible approach for treating paralysis and other disorders and disabilities in humans [55]. One of the more recent examples of an implanted BMI in humans is in [45], where a woman with tetraplegia was able to use a robotic arm to reach out, grab a beverage container, bring it to her mouth, and drink from it.

A state-of-the-art fully implantable BMI device provides wireless transmission of recordings from a 100-channel micro-electrode array (MEA), for hours at a time, with the ability to be recharged [11]. While the capability of such a device is impressive, there is still a long way to go before implantable BMIs become reliable, unobtrusive and widely used in humans. Conventional BMIs rely on microelectrodes to electrically interface with neurons. In order to improve on existing BMIs and overcome some of their limitations, researchers are considering different modalities. In particular, research on *optogenetics* and its use in neural recording and stimulation is becoming more popular. Different groups have begun pursuing optical interfacing based on optogenetic techniques [53], [97], [78], [69], [36]. Optical interfacing has advantages over electrical interfacing, and shows promise for future BMIs. These advantages are discussed in Chapter 2. Optical neural recording systems have been implemented, but are typically restricted to large bench-top systems; there has been some work on shrinking large bench-top systems down to head-mountable systems (for mice) [97], [78], but until very recently there has been very little work done on shrinking these systems down to sizes appropriate for full implantation, as was done in [53].

1.2 Aim and Contributions

The high-level aim of this research work is to advance the state of the art in implantable neural interfaces in order to improve capabilities for learning about nervous system function, as well as to further enable the design of devices for restoration of nervous system function. The focus in this thesis is on neural recording (as opposed to stimulation). More specifically, the goal is advancement of neural recording through the design of implantable optical neural recording devices based on fluorescence detection. The contributions of this work are as follows:

1. An optical modeling framework for prediction of fluorescence detection device performance.
    - The modeling framework can be generalized to other near-field fluorescence detection systems.
    - The framework is used to model the implantable optical neural recording device described below.
2. A CMOS imaging chip designed to detect low-intensity light, which is to be used in an optical neural recording device.
3. A fluorescence detection-based optical neural recording device.
    - The device consists of a CMOS imager, light-emitting diode (LED) die excitation sources, and an optical excitation filter.
    - The device is designed to detect individual spikes in layer I of the brain (a depth of about 300μm).

4. A method for testing an optical neural recording device using a fluorescent tissue phantom.
    - The tissue phantom uses embedded fluorescent microspheres coupled with modulated excitation light to emulate the functionality of neurons containing voltage-sensitive fluorescent molecules.

1.3 Organization

The rest of this thesis is organized as follows. Chapter 2 provides background information on the characteristics of neural signals, and discusses neural recording device specifications. Electrical and optical neural recording methods are compared. Lastly, fluorescence detection and its application to neural recording is described.

Chapter 3 presents work done towards electrical neural recording circuitry.

Chapter 4 presents a low-complexity model for predicting the near-field distribution of light from an excitation source. The model is applied to predicting the light distribution in a CMOS imager chip with LED dies bonded on top of the chip. The design and testing of the first optical chip with LED dies bonded on top of the chip is presented, and the experimentally measured results are compared to the model results.

In Chapter 5 the basic optical model is extended to include light propagation in tissue and fluorescence. The extended model is used to predict the recorded signal in an implantable optical neural recording device consisting of a high-sensitivity CMOS imaging chip with an optical excitation filter and LED dies bonded beside it.

Chapter 6 presents the design, packaging, and testing of the optical neural recording device described in Chapter 5. This includes the design and testing of a second CMOS imager chip, and fluorescence testing with a tissue phantom.

Chapter 7 summarizes the contributions of the thesis and discusses directions for future work.

Chapter 2

Background

This chapter provides background information on the characteristics of neural signals and the characteristics of ideal BMIs. Neural recording device specifications are discussed. Next, electrical and optical neural recording methods are compared. Lastly, fluorescence detection and its application to neural recording is described.

2.1 Characteristics of Neural Signals

The following subsection describes the neural signals that are of interest in neural recording. The description is based on [7] and Chapter 7 in [50].

2.1.1 Action Potentials

Neurons are the excitable cells of the nervous system that are capable of generating *action potentials* (APs), or *spikes*. Note that muscle cells can also generate potentials. Like all cells in the body, neurons have a cell membrane. In the resting state, there exists a potential difference across the membrane, which is called the *resting potential*. The resting potential is due to the electrical charge of ions on either side of the membrane. *Ion pumps* fed by ATP (adenosine triphosphate) maintain the intracellular (inside the cell) and extracellular (outside the cell) ion concentrations in the resting state. The most significant pump is the Na^+-K^+ pump, which exchanges Na^+ ions inside the cell for K^+ ions from outside the cell. Thus, the extracellular Na^+ concentration is high, and the intracellular K^+ concentration is high. The resting potential due to the concentrations of Na^+, K^+ and other ions is about -70mV. This is shown as part A in Fig. 2.1. The figure shows a theoretical plot of an AP. These pulses typically have an amplitude of approximately 100mV relative to the resting potential, and a duration of approximately 1ms.

In addition to ion pumps, the membrane has *voltage-gated ion channels*. These ion channels are made of proteins, and can change their shape. When triggered by an increase in membrane potential, the ion channels open, permitting the diffusion of certain ions across the membrane. When the ion channels are closed, ions are prevented from diffusing across the membrane.

When a stimulus excites a neuron a small increase in the membrane potential occurs, some Na^+ channels open and Na^+ ions diffuse into the cell, carrying with them a net positive charge. If the stimulus is too small, the membrane potential will not increase past the *threshold potential*, no AP will be generated, and the membrane potential will drop back to the resting potential. However, if the stimulus is large enough, the membrane potential increases past the threshold

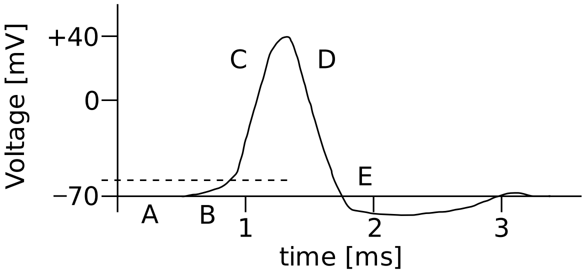


Figure 2.1: Theoretical plot of an action potential. Each letter on the plot indicates a different portion of the action potential.

value, and an AP occurs. See part B in Fig. 2.1 (the dashed line is the threshold potential). When an AP occurs the membrane potential is high enough to open more Na^+ ion channels, increasing the diffusion of Na^+ ions into the cell, and positive feedback continues to sharply increase the membrane potential, but not indefinitely. See part C in Fig. 2.1. Na^+ ion channels will close after about 1ms despite the depolarization of the membrane potential. Also, K^+ ion channels begin to open, allowing K^+ ions to diffuse out of the cell. These channels require a higher membrane voltage to open than do the Na^+ channels. As the Na^+ channels close and the K^+ channels open, the membrane potential is driven back to the resting potential. See part D in Fig. 2.1. The K^+ channels are slower to deactivate, and so there can be an overshoot of the resting potential during repolarization. See part E in Fig. 2.1.

When an AP is generated (the membrane potential exceeds the threshold), the full AP waveform is generated. i.e. it is an all-or-none response. Each spike from the same neuron will have nearly identical amplitude and duration. Therefore, though spikes are inherently analog signals, a spike train can be thought of as a digital signal. An active neuron typically generates 10-100 spikes per second, whereas a resting neuron generates around 1-10 spikes per second [39].

2.1.2 Recording Neural Signals

While the underlying processes of the nervous system are chemical in nature, we indirectly measure them by recording the resulting neural voltage waveforms. Typical electrical recording of neural signals through an implantable device relies on implanting electrodes in the neural tissue. The types of electrodes and their locations determine the type of signals that are recorded.

Cell Potentials

Ideally we could simultaneously measure the full-scale APs (cell potentials, $\sim$100mV) of many individual neurons. With electrical recording this would require *intracellular* penetrating electrodes. Unfortunately, this type of measurement will quickly result in cell death, and cannot be used for chronic implants [39].

Extracellular Signals

Extracellular signals are recorded from electrodes that do not penetrate cell bodies, but rather sit nearby in the extracellular fluid. Extracellular voltages resulting from APs are much lower outside of the cell than inside the cell, and are increasingly attenuated as the distance between the electrode and cell increases. In the case where an electrode is very close to a single neuron, *single-unit recordings* (SURs, spikes from a single neuron) can be obtained.

Another possibility is recording the superposition of signals from several neurons grouped closely together, as either *multi-unit recordings* (MURs) or *local field potentials* (LPFs). MURs are from a few cells, while LFPs are from a larger number of cells. Another option is to record signals that are a result of an even larger number of neurons over several square centimeters, called *electroencephalography* (EEG) [80]. Each of these types of recording has its advantages. For example, EEG may provide information on the big picture of brain intent but neglect the information of individual or a small group of neurons. Conversely, performing only MUR or LFP recordings may neglect the overall intent of the signals. Determining the appropriate control (stimulation) signal to send to the destination in a closed-loop system may require a combination of recording methods [80].

The nature of neuronal signals response to motion was documented four decades ago [27]. It was found that neighboring neurons may exhibit distinct firing patterns for a certain movement, but the firing of a single neuron may vary significantly in different trials of that same movement. The average of firing patterns for many trials is usually quite consistent [55]. These characteristics mean that in a neuroprosthetic device whose goal is to stimulate motion, signals from several neighboring neurons must be simultaneously recorded in order for the device to reliably extract the intent of the signals and respond with appropriate stimulation patterns. Thus, neural recording devices must be multi-channel devices.

It should be noted that several groups have investigated decoding LPFs and other lower frequency neural activity for the purposes of motor control. In their recent review of neural recording techniques, the authors of [48] state that while recordings from LFPs for the purpose of motor control have progressed in recent years and do have merit, LFP recordings still do not perform as well as spikes.

Extracellular Signal Characteristics

Typical single- and multi-unit recordings have frequency content from 100-300Hz up to 7-10kHz, and amplitudes from 50μV to 1mV (depending on the distance from the electrode to the cells) [12], [8]. Due to electrochemical effects at the electrode-tissue interface, there also exist large DC offset voltages (1-2V) at the electrodes [8]. LFPs frequency components can be below 1Hz [8], but typically are in the 10-200Hz range [40]. LFPs and MURs can be separated from a single extracellular recording with reasonably simple filtering since they occupy different frequency bands. It is also possible to distinguish between spikes from different neurons from a single recording [40], [80]. This is usually referred to as *spike sorting*. A single electrode in a MEA may see spikes from 1-4 separate neurons, whereas other electrodes in the array may see no spikes [40].

2.2 Ideal Brain-Machine Interfaces and Design Specifications

Brain-machine interfaces are devices that provide access to neural signals. In this work the focus is implantable neural recording (as opposed to stimulation) devices. From a high-level perspective, an ideal neural recording device would provide the capability to accurately monitor the activity of a population of interconnected neurons and provide a visual map of the neural connections [24]. Additionally, the device would operate in such a manner that it would not inhibit the subject in any way, and would be functional for long periods of time when chronically implanted. In the context of recording from a mouse, for example, the mouse must be able to move freely in a large area, and the device should not place excessive physical stress on the mouse, or alter the normal functionality of its nervous system and other biological systems. Similar requirements exist in the case of recording from humans. These requirements place several constraints on the design of the neural recording device. These constraints are outlined as follows.

2.2.1 Wireless Operation

An implantable device must operate wirelessly to allow for unrestricted movement of the subject and to allow for full *subcutaneous* (under the skin) implantation [8]. Subcutaneous implantation is necessary to prevent infection. A device with wires protruding from the skin provides a path for infection which can be damaging to the patient.

2.2.2 Form Factor

Unfortunately, subcutaneous implantation imposes severe contraints on the size of implanted devices. The limitations will vary depending on implant location. For example, the thickness of the device in [79] is limited to 1-2mm due to the amount of space between the dura and the skull, where the device will be implanted.

2.2.3 Power Supply

Another problem posed by full implantation is access to a power supply. An external wireline power supply is no longer an option. The device may have to include some sort of energy storage solution. Batteries may be an option. However, battery size, and thus lifetime, are severely limited by the size requirements. This is a problem because frequent surgery to replace batteries is highly undesirable. Surgery is costly in terms of money and man-power, and involves some risk to the patient no matter how reliable the surgery may be. Therefore, any batteries must be rechargeable. Inductive power coupling can be used to recharge batteries, or continuous inductive coupling could be used to power the device without a battery. Ideally, energy harvesting of ambient energy within the body could be used to recharge the batteries and power the device indefinitely.

2.2.4 Power Consumption

Perhaps the most important design constraint is power consumption. Low power design is absolutely essential for two reasons. One, as has been discussed, is the difficulty in supplying power to the device. The other is that excessive heat dissipation causes temperature gradients at the implantation site. If too much heat is dissipated, surrounding cells will die, which is not acceptable [29]. A temperature difference of even 1°C is enough to damage cells [38]. Experimental results have shown that power dissipation must be less than $150\mu\text{W/mm}^2$ of chip surface area in order to limit the temperature change to 1°C [38]. Another study has shown that the power consumption of an implanted device must be less than 80mW/cm^2 in order to avoid tissue damage [86].

2.2.5 Biocompatibility

Biocompatibility is an issue with device implantation. The body sees an implanted device as a foreign substance, and encapsulates it with a *glial* (non-neural support cells in nervous tissue) tissue, which can be 50-100μm thick and takes 6-12 weeks to form [17]. This reaction makes electrical recording of extracellular signals more difficult, because impedance is increased and neurons are further away from the electrode [17]. Hence, recording degrades over time [38]. Other problems with implanting devices are possible surgical trauma, and the movement of electrodes over time relative to the surrounding cells due to motion of the subject [17].

2.2.6 Density of Recording Sites

It is desirable to have a high density of recording sites on the device in order to extract information from as many neurons as possible. The density of recording sites will impact device form factor, packaging, data rates, power consumption and potentially biocompatibility.

2.2.7 Data Rates

MEAs include up to 100 channels [65], and future recording systems may include multiple recording sites with multiple MEAs. In order to capture the shape of spikes, each of these channels will be continuously sampled at a sampling rate in the vicinity of 15-20kSps, such that all of the frequency content (up to 7.5-10kHz) can be captured without aliasing. Additionally the recordings should reproduce the actual shape of the waveform, which is necessary for spike sorting algorithms to operate. The desired sample resolution (amplitude resolution) is typically 8 bits.

It is easy to produce tens of megabits of data per second in this case with continuous recording [14], [40]. Given the power consumption constraints the wireless communication circuitry likely cannot reliably transmit these high data rates through tissue. Therefore, the system must include some method for reducing the transmitted data rate [40].

2.2.8 Functionality

The device should have some sort of built-in testing circuitry to verify that the device is working properly, and that it won't harm the patient [71]. One example is a temperature sensor to ensure heating limits are satisfied.

2.2.9 Generic, Cheap, Programmable

In the case of commercial adoption of an implantable device, ideally the developed device would be capable of implantation anywhere in the body, and capable of recording any type of neural signal. It would be wirelessly programmable such that the user could optimize its performance for the specific application [29]. With a wirelessly programmable device, security is also an issue. Without security, anyone with malicious intent could take control of the device. The device should be designed and fabricated with standard technologies to keep costs low.

2.2.10 Extensive Testing

Lastly, extensive testing (electrical, long term biocompatibility) of the devices is imperative. The ultimate goal is to have a fully functioning device chronically implanted in a human. Existing devices have been tested in animals, and successful tests may create a push to implant in humans. However, testing must be sufficient to ensure the safety of the patient [55].

2.3 The State of Electrical Neural Recording Devices

Partially implantable wireless BMI devices have been used to successfully record neural signals in animal experiments [41], [10], [84], [83]. In [1], a fully implantable device has been designed and its functionality is verified in bench testing. In [79], an implantable device was tested in a rat. The device is not currently wireless, but rather uses a fiber-optic system to transmit recorded data to the implant. They state that in the future data could be transmitted through the skin with a near-infrared laser. Perhaps the most advanced fully implantable device is presented in [11]. The hermetically sealed device connects to a 100-channel MEA, and wirelessly transmits the 24Mbps of recorded data to a receiver 1 meter away. The system is battery powered and can run continuously for 7 hours between recharge cycles (through an inductive link). The device is 5.6 x 4.2 x 1.1 cm^3. The device was tested in a non-human primate and was shown to operate safely and effectively for over 1 year.

Despite these achievements, there are still several milestones to achieve before these devices can truly become clinically relevant. Remaining issues for chronic fully implantable devices include the physical size and tissue damage, power dissipation being small enough for long term recording and supplying power to the devices, increasing the number of channels and long term signal quality and reliability, and biocompatibility [35]. New techniques and novel design work are required to address these current limitations.

2.4 Electrical vs. Optical Recording

Conventional BMI technologies focusing on MUR or LFP recordings interface electrically with the nervous system through miniature MEAs. An example of such an array is the Utah Electrode Array, which is a 10 x 10 array of platinum-tipped silicon microelectrodes. The MEA is 4 x 4 x 1.5mm^3 [65]. MEAs can be connected to a microchip with recording circuitry, and then implanted. Unfortunately, recording through an electrical interface with MEAs has several drawbacks, from both the biological and circuit design points of view, that make the implementation of a fully implantable recording interface a difficult task.

Implanting an MEA damages tissue. As discussed previously, the response of the body is to encapsulate the micro-electrodes, which degrades the fidelity of recording and makes chronic recording difficult. The size of the micro-electrodes is limited by the amount of tissue damage that is deemed acceptable. Limited micro-electrode size and spacing places limits on the resolution of recording. A micro-electrode cannot necessarily target and record the membrane potential from only individual cells; often the recorded signal is an MUR or LFP. Such recordings are useful for certain applications, but not necessarily useful for others where single-unit recordings are required.

From a circuit design point of view there are several issues that must be overcome. Charge balance in the tissue must be maintained so as not to upset the electrochemistry of the tissue. Neural signals occupy a low frequency band, which means filtering circuitry requires very large RC time constants. Large on-chip capacitors and resistors require significant chip area. Large (10s of MΩ and larger) resistances are typically implemented with transistors operating in the subthreshold regime where process variation reduces precision, which then requires extra circuitry for programming/tuning the correct resistances. Stimulation is also typically performed electrically, and stimulation artifacts affect recorded waveforms [42].

Given these issues, there is motivation to find other ways to implement neural recording. Replacing electrical interfacing with the nervous system with optical interfacing is a possible option. Optical recording and stimulation have the following potential advantages over electrical recording and stimulation:

- Reduced tissue damage compared to an MEA, assuming the implant sits on the surface of tissue rather than puncturing the tissue.
- Optical recording may not degrade over time, and may be a better solution for chronic implants.
- Higher spatial resolution is possible (more photodetectors or light sources per unit area than electrodes).
- Light does not interfere with neuronal function [85].
- Specific cell types can be targeted (using optogenetic techniques).
- Electrical frequency response requiring large time constants may be avoided.

Due to these potential advantages, optical interfacing is becoming an area of increasing interest to the research community. Several groups have designed implantable devices with optical stimulation [74], [97], some with simultaneous electrical recording [100], [96]. However, few groups have designed devices for implantable optical neural recording. The basis for an optical recording system is fluorescence detection, which is described in the next section. The state of the art in implantable optical neural recording devices is then reviewed in Section 2.6.

2.5 Neural Recording With Fluorescence Detection

In a *fluorescence detection (FD)* system the goal is to obtain an image of some system of interest by detecting fluorescent light emitted from the system. The system of interest must contain fluorescent molecules, which contain *fluorophores*. A fluorophore is a part of the molecule that absorbs light of a certain energy (or energies) and then in response emits light of a certain lower energy. The most basic FD system includes an excitation light source for the fluorophores, and a photodetector to record the light emitted by the fluorophores. A photodetector captures photons that are emitted from the fluorophore and produces an electrical signal proportional to the amount of captured photons. The signal can then be processed to extract relevant information about the system being imaged.

Most conventional FD systems are rather large, and include several optical components, such as filters, mirrors and lenses, as well as a lamp or laser excitation source and some sort of camera as the photodetector. An example of a bench-top FD system for in vivo imaging is seen in [85]. Their particular system includes a mercury arc lamp connected through a fiber to a structure that includes excitation and emission filters, a dichroic mirror and a projection lens. The dichroic mirror allows detected light to reach an EMCCD (electron-multiplying charge-coupled device) camera. The structure is connected through a fiber optic cable to a small headstage on a rat. The system allows for detection within a somewhat freely moving mouse, though the fiber optic cable tethering still restricts movement.

2.5.1 Fluorophore Characteristics

A given fluorophore has a characteristic *excitation (absorption) spectrum* and *emission spectrum*, which define the energies (or wavelengths) of photons that it will absorb and emit. Fig. 2.2 shows the excitation and emission spectra for Cy3, a common fluorophore [18]. The wavelength of emitted light is longer than the wavelength of absorbed light (or the energy of the emitted light is lower than the energy of the absorbed light). This shift in energies between absorbed and emitted light is known as the *Stokes shift*.

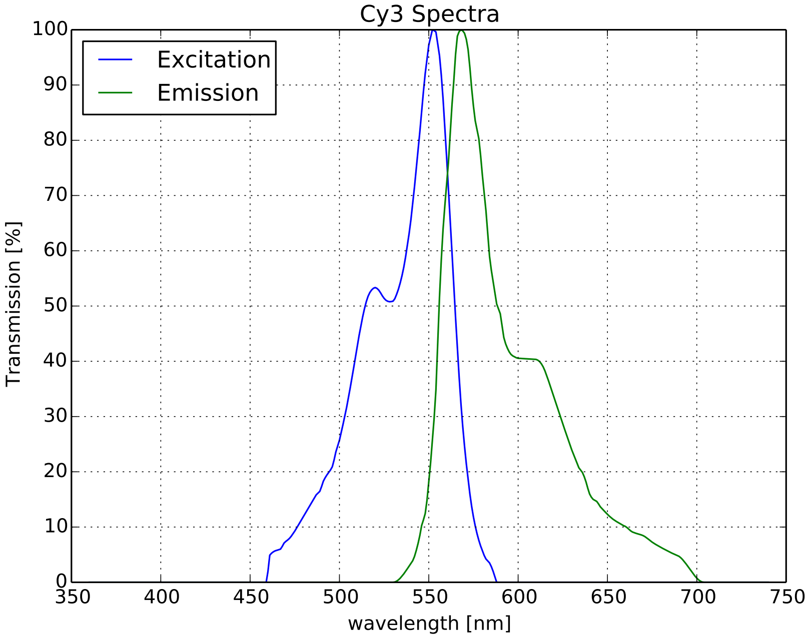


Figure 2.2: Excitation and emission spectra of Cy3.

Different types of fluorophores are used for different applications. Many FD systems are designed to produce a static image of some biological system, and therefore use a relatively long exposure time when recording the image. Such a system might be used simply to determine if a certain type of cell exists in the system of interest. In this case the sensitivity of the fluorophore to some biological signal is not very important, nor is its response time. However, in other FD systems, such as those used in neural recording, the sensitivity of the fluorophore to the signal of interest and the speed with which it can respond are very important characteristics.

2.5.2 Types of Fluorophores

The two main types of fluorophores are *fluorescent dyes* and *fluorescent proteins*. Two common types of dyes for imaging neural activity are *voltage-sensitive dyes (VSDs)* and *calcium-sensitive dyes (CSDs)*. A VSD changes its intensity of fluorescence in response to a change in membrane voltage of a neuron. A CSD changes its intensity of fluorescence in response to a change in calcium ion concentration near the fluorophore. This *change in intensity* is the signal of interest. The advantage of dyes in neural recording is that they can have good sensitivity and are very fast to respond to neural signals, and support high temporal resolution. The disadvantage of dyes is that they stain all of the cells in the area where the dye is injected and do not target specific cell types. This means that the FD system will collect a lot of undesired light, which is often referred to as *noise* or *background light*. Dyes can also be toxic and can not be used for chronic imaging.

Fluorescent proteins are similar to dyes except that they can be genetically programmed. The genes encoding a protein can be introduced into an organism, which then can produce the protein. Through appropriate genetic coding, fluorescent reporter proteins can be targeted to specific cell types, and even to certain locations in cells [85]. Fluorescent reporter proteins can be designed to sense voltage or calcium signals, or other signals that may be of interest. The capability to design such light-sensitive proteins has brought rise to the field of optogenetics, which combines genetics with optical systems for the purpose of monitoring or controlling cells in vivo [101].

Voltage-sensitive fluorescent proteins (VSFPs) are proteins that sit inside the plasma membrane of neurons. In response to a changing membrane potential, VSFPs alter their optical characteristics; usually the intensity of fluorescence [90]. Early generations of VSFPs suffer from low SNR, low response time, or poor targeting to the plasma membrane [6]. However, development of VSFPs is occuring at a rapid pace [85]. VSFPs are much more promising than VSDs for chronic recording.

Similar to VSFPs are genetically encoded calcium indicators (GECIs). In neurons, "action potentials trigger large and rapid changes in cytoplasmic-free calcium concentration" [95]. Instead of directly sensing the neuron membrane potential, GECIs respond to the concentration of calcium ions. GECIs are commonly used to image neural activity [95].

2.5.3 Comparison of VSFPs

There are several important characteristics of VSFPs that must be taken into consideration when choosing an appropriate VSFP for an application:

- **Targeting Specificity:** This is how well the VSFP targets the correct cell type/location. Attachment to the wrong location contributes to noise in the recorded signal [6].
- **Capacitive Loading:** Adding molecules to the cell membrane contributes to extra capacitive loading of the neuron. High levels of expression can increase the capacitive loading enough to alter the functionality of the neuron [56]. However, as shown in [2], sufficient

SNRs could be achieved for VSFP2.3 at concentrations that did not affect action potential generation in response to applied synaptic stimulation.

- **Photostability:** This is a measure of how well the VSFP maintains performance when exposed to excitation light for a long period of time.
- **Wavelengths:** The absorption and emission wavelengths of the VSFP should be sufficiently far apart such that excitation light can be separated (filtered out) from the signal of interest. Higher wavelength light can penetrate deeper into tissue. Red-shifted spectra overlap less with hemoglobin absorption and tissue *autofluorescence* [6]. Autofluorescence is due to light emitted by naturally occuring fluorescent molecules. For example, NADH and flavins, which are products of metabolism, produce strong autofluorescence [64].
- **Sensitivity:** This is how sensitive the VSFP signal (change in fluorescence intensity) is to the voltage signal. This is often expressed as a percent change in fluorescence intensity per 100mV change in membrane potential. It is argued in [54] that this metric of sensitivity is inaccurate, but nevertheless it is widely used. The sensitivity of the VSFP directly impacts the required power of the excitation source.
- **Speed of Response:** This is how fast the VSFP can respond to a change in membrane voltage. The response time should be faster than the desired sampling rate.

Fig. 5 from [54] shows a comparison of the speed and sensitivity of several VSFPs from the literature. Most VSFPs are not fast enough to allow for the recording of individual action potentials, and have fairly low sensitivity (typically less than 10%/100mV). However, recently Kralj et. al. [54] reported a new VSFP called *Archaerhodopsin 3 (Arch)* which greatly exceeds the performance of existing VSFPs, both in terms of speed and sensitivity. They could not accurately measure the response speed of Arch because of the limitations of their equipment, but it was faster than 500μs. The sensitivity is around 65%/100mV, while the phototoxicity is comparable to other VSFPs. The quantum yield is quite low, but high sensitivity somewhat makes up for that. The absorption and emission wavelengths are relatively high, where tissue is relatively transparent and where corruption due to autofluorescence is low.

The authors were able to resolve individual spikes in mammalian neurons in vitro with high SNR and submillisecond temporal resolution, as shown in Fig. 2 of [54]. These results are very promising, and should make it possible to design a system to perform high temporal resolution neural recording based on VSFPs. The authors suggest that improvements on Arch may be found within the family of about 5000 known microbial rhodopsins.

2.6 The State of Integrated Optical Neural Recording Devices

Numerous integrated optical detection systems have been developed for lab-on-a-chip type applications [34], and contact imaging applications [88], [9], [89], but only a handful of groups have looked at developing integrated optical neural recording devices.

One notable work is from Schnitzer's group [33]. They presented a miniaturized fluorescence microscope for brain imaging that is head-mountable on a rodent, though tethered. The device weighs 1.9g and is made entirely from mass-producable parts. The device was used to image Ca^{2+} spiking over a $0.5mm^2$ area of a mouse brain at up to 100Hz. The imager is 300 x 300 pixels, and the imaging resolution is about $2.5\mu m$.

Coskun et. al. designed a lensless wide field-of-view fluorescence microscope [20]. Compressed sensing/sampling algorithms are used to improve the spatial resolution of the imaging. Long integration times around 1 second are used. Instead of optical filters, the device includes a large prism to reject excitation light through total internal reflection. The prism is 25mm x 35mm x 17mm, which is too large for an implant.

A few groups have designed integrated partially implantable devices for optical neural recording. Several papers by one group describe the evolution of their implantable optical neural recording devices [76], [78], [77], [72]. The latest device can be mounted on the head of a rodent, though it is quite a large system. Their design specifications state that "the system must contain all the elements of a traditional fluorescent microscope: illumination source, collector lens, dichroic excitation filter, dichroic mirror, objective lens, emission filter, transfer lens and image detector." They use a high power LED, driven at up to 750mA. It includes a 32 x 32 pixel image sensor and images at 900fps. It can sense a 0.2% change in fluorescence. The pixels have a well capacity of 2.1Me^{-}, a peak SNR of 61dB, and a conversion gain of 4.9$\mu V/e^{-}$.

Another group has several papers presenting their work [67], [62], [68], [69]. Their earlier work is geared towards a head-mounted imaging device for a mouse. The head-mounted device includes lenses, a CMOS imager and an excitation LED. The device in [68] includes a CMOS imager with one LED on opposite sides on a flexible substrate. The imager is 90 x 30 pixels, with 7.5μm x 7.5μm 3T (three transistor) active pixel sensor (APS) pixels, designed in the ams 0.35μm 2P3M (two poly three metal) technology. The pixel fill factor (FF) (the percent of the total pixel area allocated to photodetector area) is 35%. The imager is coated with an excitation filter (through a spin coating process). They state that, "an interference filter, which is conventionally equipped in a fluorescence microscope, is not applicable to our sensor, because fluorescence is so strongly scattered in the tissues that it enters into the sensor from every direction." Note that an interference filter is most useful in cases when the light to be filtered is incident perpendicular to the filter surface. The implantable portion of the device is about 0.5mm wide and 0.2mm thick, which is smaller than the previous device. As part of their testing they placed fluorescent beads ($15\mu m$) on the surface of the device to measure the spatial resolution and excitation light distribution of the device. $7\mu W$ LED optical power (from $50\mu A$) was used in this test. They mention a new device design where the LED dies are buried into the image sensor substrate after a deep reactive ion etching (DRIE) step to hollow out a place for them. This allows the LED surfaces to line up with the imager surface.

They have also developed an implantable device for imaging in deeper tissue [69]. This device includes no imaging optics and utilizes proximity-based imaging. The device is packaged with with LEDs and an excitation filter on a flexible polyimide substrate, and is designed to be implanted much like a microelectrode or needle. It is encapsulated with Parylene-C.

Murari et. al. present a high-sensitivity CMOS fluorescence imager in [60]. The imager is 132 x 124 pixels with $20.1\mu m$ pixel pitch, fabricated in a standard $0.5\mu m$ technology. The pixels include CTIA-based amplification. The chip includes delta differencing circuits to "calculate the difference between a pixel's light-dependent signal value and the subsequent reset value as a measure of the photon flux". With the nominal frame rate (FR) of 70fps, the sensitivity is about 4nW/cm^2. The imager is used in a head-mounted fluorescence microscope system in [87]. The laser speckle contrast imaging system is used to visualize micro-vessels and monitor blood flow in a rat brain.

An interesting device is presented in [53]. The device is a miniature injectable device capable of optical stimulation through 4 addressable microscale LEDs. Adjacent to the $50\mu m$ x $50\mu m$ micro-LEDs are a $20\mu m$ x $20\mu m$ Pt microelectrode for electrical stimulation or recording, a $200\mu m$ x $200\mu m$ photodiode, and a precision temperature sensor or microheater. The important advantage of the device is that it can be implanted deep in tissue. The device must still attach to wires for power and control, and is thus not fully implanted. The external components are, however, wireless, allowing for freely moving rodents.

A more detailed comparison of image sensors from the literature described above is provided in Section 6.4.7.

2.7 Summary

This chapter described the signals of interest for neural recording, and the constraints involved in the design of implantable neural recording devices. The potential for optical recording devices based on fluorescence detection to improve on conventional electrical recording devices was discussed. The existing literature for both electrical and optical recording devices was reviewed.

While good progress has been made in wireless implantable electrical recording devices, optical recording devices are considerably further behind. A key parameter in optical recording devices that must be improved is the frame rate, which should be fast enough to detect individual spikes. If optical recording devices are ever to be fully implanted, the level of integration must also be improved; there is plenty of room for optimization of device design left. Wireless operation, power delivery and heat dissipation are also important factors that must be considered.

In this work the goal is to design a prototype implantable optical recording device with sufficient frame rate to detect individual spikes while eliminating as much of the optics as possible and relying on imaging from the surface of tissue in order to avoid any serious damage due to puncturing of the tissue. As will be seen in chapters 5 and 6, the target imager specifications used in this work are as shown in Table 2.1. The specifications are derived from the modeling results of chapter 5. The sampling rate (frame rate) is relaxed from the Nyquist rate as the goal is to detect spikes rather than record their shapes.

Pixel pitch	$100\mu m$
Frame rate	2.5kHz
Resolution	1mV

Table 2.1: Optical recording device specifications for this work.

The remainder of this thesis presents the work done towards electrical and optical recording front ends.

Chapter 3

Electrical Recording Chips

This Ph.D. work began with the goal of designing multi-channel front-end electrical neural recording circuitry to interface with a microelectrode array in an implantable wireless neural recording device. This chapter describes the work that was done for electrical neural recording circuitry that was to be integrated on a system-on-chip with wireless communications and powering, as well as any other circuitry required for a fully implantable device. Ideas for future work on the recording circuitry are discussed.

3.1 AF4 Design

In collaboration with colleagues from the Alberta Innovates - Health Solutions (AIHS) Sensory Motor Adaptive Rehabilitation Technology (SMART) multidisciplinary team in neural prostheses, we arrived at specifications for the front-end circuitry that would interface with a microelectrode array:

- 16 channels
- amplification of neural signals as large as $200\mu V$ (peak)
- 750Hz to 7.5kHz bandpass frequency response
- 15kSps sampling rate (per channel)
- 8-bit resolution
- low power and low area

The key circuit for the front-end circuitry described here is the low-noise amplifier (LNA), which connects to the microelectrodes ($\sim 200k\Omega$ impedance). The LNA must have the following characteristics:

- very low input-referred noise so as not to corrupt the small and noisy neural signals.
- DC-coupling at the input to reject potentially large DC offsets
- bandpass filter characteristic, with high-pass cutoff at 750Hz and low-pass cutoff at 7.5kHz

It was also decided that the LNA should be fully differential in order to have very good common-mode and power-supply rejection ratios, given that the power supply of an implanted device may be noisy.

Fig. 3.1 shows the chosen circuit schematic for the LNA, including common-mode feedback (CMFB) circuitry. The telescopic amplifier was chosen for its combination of low power consumption (only one tail current branch) and high gain. The limited swing of the telescopic amplifier is not an issue in this case because the output signal is on the order of 10s of mV at the LNA output (a second gain stage is used to provide the rest of the gain). Table 3.1 shows the LNA device sizing.

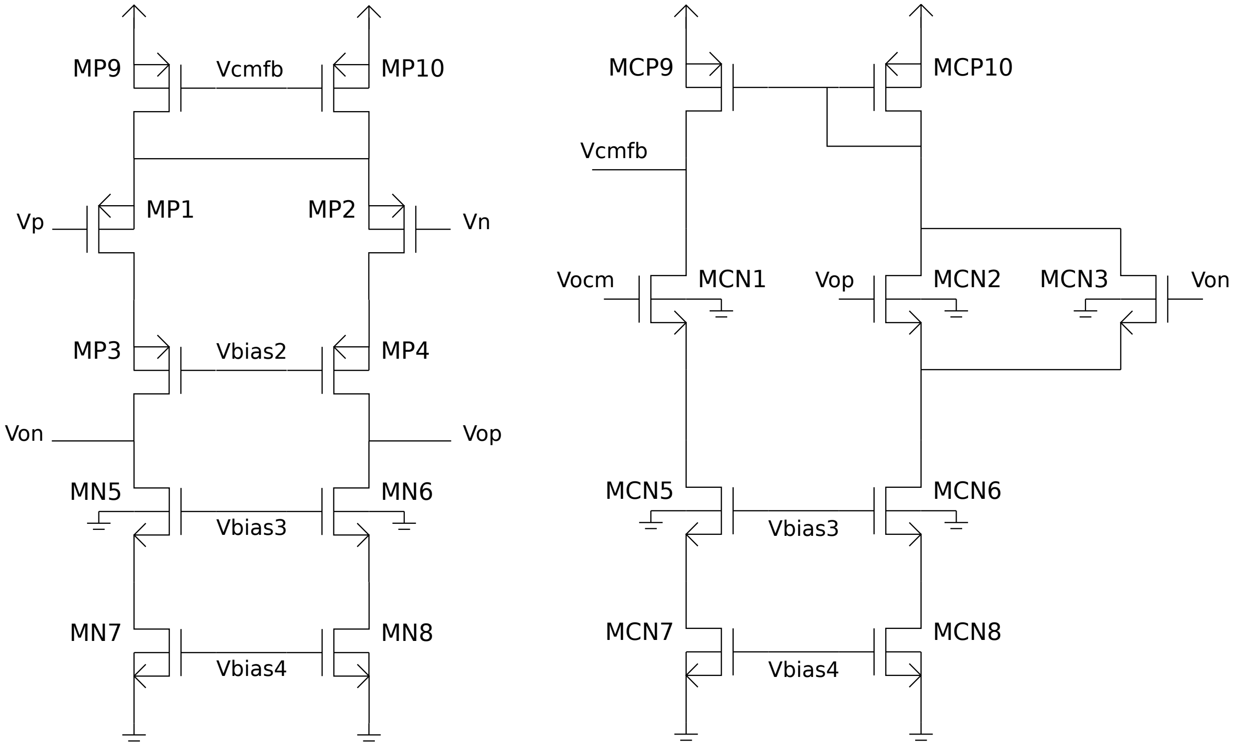


Figure 3.1: LNA schematic.

	MP1	MP3	MP5	MP7	MP9	MCN1	MCN3	MCN5	MCN7	MCP9
W	40	8	2	4	8	4*2	4	2	4	8
L	20	10	10	20	10	10	10	10	20	10

Table 3.1: LNA device sizing (all sizes in μm).

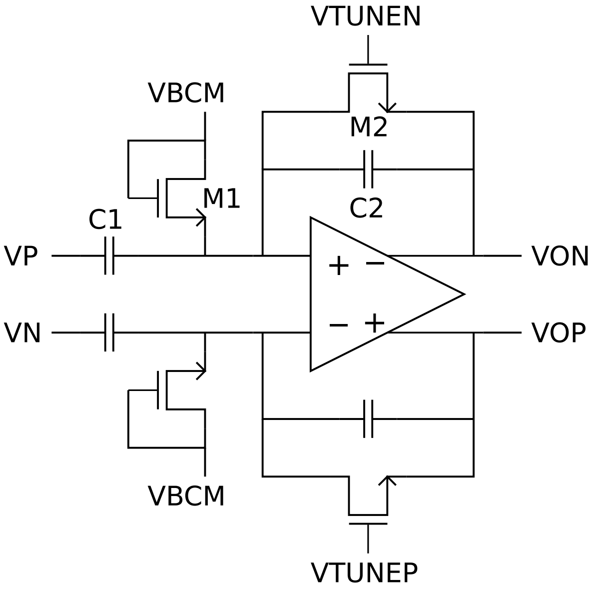


Figure 3.2: LNA feedback schematic. The transistors operate in the subthreshold region and implement the large resistance values. The M2 devices correspond to R_2.

Fig. 3.2 shows the LNA feedback circuitry, which implements a high-pass filter, as described by Eq. 3.1.

$$A(s) \approx \frac{-sC_1R_2}{1 + s\left(\frac{C_1R_2}{1+A} + C_2R_2\right)} \; . \tag{3.1}$$

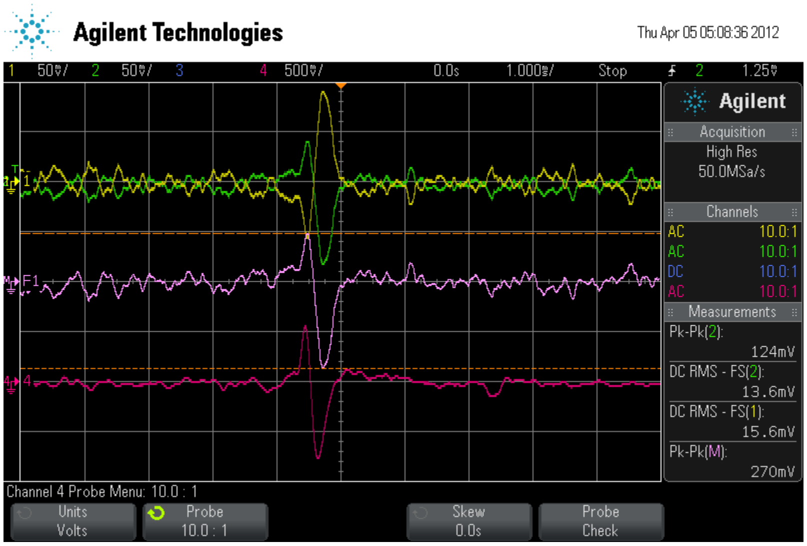
The mid-band gain is set by the capacitor ratio C_1/C_2. The high-pass cutoff frequency is set by the C_1R_2 time constant. A main issue in the design of this circuit is the very large resistance value required for the high-pass filter time constant. The input capacitors C_1 are chosen to be 10pF. Their size is limited by circuit area. This means the required feedback resistance is:

$$R_2 = \frac{1}{2\pi f_L C_1} = \frac{1}{2\pi * 750 * 10x10^{-12}} \approx 21.2M\Omega \; . \tag{3.2}$$

This is too large to implement with a typical polysilicon resistor, which would require significant area. Instead, the resistance is implemented with a MOSFET operating in the subthreshold region, where the channel resistance can be made very large. The problem with using a MOSFET in subthreshold is that process variation has a significant effect on the device's channel resistance. Simulation shows that the filter cutoff frequency can vary by about 4 orders of magnitude. This means that the device bias voltage must be tuned post-fabrication in order to obtain the desired cutoff frequency. This tuning requirement could be thought of as an advantage, as different high-pass cutoff frequencies could be used to record neural signals other than action potentials. However, when a precise cutoff frequency is desired, this variability due to process variation is a problem. Section 3.5 discusses some potential solutions to this problem. Another issue with the MOSFET device in feedback is linearity, in the case where the output signals are larger than a few 10s of mV, which is the case in the second stage amplifier.

3.2 AF4 Results

The LNA, along with a second stage amplifier, was fabricated in a $0.35\mu m$ CMOS technology and bench tested. Figs. 3.3 and 3.4 show oscilloscope captures of the channel outputs (green and yellow are positive and negative) with a pre-recorded neural signal used as the input (the dark pink at the bottom). The light purple waveform in the middle is the difference between the differential outputs. Note that the input signal scale shows the signal as much larger than it actually is. The signal from the waveform generator is divided down to typical microvolt neural signal levels before being fed to the LNA.



High Res
50.0MSa/s

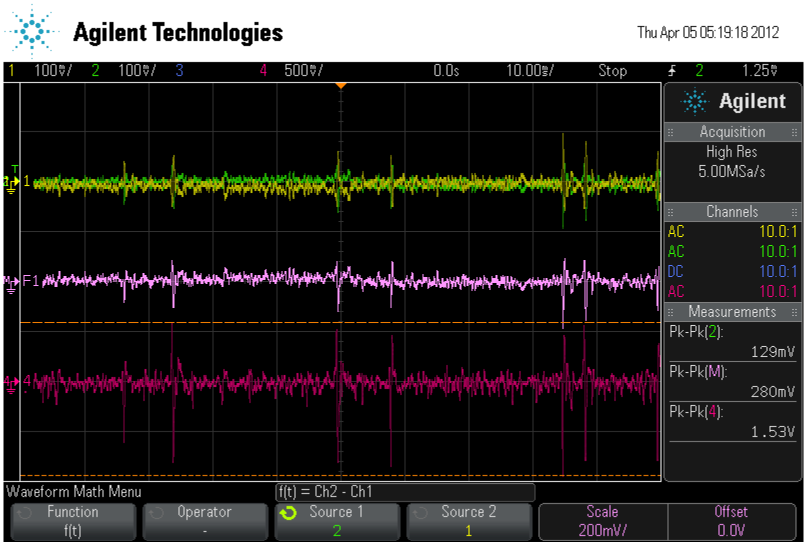
AC 10.0:1
AC 10.0:1
DC 10.0:1
AC 10.0:1

Pk-Pk(2): 124mV
DC RMS - FS(2): 13.6mV
DC RMS - FS(1): 15.6mV
Pk-Pk(M): 270mV

Channel 4 Probe Menu: 10.0 : 1

Units Volts
Probe 10.0 : 1
Skew 0.0s

Figure 3.3: Analog channel output (LNA and second stage), single spike.



High Res
5.00MSa/s

AC 10.0:1
AC 10.0:1
DC 10.0:1
AC 10.0:1

Pk-Pk(2): 129mV
Pk-Pk(M): 280mV
Pk-Pk(4): 1.53V

Waveform Math Menu
f(t) = Ch2 - Ch1

Function f(t)
Operator -
Source 1 2
Source 2 1
Scale 200mV/
Offset 0.0V

Figure 3.4: Analog channel output (LNA and second stage), multiple spikes.

3.3 AF5 Design

After verifying the functionality of the channel design, a multi-channel chip, internally referred to as AF5, was designed. The chip has 4 blocks of 4 channels. The channels in each block had slightly different feedback circuitry - some used NMOS devices in feedback, some used PMOS, some were single-ended output and some were fully differential. Within each block, each channel output is wired to a multiplexer before going off chip. Also included on the chip was a 5-bit resistor string digital-to-analog converter (DAC) for controlling each feedback device tuning voltage, as well as the CMFB voltages. The DACs are programmed through a serial shift register. This chip is to be used in live recording experiments.

3.4 ADC

An 8-bit 1MSps successive approximation register (SAR) analog-to-digital converter (ADC) was designed to be used with the analog front-end. The design was not included in the AF5 chip, but was included in the OR2 chip discussed in Chapter 6. The ADC design and test results are discussed in Appendix B.

3.5 Future Work

Regarding electrical recording, there are a few directions that can be investigated for improving the analog front-end circuitry.

3.5.1 Auto-Calibration

In order to solve the problem of process variation affecting the high-pass cutoff frequency, auto-calibration of the bias voltages could be implemented. The idea is to use a DAC to generate a sine wave as an input to the channel, and use the (already existing) ADC to find the gain at an input frequency in the tunable range, and compare it to the desired value. Digital logic would control the tuning voltage through a DAC, and implement the control loop. Such an auto-calibration scheme could also be used to compensate for any drift in the tuning over time if it was run periodically.

Reducing the channel electronics to a single amplification stage is feasible, and would greatly simplify any calibration routine, as it eliminates combinations of variables that need to be optimized.

Originally, small resistor-string DACs were used to control the tuning voltages. A much more compact and potentially lower power solution may be to use a dynamic switching DAC that periodically charges a capacitance and relies on slow leakage to maintain the voltage. Such a DAC could be multiplexed between several different outputs.

3.5.2 Switched-Capacitor Operation

Using a switched-capacitor resistor in place of the subthreshold MOSFET for the large resistance value is another possible option. In this case the resistance is determined by a capacitor value, and a clock frequency, which is well controlled and can be programmed. A key factor in such a design would be the kT/C noise contributed by the switched capacitor.

3.5.3 Other Improvements

A few other improvements would be including the ability to power-down individual channels, and including a reset switch across the feedback resistance to provide a quick way to bias the amplifiers and to avoid any drift in DC levels.

3.6 Summary

The experience with the electrical recording circuitry design uncovered the key design challenges that must be tackled in future work, and also led to the exploration of alternative methods for neural recording. With a new M.Sc. student to take over the electrical recording work under my guidance, my focus shifted to optical neural recording, which is the focus of the remaining chapters of this thesis.

Chapter 4

Basic Optical Model and OR1 Chip

Given the potential benefits of optical recording and promising results from VSFP testing discussed in Chapter 2, I decided that the feasibility of an implantable optical recording system should be explored through optical modeling of an envisioned system. Modeling is an important first step for two main reasons: 1) it is very costly in terms of time and effort to design, fabricate, and test a microelectronic device; modeling is much quicker and modeling results can indicate whether or not undertaking the design of a device is worth the time and effort, and 2) having a parameterized model allows for optimization of system parameters and derivation of appropriate design specifications.

One of the main goals of this work is to shrink the size of existing optical systems down to something that could be fully implanted. In order to do that, the idea is to eliminate all of the optical components (other than excitation filter) and rely on contact (proximity-based) imaging. Note that eliminating lenses from the design will mean that the images are out of focus due to scattering of the light we are trying to image. Thus, the source of fluorescent light must be close to the image sensor, which means the useful imaging depth is limited. The initial idea was to bond LED dies directly on top of a CMOS imager chip.

The rest of this chapter first describes the aforementioned system, then proceeds to describe a basic general model to predict where light from an excitation source will end up in a system, and in particular applies the model to the system described here. Next, the design of the first optical chip, referred to as OR1, is presented. Lastly, experimental measurements from the OR1 chip are compared with the modeling results. The following chapter extends the optical model presented here to an implantable neural recording device.

4.1 OR1 System Design and Basic Optical Model

Fig. 4.1 shows a drawing of the initial OR1 device design, which guided the optical modeling. The design has the following characteristics:

- A pixel array in the center of the chip
- LED dies bonded to large bondpads on the chip, adjacent to the pixel array
- Nothing blocking the line of sight between the p-n junctions of the LED dies and the pixel array (i.e. free space)

To model the system we must consider how light is emitted from an LED. An LED radiation pattern is typically specified from a far-field point of view, where the LED can be modeled as a point source within the larger system. The radiation pattern of a packaged LED can be assumed

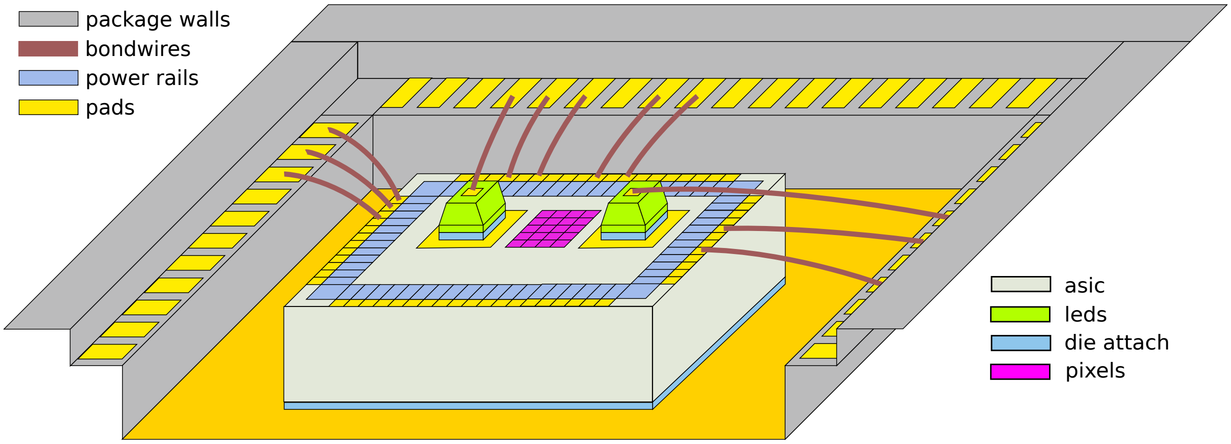


Figure 4.1: OR1 device design.

to be either *Lambertian*, or *isotropic* (but with light only emitted in the upwards direction, as any light emitted downwards from the junction would be reflected back up by the package). Light from a Lambertian source is not equally distributed in the beam angle, but rather most of the light is emitted directly upward, and less is emitted as the angle from the normal increases. The packaging of the die can affect the radiation pattern. However, in this case we are using bare LED dies, not packaged LED dies.

For an isotropically radiating point source of light (in free space), the optical intensity decreases with radial distance from the source due to spreading according to

$$I(r) = \frac{P_0}{4\pi R^2} \quad \left[\frac{W}{m^2}\right], \tag{4.1}$$

where P_0 is the total power emitted by the source, and R is the radial distance from the source [37]. In the case where the source emits not isotropically, but rather in a given solid angle ω, the previous equation can be modified as follows:

$$I(r) = \frac{P_0}{\omega R^2} \quad \left[\frac{W}{m^2}\right]. \tag{4.2}$$

The total point source power P_0 can be found from the radiant intensity, I_e, in Watts per steradian, where the steradian is the unit of solid angle [91], and from the solid angle of emission, ω, as follows:

$$P_0 = I_e\omega \quad \text{[W]}. \tag{4.3}$$

Though the LED dies are above the imager array and mainly shining away from the array, we cannot assume that no light hits the array. One might assume that because the LED die is bonded to a large metal bondpad, any light emitted downward would just be reflected upwards, and no light should hit the pixel array. However, the p-n junctions of the LED dies are at a non-zero height above the LED die bottom surfaces. Photons can be emitted downwards and exit the side of the die without reflecting back upwards, and can actually hit the pixels below the dies, hence contributing to an undesired baseline excitation signal. Therefore, the main goal of the basic model is to predict how much baseline excitation power from the LED die will directly hit the pixel array, and how the absorbed light signal changes as a function of distance between the pixel and the LED die.

Because the system geometry is comparable to the LED die dimensions, a far-field model of the LED die emission is not applicable. Thus, here the LED die is modeled as a set of distributed point sources in a plane at the location of the p-n junction in the die. The total power of the LED is divided equally into each of the distributed point sources. By using tiny distributed point sources, the distance between each point source and each pixel in the array is large enough to use the basic far-field equation for isotropic radiation. The contribution of each point source to each pixel is summed to determine the total power absorbed in that pixel. The model neglects any potential reflections in the system, assumes there is a direct line of sight from each point source to each pixel, and assumes no photon reabsorption or scattering inside the LED die.

The power absorbed in a pixel is then converted to the change in voltage during the integration time according to the equations for the imaging circuitry. The 3T APS circuit shown in Fig. 4.2 is a well known basic pixel circuit, and is used in OR1. By inspection, the equation for this circuit is:

$$\Delta V_{px} = P_{px} Resp \frac{T_{int}}{C_p} A_v \ , \tag{4.4}$$

where P_{px} is the power absorbed in the pixel [W], $Resp$ is the responsivity of the pixel [A/W], T_{int} is the integration time [s], C_p is the pixel capacitance integrating the photocurrent [F], and A_v is the additional voltage gain of the readout circuitry [V/V].

Fig. 4.3 shows example model results for the OR1 model. The black square represents the LED junction. The colors represent the change in voltage in each pixel.

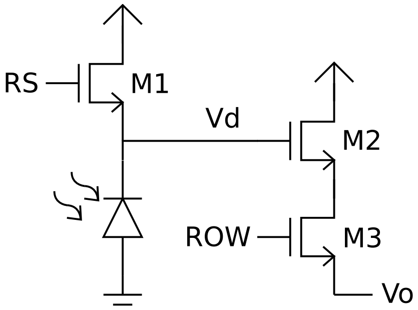


Figure 4.2: The basic 3T APS circuit.

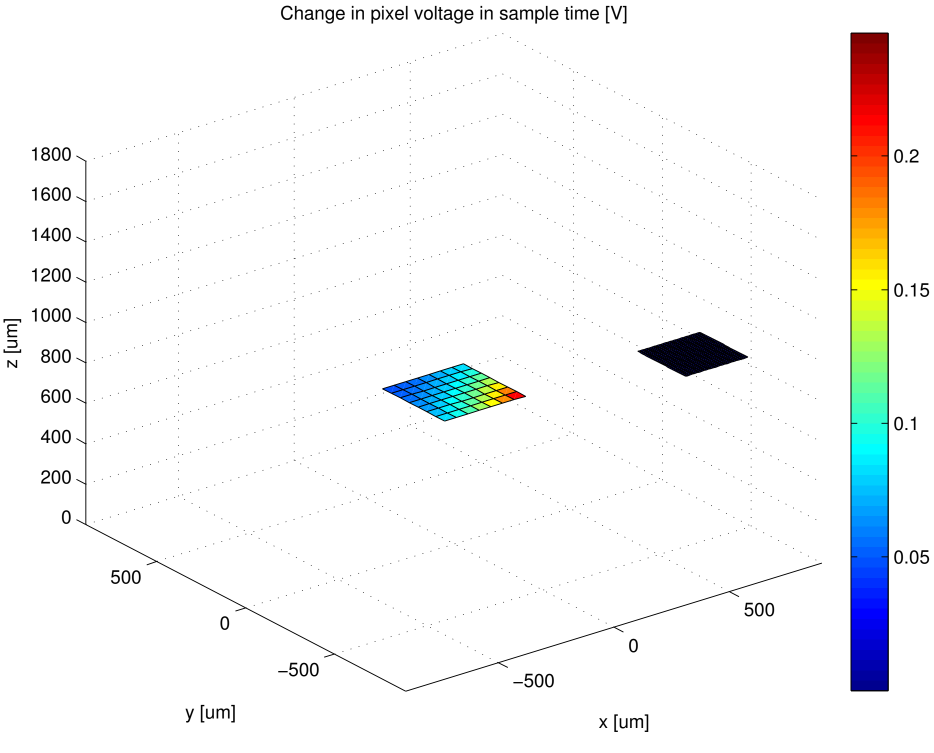


Figure 4.3: Plot of typical OR1 model results. The black square represents the LED junction. The colors represent the change in voltage in each pixel.

4.2 OR1 Chip Design

A test chip, OR1, was designed to verify the optical modeling presented above. For OR1 the concern was not with imaging fluorescence or optimizing the imaging circuits. The goals for the OR1 design were to:

- Determine the baseline light signal hitting the imager as a function of distance between an LED die and the pixel
- Verify the mixed-signal design flow for the *austriamicrosystems* 0.35μm (amsp35) technology
- Verify the basic imager architecture and digital control logic
- Test the bonding of LED dies to large pads on a CMOS chip, and the simple LED driver circuits

In summary, the chip includes an 8 x 8 array of 50μm x 50μm 3T pixels, programmable imager control and readout logic, and four 500μm x 500μm bondpads adjacent to the pixel array, with an LED die bonded to each. Fig. 4.4 shows the die photo of the chip with the LED dies bonded. The pixel array rows and columns are labeled. Pixel 1 is row 1, col 1. Pixel 2 is row 1, col 2. Pixel 9 is row 2, col 1, etc. LED2 and LED4 have single-transistor on-chip drivers whose gate voltages are set with off-chip DACs, while LED1 and LED3 are driven through off-chip circuitry. Fig. 4.5 shows a picture of OR1 with a side perspective. The design details for the chip are discussed in the following sections.

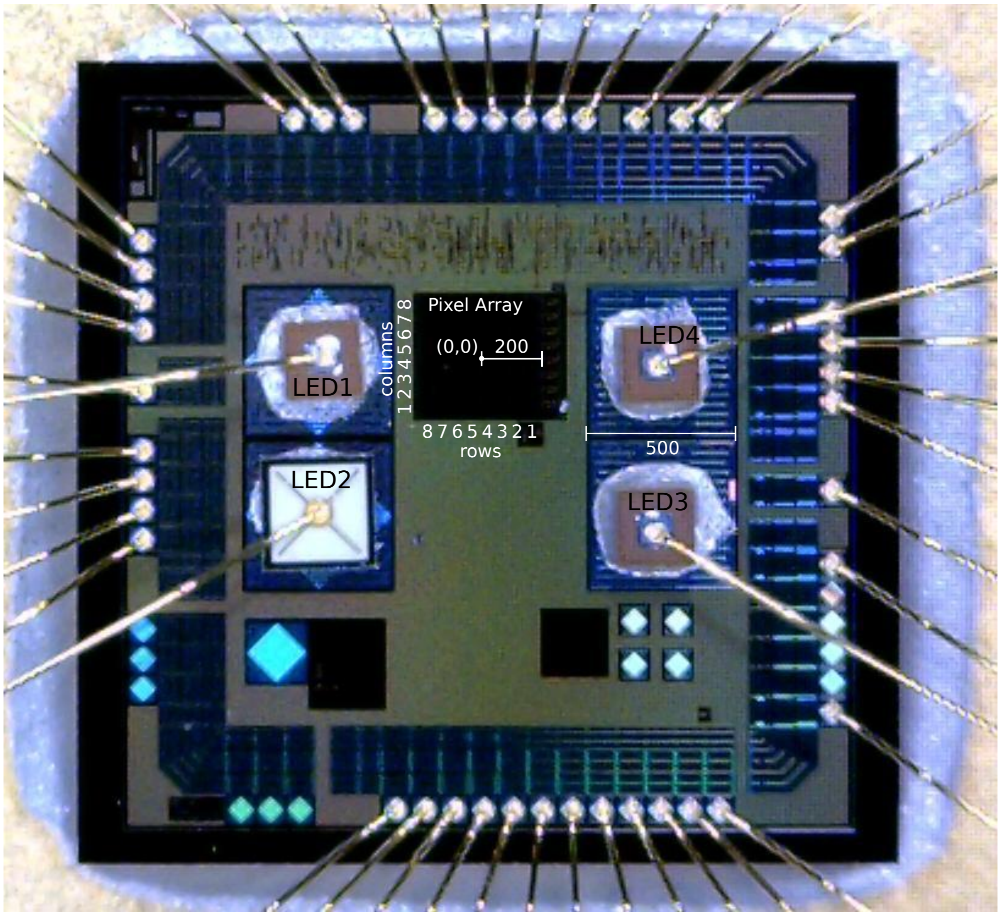


Figure 4.4: Die photo of ICMWTOR1 with LED dies bonded.

4.2.1 Photodiodes

The authors of [59] did a study comparing the performance of different types of photodiodes and determined that N-well p-substrate (NW-PSUB) photodiodes were the best choice. NW-PSUB diodes are used in this work.

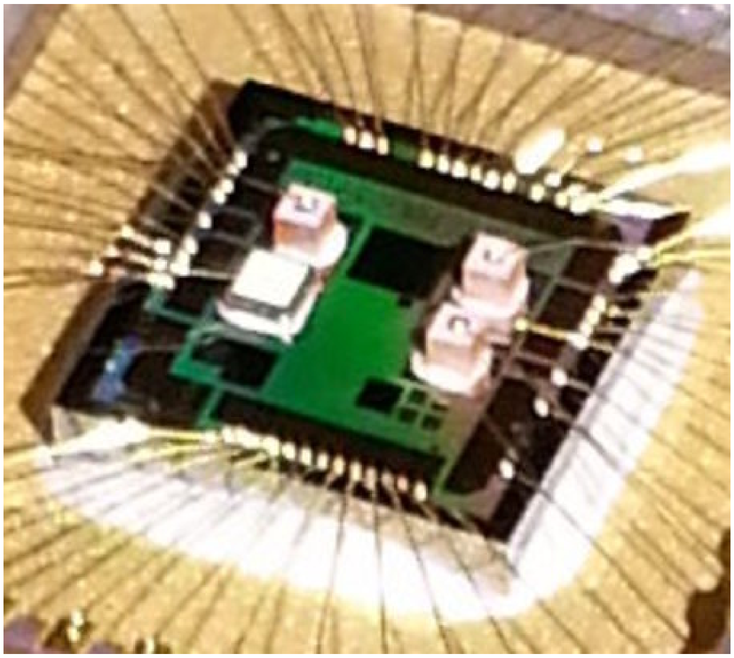


Figure 4.5: Side view of ICMWTOR1 with LED dies bonded.

4.2.2 Imaging Circuitry

For this first chip a basic 3T APS pixel circuit was designed, as shown earlier in Fig. 4.2. RS is pulsed high to reset the pixel (setting $V_d = V_{RS} - V_{thn}$). When RS goes low again, photocurrent discharges the pixel capacitance, causing V_d to decrease. M2 and the bias device at the bottom of the column line (shown in Fig. 4.6) form a source follower, which buffers the pixel voltage from the column line, V_o. When ROW is set high, the switch M3 is turned on, and the row is selected. The source follower drives the column line V_o to $V_d - V_{gs2}$.

The architecture of the image sensing circuitry is shown in Fig. 4.6. The figure indicates an array with N rows and M columns. The column amplifier used in this design accomplishes a few things: 1) it buffers the column line from the output line (V_{out}); and 2) it shifts and am-

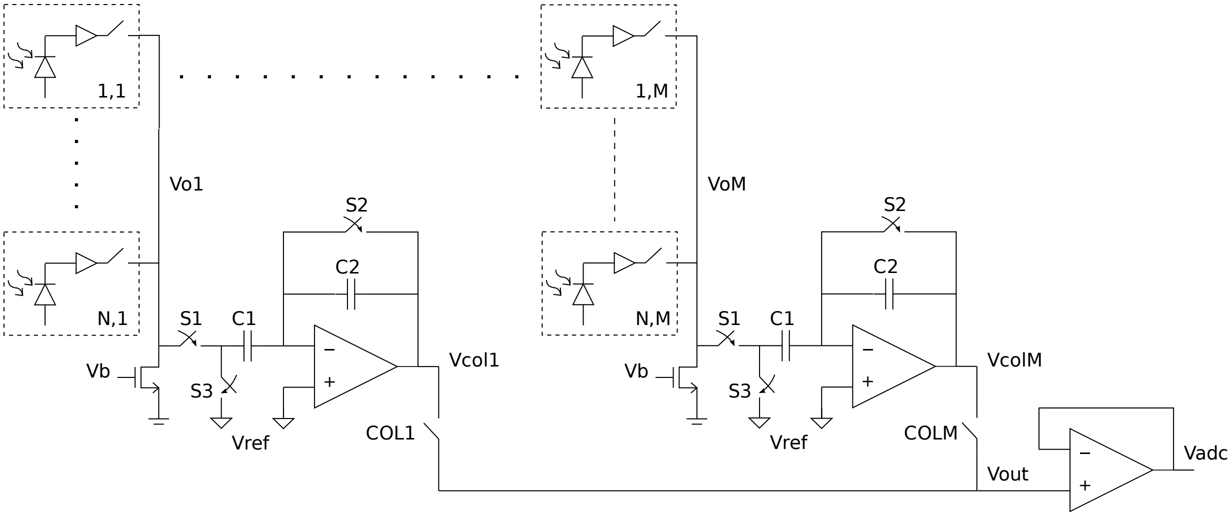


Figure 4.6: The architecture of the image sensor circuitry.

plifies the column voltage to maximize signal swing. Regarding the timing, essentially phase 1 ($T_{SAMPCOL}$) is the sampling phase where S1 and S2 are closed, and S3 is open. In phase 2 ($T_{HOLDCOL}$) the switches all flip, and the input voltage is amplified at the output. The gain is set by the ratio $C1/C2$.

The voltages from the pixel source followers are decreasing voltages. When no light hits the pixel, the pixel voltage remains high, and the output voltage remains high. When light hits the pixel, the pixel voltage decreases, and the output voltage decreases. Ideally V_{ref} would be ground in order to maximize the signal swing. However, V_{ref} can only be as low as the minimum input common mode voltage of the amplifier.

In the time that the pixel outputs from one row are sampled, each of the column outputs must be multiplexed to V_{out}, which is then driven off-chip through a buffer to a sampling circuit and ADC.

The ROWx signals are the outputs of a circular shift register with one 1 loaded into it. Each cycle of ROWCLK activates the next row. The COLx signals controlling the column multiplexer (MUX) are set in the same way.

4.2.3 LED Die Bondpads

Four large bondpads were placed on the top metal layer adjacent to the imaging circuitry for the purposes of bonding LED dies to the chip. These bondpads were each 500μm x 500μm (refer to Fig. 4.4). All of the bondpads have metal slotting, due to their large size. Typically bondpads include a full stack of all metal layers to improve their mechanical stability and prevent damage during the bonding process (the amsp35 technology design rules require all metal layers to be included in the bondpad stack). There was concern that this would be even more important for the extra large bondpads. However, when bondpads for flip-chip bonding are included in a design, it is beneficial to only have the top metal included, and not the full metal stack, because that would allow for circuitry and routing underneath the bondpad, which would save significant area. This is especially important in this design with such large area bondpads. In OR1, bondpad 1 (LED1) included the full metal stack (MET1 to MET4). Bondpad 2 (LED2) included MET2 to MET4, since its driver circuit required MET1 for connections. Bondpads 3 (LED3) and 4 (LED4) included only MET4. It turned out that both types of bondpads worked fine in this case.

4.3 Packaging

Packaging of OR1 required extra steps to bond the LED dies onto the large bondpads on the chip. The top electrodes of the LED dies were then wirebonded out to package pins. This special bonding was arranged through CMC Microsystems [57].

4.3.1 The LED Dies

Samples for two different LED dies were acquired from two different companies. The green die (LED2) is the CPR3ED (C527EZ400-0205-2) from Cree [22]. The yellow dies (LED1, LED3, LED4) are the ED-113SG from Optotech [70]. The LEDs have the specifications shown in Table 4.1.

LED	CPR3ED	ED-113SG
Wavelength	527nm	570nm
Full-Width Half Max.	35nm	26nm
Current	25mA	20mA
Radiant Flux	10mW	3mcd (min)
Width	380μm x 380μm	285μm x 285μm
Height	170μm	255μm
Junction Height	$\sim$170μm [a]	165μm
Radiation pattern	Lambertian	isotropic

[a] The junction is made from thin-films deposited at the surface of the substrate.

Table 4.1: LED die specifications. The CPR3ED is C527EZ400-0205-2.

4.4 Results

This section discusses the testing results for the chip and compares the results to predicted MATLAB model results.

4.4.1 Test Setup

A custom printed circuit board (OR1PCB) was designed with Altium Designer [4], obtained through CMC Microsystems. OR1PCB interfaced with an Opal Kelly XEM6010 field-programmable gate array (FPGA) circuit board [52], through high-density connectors. Opal Kelly FPGA boards along with their HDL (hardware description language) endpoints and API (application programming interface) software (Python, CPP, MATLAB) provide a USB interface to talk to the FPGA. Fig. 4.7 shows a high-level block diagram of the testing setup. A black cover is placed over the chip to eliminate ambient light and unwanted optical reflections.

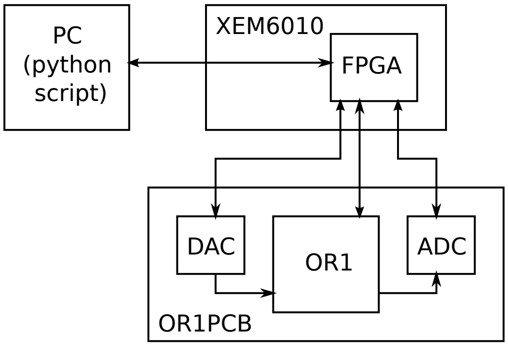


Figure 4.7: Block diagram of OR1 testing setup.

4.4.2 Testing Procedure

Initial electrical testing confirmed that the chip was functioning correctly, which verified that the mixed-signal design flow is complete and trustworthy. The simple on-chip LED driver circuits worked correctly (for LED4), as did the off-chip driver circuit (LED2 and LED3). The imaging circuitry (pixel circuitry and digital control) worked correctly as well.

In order to characterize how much baseline light from an LED die hits each pixel in the array, the LED intensity (proportional to the LED current, I_{LED}) is set to a level such that the pixels do not saturate during the integration time correponding to the desired frame rate. In order to determine the actual light intensity, I_{LED} is measured, and the current value is mapped to intensity using the LED datasheet (this is not an exact measurement of light intensity, but is sufficient for our purposes given the other sources of error in the system, discussed below). Once the intensity is set, several consecutive frames are captured and digitized by an off-chip ADC (AD7985) on OR1PCB, and the captured data is processed offline with Python.

4.4.3 Sources of Error and Data Normalization

There are several variables affecting the absolute voltage levels (in both the model and experimental results) that are not known exactly. The most significant sources of error are:

- the optical characteristics (peak wavelength, conversion efficiency, etc.) and geometrical characteristics (width and length of emission area) vary from die to die
- the mapping of current to intensity through the datasheet is quite prone to error in this case because of the poor resolution of the curve at the current levels used in this work

- the ED-113SG datasheet only provides a minimum value for luminous intensity, but no typical or maximum value; this is likely the most significant source of error
- the thickness of epoxy bonding the LED die to the bondpad is not known exactly, which means the LED junction height relative to the imager surface is unknown

Other sources of error include:

- Fixed pattern noise due to component variation in the circuitry.
- The column amplifiers in the imaging circuitry can have significant variation in input-offset voltage (10s of mV), and each pixel will have some variation as well. This is verified by observing the imager output voltage when the LEDs are turned off. Ideally the voltage should be constant for all pixels, but instead has variation. However, the effects of variation are somewhat eliminated by considering the change in voltage between when the LEDs are on and off, and not just the voltages when the LEDs are on. Averaging the results of many frames reduces noise and the remaining variation is indicative of pixel offsets.
- The conversion from candelas to watts in the model is not exact, because 555nm monochromatic light was assumed. The LED peak wavelength is nominally 570nm (ED-113SG), but it is not monochromatic either. There is a spectrum with 26nm FWHM (however, assuming the spectrum is symmetric, the average energy would be that correponding to the peak wavelength).
- The LED operating temperatures are unknown.
- The exact height of the p-n junctions inside the LED dies is unknown (eyeballed from datasheet for ED-113SG, near the surface for CPR3ED).

- The pixel circuitry is not shielded from light in this chip. Therefore, there could be some modulation of the biasing if light is hitting the pixel circuitry (however, given that the circuits are working, this may be insignificant).

- There is some nonlinearity in the 3T APS circuit.

- (x,y) coordinates of the array and LED for the model are "measured" from the die photo.

- It is not certain that there are no reflections inside the chip package.

Given these sources of error, the interest here is not in comparing the absolute values of the predicted and measured voltages. Rather, the relative change in voltage from pixel to pixel, i.e. the shapes of the curves, and how well they match, is more important. Worst-case parameter values can be used to account for absolute voltage levels when modeling future system designs.

Therefore, the difference between predicted model results and experimentally measured results is calculated after normalizing each of the data sets. The model data is first shifted, and then scaled, such that the resulting root mean square error (RMSE) between the datasets is minimized. The RMSE is calculated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{x}_i - x_i)^2} \tag{4.5}$$

where x_i are the ideal model values, $\hat{x}_i$ are the experimental values, and n is the number of data points (number of pixels in this case).

4.4.4 Dark Levels

The average zero-light voltage is found for each pixel by averaging its zero-light voltage over 192 frames with a 1 second delay between each frame capture. The resulting average zero-light voltages for all pixels are in the range of 1.7525±0.0015V.

4.4.5 LED1

There was an issue with LED1, and it would not produce light (in both chips that were tested), so there are no results. There is no obvious problem in the chip or PCB design, or packaging, so the reason is unknown.

4.4.6 LED2

LED2 has a Lambertian radiation pattern. Modeling of a Lambertian radiation pattern indicates that zero light should exit the die in the downward direction, and that zero light should hit the pixels. Therefore LED2 is not compared with models in this chip. LED2 was included in the design to test its functionality and evaluate it for use in future system designs.

4.4.7 LED3

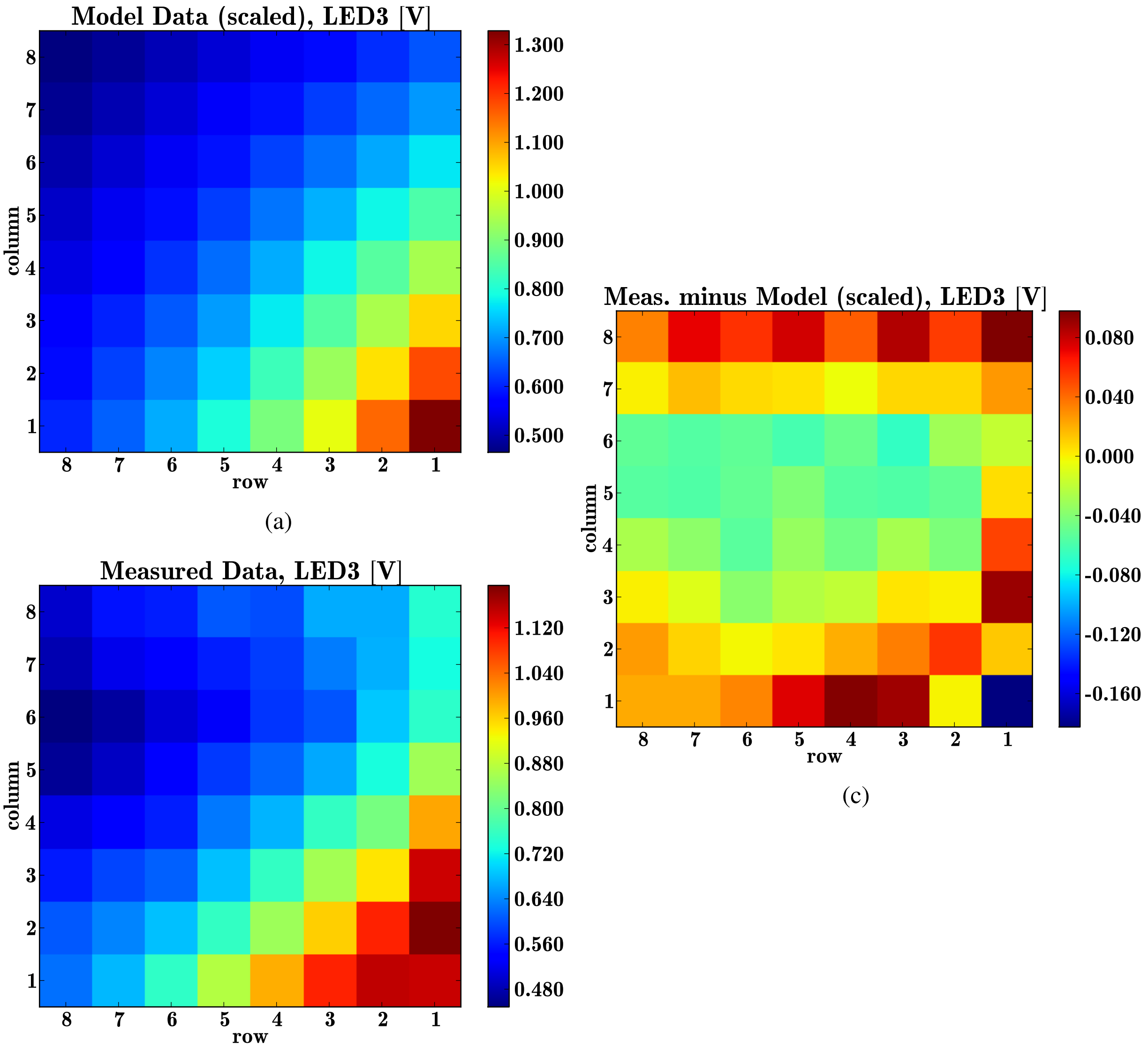
Fig. 4.8 shows: (a) the predicted *change* in voltage for each pixel, shifted and scaled to minimize the RMSE; (b) the measured *change* in voltage for each pixel with LED3 turned on (the difference between the sampled voltage and the average zero-light voltage); (c) the difference between the measured values of (b) and the normalized model values of (a).

The change in voltage increases as more light hits a pixel during its integration time. As expected, pixels closer to the LED have a larger change in voltage than pixels further from the LED. The top right corner of Fig. 4.8(b) indicates some deviation from predicted results; this is likely due to reflections off of bondwires near that corner of the imager. For LED3 the optimized values for the shift and scale factor are 0.068V and 4.23V/V, resulting in an RMSE of 0.0509V.

4.4.8 LED4

Fig. 4.9 shows: (a) the predicted *change* in voltage for each pixel, shifted and scaled to minimize the RMSE; (b) the measured *change* in voltage for each pixel with LED4 turned on (the difference between the sampled voltage and the average zero-light voltage); (c) the difference between the measured values of (b) and the normalized model values of (a).

As with LED3, the change in voltage increases as more light hits a pixel during its integration time. As expected, pixels closer to the LED have a larger change in voltage than pixels further from the LED. Here, there is no obvious deviation from the model. For LED4 the optimized values for the shift and scale factor are 0.054V and 1.62V/V, resulting in an RMSE of 0.0437V.

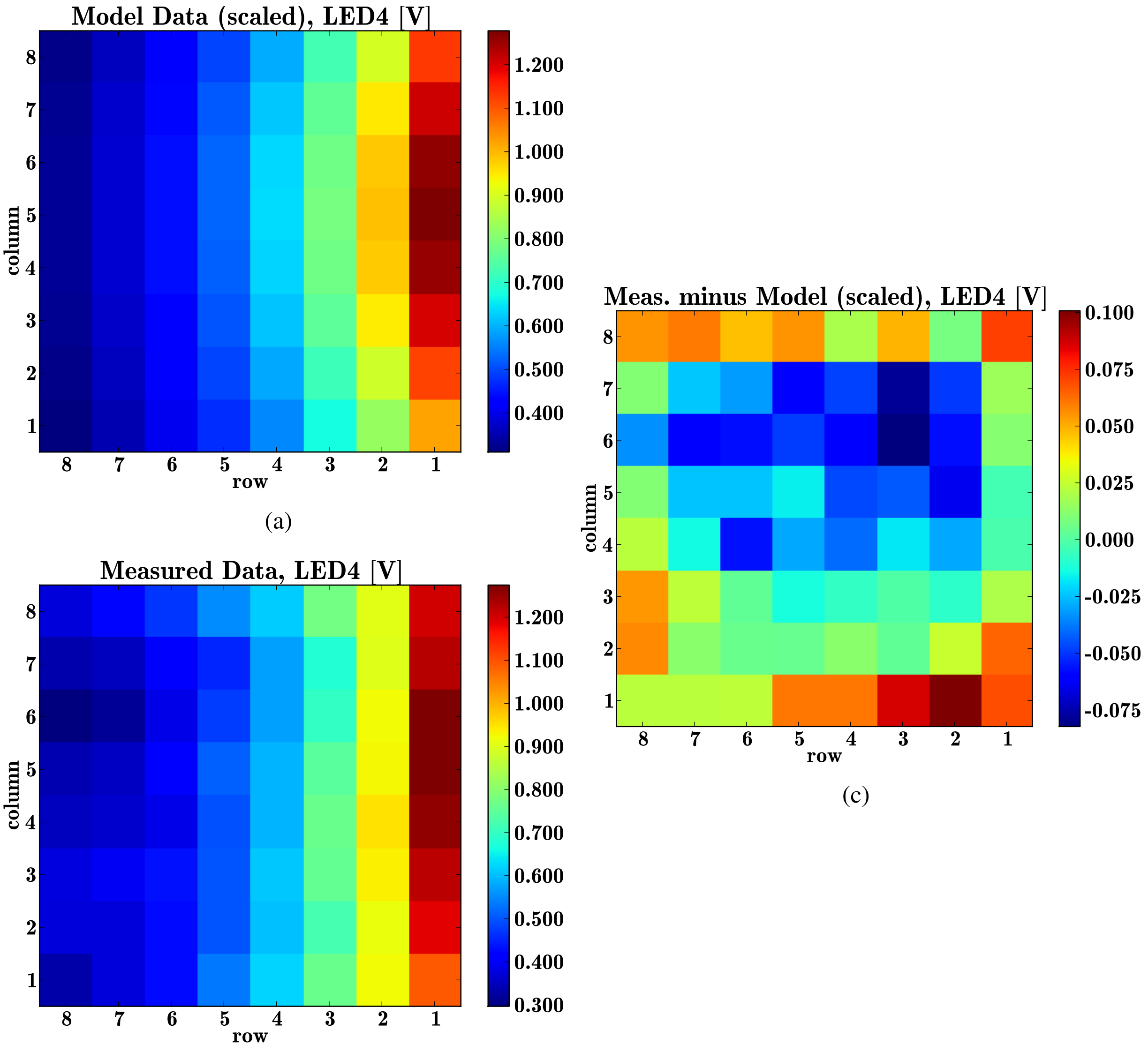


(a)

(b)

(c)

Figure 4.8: LED3 results: (a) predicted voltages (scaled to minimize RMSE); (b) measured change in voltage; (c) difference between normalized model data and measured data.



(a)

(b)

(c)

Figure 4.9: LED4 results: (a) predicted voltages (scaled to minimize RMSE); (b) measured change in voltage; (c) difference between normalized model data and measured data.

4.5 Summary

A low-complexity distributed point source model for predicting the near-field light distribution from an excitation source was described. The model was used to predict the signal in a CMOS imager due to LED dies bonded directly on top of the chip surface to large bondpads. A test chip for verifying the optical modeling was successfully designed and tested. Prior to this work it was not clear what kind of light distribution would be seen below the bottom surface of an unpackaged LED die. Experimental measurements of the ED-113SG isotropically radiating dies show that significant amounts of light do exit the die in the downwards direction. For the ED-113SG dies, the distributed point source model is applicable, which indicates that scattering of light inside the LED die substrate is not a significant effect. In general, such a model can be used to predict the excitation light power incident at different locations within this type of integrated device with near-field lighting conditions.

Chapter 5

Optical Model for Implantable Recording Device

The previous chapter established the usefulness of optical modeling for predicting the near-field light distribution in an optical device. We are now interested in a full neural recording system with the device implanted in tissue, or alternatively a tissue sample sitting on top of the device.

This chapter begins by describing light propagation in tissue, which is a very important consideration for implantable optical devices. The effects of light propagation in tissue, as well as fluorescence, are added to the basic model to create a model for an implantable optical neural recording device. The model works from excitation power from off-chip LED sources to predict fluorescent power from fluorophores in tissue and ultimately the voltage signal in the imager that absorbs the fluorescent light. The model is used to choose design specifications for the OR2 chip and the overall system to perform optical detection of individual spikes from individual neurons.

5.1 Light Propagation in Tissue

The amount of fluorescence produced by a fluorophore is dependent on the local excitation photon flux, which can be defined as the number of incident photons per unit area per unit time [101]. The photon flux at the fluorophore is highly dependent on the propagation of light through tissue. Therefore, when designing an implantable optical recording system it is essential to model light propagation in tissue. Usually the photon flux is reported as a light power density in mW/mm^2. Light power density is the photon flux multiplied by the energy of a single photon.

As light propagates through the system, it will be absorbed and scattered, and the intensity of light at the end of the optical path will be lower than at the start. We need to model how the light intensity decreases from its starting point, the LED, to the end point, the photodetector. The attenuation will dictate the excitation power required in the LED that will ultimately produce a detectable signal with a certain SNR at the photodetector.

The following description is based on that in Chapter 2 from [47] and from [58]. The absorption of light in a material can be modeled by the Beer-Lambert Law:

$$I = I_0 e^{-\mu_a d} \quad [\mathrm{W/m^2}], \tag{5.1}$$

where I is the intensity of light transmitted through the material in W/m^2, I_0 is the intensity of incident light in W/m^2, μ_a is the *absorption coefficient* of the material in mm^{-1}, and d is the distance traveled through, or thickness of, the material in mm.

Elastic (no energy loss) scattering in a material can also be modeled by the Beer-Lambert

Law:

$$I = I_0 e^{-\mu_s d} \quad [\mathrm{W/m^2}], \tag{5.2}$$

where I is the intensity of light transmitted through the material in W/m^2, I_0 is the intensity of incident light in W/m^2, μ_s is the *scattering coefficient* of the material in mm^{-1}, and d is the distance traveled through, or thickness of, the material in mm.

When scattering is the dominant factor the material can be modeled as a diffusing substance. The amount of light that is scattered in the forward direction after a scattering event is modeled by g, the *anisotropy coefficient*, or cosine of scatter. A rule of thumb for determining when the diffusion approximation is valid is to test when the following condition is true [58]:

$$\frac{\mu_s(1-g)}{\mu_a} > 10\ . \tag{5.3}$$

When the diffusion model is valid, scattering can be represented with a modified scattering coefficient, $\mu_s^{'} = (1-g)\mu_s$. In tissue, it is usually the case that scattering is significant, and the anisotropy coefficient is usually taken to be $g = 0.9$.

Combining the effects of absorption and scattering, the attenutation of light in tissue can be modeled by the following equation:

$$I = I_0 e^{-(\mu_a + (1-g)\mu_s)d} = I_0 e^{-(\mu_a + \mu_s^{'})d} \quad [\mathrm{W/m^2}]. \tag{5.4}$$

It is important to note that the absorption and scattering coefficients are dependent on wavelength. The authors in [32] did a study characterizing the absorption and scattering coefficients of brain tissue. Through *in vitro* experiments on 82 samples of brain tissue they found μ_a and $\mu_s^{'}$ for a range of wavelengths for grey matter, white matter, and glioma tissues. Their results are used here in the optical modeling.

Certain fluids, such as HbO_2 (oxygenated hemoglobin), HbR (de-oxygenated hemoglobin), and water have a significant effect on the propagation of light in tissue [44]. The effects are wavelength dependent. Light of less than 600nm is strongly absorbed by hemoglobin molecules. Light of more than 900nm is strongly absorbed by water. As mentioned previously, certain structures like flavins autofluoresce, which contributes to background noise. These structures generally emit in the range of about 500-550nm [44]. Considering light absorption in neural tissue and autofluorescence, wavelengths in the near-infrared (NIR), between 600nm and 900nm are preferred for deep-tissue imaging. However, at higher wavelengths where absorption is lower, scattered light also will travel further. In this work where proximity-based imaging is used, this extra scatter can reduce spatial resolution. Therefore, the choice to use higher wavelengths would require more consideration.

5.2 General Optical Path Model

This section combines the equations for light attenuation in tissue with the modeling techniques of Chapter 4 to derive a general model and calculations that can be used to describe an implantable fluorescence detection system from LED optical power to photodetector voltage signal.

5.2.1 Noise and Signal-to-Noise-Ratio

The Signal in Optical Recording

When a VSFP is excited it will fluoresce, but the intensity of fluorescence changes based on the neuron membrane voltage. Therefore, the signal of interest is the *change* in fluorescence intensity at the detector due to a change in membrane voltage. The signal power is proportional to the change in fluorescence intensity,

$$\Delta I_{fl} = I_{fl,peak} - I_{fl,base} \ , \tag{5.5}$$

where $I_{fl,peak}$ is the peak fluorescence level corresponding to the peak membrane voltage and $I_{fl,base}$ is the baseline fluorescence level at the resting potential.

Noise in Optical Detection

For a fluorescence detection system the lower bound of the noise (neglecting circuit noise) is set by the stochastic nature of photon arrival at the detector [61]. When the number of photons arriving in a given time interval is relatively small, fluctuations around the mean can be significant, which reduces the SNR. This is called shot noise [85]. The number of photons arriving at the detector in a given time interval T_s (the sample time) can be modeled by the Poisson distribution. If the average number of events (photons arriving at the detector) is given by N, then the probability of k events occuring in the time interval is given by the Poisson probability mass function (PMF) [90]:

$$f(k,N) = \frac{N^k e^{-N}}{k!} \ . \tag{5.6}$$

When N is large, the Poisson distribution can be approximated as a Gaussian distribution where the variance is equal to the mean; in this case N, and the standard deviation is $\sqrt{N}$. This defines the Poisson noise, or shot noise. The SNR for an optical detection system obeying a Poisson distribution can be defined as follows [103]:

$$SNR = \frac{N}{\sqrt{N}} = \sqrt{N} \ . \tag{5.7}$$

The SNR of Optical Detection

For a fluorescence detection system, the optical SNR (before conversion to an electrical signal and corruption with electrical noise) is proportional to the ratio of ΔI_{fl} to the shot noise of the baseline signal, $\sqrt{N_{bs}}$, where N_{bs} is the number of baseline photons detected [85].

An implantable fluorescence detection system has other sources of noise besides the shot noise of the optical signal. Spatial movement caused by the heartbeat and breathing, tissue autofluorescence, changes in hemoglobin oxygenation all contribute to noise [2]. There is also electrical noise introduced by the detector, including thermal, flicker and shot noise, and noise due to dark current [23]. Ideally the system would introduce minimal noise such that only the shot noise of the optical signal limits the SNR. Therefore, we are interested in translating the photon shot noise to an equivalent noise voltage, which would become the design specification for the noise performance of the detection circuitry.

Assuming a general CTIA design as shown in Fig. 5.1, the photocurrent is integrated on the feedback capacitance for a time period T_s to generate the output voltage, as described by:

$$v_{out}(t) = \frac{1}{Cfb} \int i_{pd} dt \ , \tag{5.8}$$

where, Cfb is the amplifier feedback capacitance, and i_{pd} is the photocurrent. Given the rate, r_{bs}, of baseline photons hitting the detector per second, the output RMS noise voltage can be calculated as follows. The time interval for the Poisson distribution is T_s, so the number of baseline photons that hit the detector in T_s is

$$N_{bs} = r_{bs} T_s \quad \text{[photons]}, \tag{5.9}$$

and the number of baseline electrons generated in T_s is

$$N_{be} = N_{bs} QE_{pd} \quad [\mathrm{e}^-], \tag{5.10}$$

where QE_{pd} is the quantum efficiency of the photodetector. The corresponding electrical shot noise due to the optical signal is then $\sqrt{N_{be}}$, which can be considered an RMS value. The RMS noise charge integrated on the feedback capacitance, C_{fb}, during T_s is $\sqrt{N_{be}}q$, where q is the

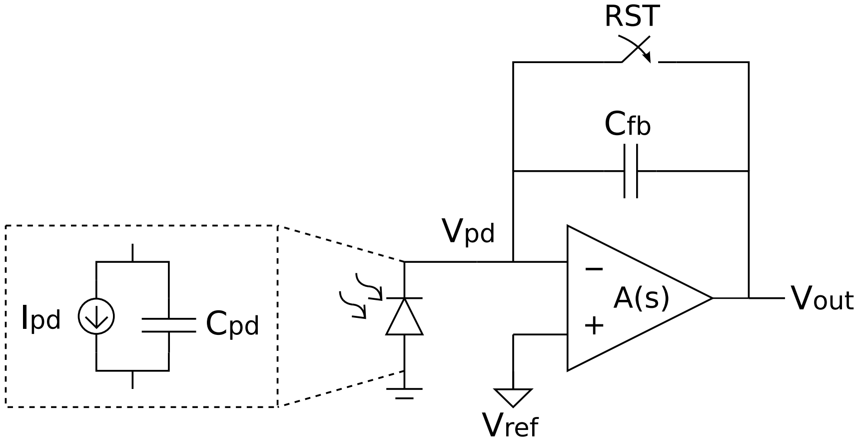


Figure 5.1: Generic capacitive transimpedance amplifier (CTIA) circuit. Photocurrent is integrated on the feedback capacitor to produce the output voltage.

charge of an electron. Finally, the RMS noise voltage is given by

$$V_{n,rms} = \frac{\sqrt{N_{be}}q}{C_{fb}} = \sqrt{r_{bs} Q E_{pd} T_s} \frac{q}{C_{fb}} \quad [\mathrm{V_{rms}}]. \tag{5.11}$$

The average signal voltage can be found in a similar manner. The extra number of photons hitting the detector during T_s assuming a spike with voltage V_{ap} during all of T_s is:

$$N_s = N_{bs} \frac{S_{fl}}{100} \frac{V_{ap}}{100} \quad [\text{photons}], \tag{5.12}$$

where V_{ap} is the actual action potential magnitude $[mV]$, and the sensitivity of the fluorophore is $S_{fl} = \Delta I_{fl} / I_{fl}$ in %/100mV. The peak number of electrons generated during T_s is then

$$N_{se} = N_s Q E_{pd} \quad [\mathrm{e}^{-}]. \tag{5.13}$$

The signal voltage is then given by

$$V_s = N_{se}\frac{q}{C_{fb}} \quad [\mathrm{V}]. \tag{5.14}$$

Finally, the SNR (in terms of electrical power) is given by

$$SNR_e = \frac{V_s^2}{V_{n,rms}^2} = r_{bs}T_sQE_{pd}\left(\frac{S_{fl}}{100}\frac{V_{ap}}{100}\right)^2 \quad [\mathrm{V}^2/\mathrm{V}^2]. \tag{5.15}$$

The SNR_e can be increased by improving the fluorophore sensitivity, increasing the photon flux at the detector, improving the quantum yield of the detector, and increasing the sampling time.

It is important that the sum of the baseline and signal voltages is not enough to saturate the detector. Perhaps the most important factor is the distance between the fluorophore and the detector, given its significant affect on the optical signal path, which will be seen in the next section. An implantable detector might have to calibrate or dynamically adjust the power of each LED, or the sampling time, to ensure saturation does not occur.

5.2.2 General System

There are several possible geometrical designs for the implant, in terms of the sizes of the excitation sources and photodetectors and their relative locations. The general model presented here considers a single excitation point source, a single neuron, and a single photodetector; these comprise a single optical path. The center of the photodetector is at the origin, and the other objects are arbitrary coordinates relative to the photodetector. The geometry of the model is shown in Fig. 5.2.

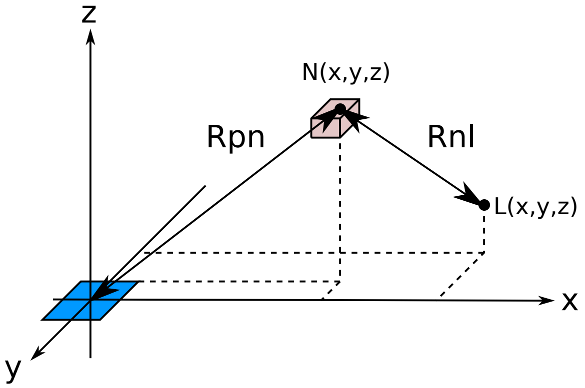


Figure 5.2: Coordinate system and geometry for the general optical path model. The blue rectangle is the photodetector (surface in the x-y plane); the pink cube is a neuron.

5.2.3 LED-to-Neuron Path

Consider the model for light propagation from the LED to the neuron in Fig. 5.3, where the LED is assumed to be a point source, the neuron has a surface area (facing the LED) of A_n, a thickness of d_n, and $d_n \ll R_{nl}$, where R_{nl} is the distance between the neuron and the LED. This point-source model is also used for the path from the neuron to the photodetector, where the neuron is assumed to be a point source, and the photodetector has a surface area of A_{pd}.

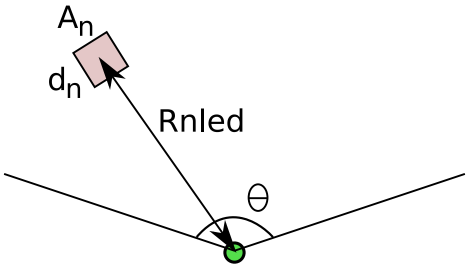


Figure 5.3: Model of the optical path from the LED to the neuron. The LED is modeled as a point source. The neuron has a depth d_n and surface area facing the LED A_n.

Now, the Beer-Lambert law is valid when the excitation source is a laser, where there is very little spatial spreading of light from the source. When the excitation source is an LED, this model is not entirely accurate because LEDs emit light with spreading. Here, the same calculations as were used in the basic model for OR1 can be used to find the light power hitting the neuron from the excitation point source. Combining Eq. 4.2 with the Beer-Lambert law, the optical power from the LED seen at the neuron can be modeled by

$$P_n = \frac{P_0}{\omega R_{nl}^2} \cdot A_n \cdot e^{-\mu_t R_{nl}} \quad [\mathrm{W}] \ , \tag{5.16}$$

where $\mu_t = \mu_a + \mu_s^{'}$ is the total coefficient of absorption and scattering in tissue at the excitation wavelength.

5.2.4 Fluorophore Absorption

If we assume that the fluorophores are evenly distributed throughout the volume of the neuron, the optical power absorbed by the fluorophores in the neuron can then be found using the Beer-Lambert law:

$$P_{n,abs} = P_n \cdot (1 - e^{-\mu_n d_n}) \quad [\mathrm{W}] \ , \tag{5.17}$$

where μ_n is the absorption coefficient of the neuron (due to the fluorophores). The optical power of fluorescence is then given by

$$P_{fl} = P_{n,abs} QY_{fl} \quad [\mathrm{W}] \ , \tag{5.18}$$

where QY_{fl} is the quantum yield of the fluorophore (the number of photons emitted for each photon absorbed).

5.2.5 Neuron-to-Photodetector Path

The optical path from the fluorophores (neuron) to the photodetector can now be modeled in a similar way to the LED-to-neuron path. Assuming the fluorophores together are an isotropically radiating point source, and the photodetector has a surface area of A_{pd}, the optical power absorbed by the photodetector can be modeled as

$$P_{pd} = \frac{P_{fl}}{4\pi R_{pn}^2} \cdot A_{pd} \cdot e^{-\mu_f R_{pn}} \quad [\mathrm{W}] \ , \tag{5.19}$$

where $\mu_f = \mu_a + \mu_s^{'}$ is the total absorption coefficient of the tissue at the wavelength of fluorescence. Note that again the model is assuming that a far-field scenario applies, and that the photodetector area is perpendicular to the vector R_{pn}. The assumption is that any error in the calculation due to a non-perpendicular angle is negligible.

5.2.6 Photodetector Absorption and SNR

Given P_{pd} and the energy of a photon at the fluorophore emission wavelength $E_{ph} = hc/\lambda_{fl}$ [J], where h is Planck's constant, c is the speed of light, and λ_{fl} is the wavelength of the photons, we can find the number of baseline photons that hit the photodetector during T_s:

$$N_{bs} = \frac{P_{pd}}{E_{ph}} T_s \quad [\text{photons}] \ . \tag{5.20}$$

Referring back to the section on SNR, and combining the optical path equations, we can derive a single equation relating the SNR to all of the system parameters:

$$SNR = \frac{P_0 \cdot A_n \cdot A_{pd} \cdot QY_{fl} \cdot QE_{pd} \cdot T_s \cdot (1 - e^{-\mu_n d_n})}{\omega 4\pi \cdot E_{ph} \cdot R_{nl}^2 \cdot R_{pn}^2 \cdot e^{\mu_t R_{nl}} \cdot e^{\mu_f R_{pn}}} \left(\frac{S_{fl}}{100} \frac{V_{ap}}{100} \right)^2 . \qquad (5.21)$$

It is clear that the distances R_{nl} and R_{pn} have a very strong effect on the SNR. Consider the parameters that the circuit designer can control:

- P_0: the LED power should be high to maximize SNR, but in practice it is limited by the acceptable power consumption of the device and tissue heating effects, as well as possible saturation of the photodetector.

- A_{pd}: the photodetector area should be large to maximize SNR, but it will be limited by the required spatial resolution of the detector.

- T_s: the sampling time should be large to maximize SNR, but it will be limited by the required sampling frequency of the system.

- R_{nl} and R_{pn}: the distance between the light source and fluorophores, and between the fluorophores and photodetector, should be small to maximize SNR. These distances are perhaps the key to the system design. They can be controlled to some degree by the design of the system geometry. However, from another point of view, in a neural recording system being able to record deeper into tissue is desirable. Optimizing the geometry of the system design can only get you so far due to the physical limitations of light propagation through tissue.

5.3 MATLAB Modeling

The equations described in the previous section can be applied to a MATLAB model for a given system design. In general, the MATLAB models for different system designs are all very similar. The following is a list of key points about the MATLAB modeling, followed by pseudocode describing what the models calculate:

- The equations in the previous section assume a single point source for the LEDs (and neurons). However, that assumption is not necessarily valid in a near-field system. Therefore each LED can be modeled as a distributed point source. The point sources are distributed over the planar cross-section where the p-n junction is located. The total LED power is divided equally between the point sources.

- Each “neuron” in the model is also represented by a square surface that is split up into smaller squares. The total power hitting each neuron is the sum of the power hitting each small square. The reason for splitting up the neuron area is it allows us to model a near-field situation by creating a set of far-field situations and summing the results (far-field allows us to assume a point source model).

- The models only account for light that directly hits its target, and do not account for any reflections in the system (advanced optical ray tracing software would be needed for that). From the point of view of light collection efficiency, this could be considered a pessimistic model.

- The models including fluorophores do not necessarily account for excitation light being absorbed in the photodiodes. This is related to the last point. However, optical excitation filters and system geometry are designed to take care of this.

The pseudocode in Algorithm 1 describes at a high level what the MATLAB code is doing for the OR2 system design (certain steps may or may not be needed for different system models):

```
define system geometry and object parameters;
find LED power emitted in each direction given the radiation pattern;
foreach neuron location do
    foreach LED do
        foreach point source location on LED do
            add excitation power hitting neuron area to total power hitting neuron
        end
    end
end
foreach neuron do
    calculate baseline power of fluorescence from absorbed power
end
create spike waveform;
foreach neuron do
    modulate neuron fluorescence over time with randomly inserted spike waveform
end
foreach pixel location do
    find fluorescent power hitting pixel vs. time due to all neurons
end
foreach pixel location do
    foreach sample time do
        find total number of photons collected during sample time
        find pixel voltage at the sample time
    end
end
```

Algorithm 1: General MATLAB model pseudocode.

5.4 OR2 Optical Model

The general optical path model in the previous section provides the basis for constructing more complex system models. The goal for the OR2 system is to have multiple LED dies bonded beside the imager chip. The main reason for moving the LEDs from the top of the imager chip (as they were for OR1) to beside the imager chip is to allow the tissue to sit closer to the photodetectors, which, based on MATLAB model results, improves detection resolution.

To model the system, the LED dies are represented with distributed point sources, as they were in the OR1 model. The "tissue" sits on top of the imager chip (and LED dies). Multiple neurons are placed at various locations within the tissue space. The imager chip includes a multi-pixel array. In order to ensure the far-field equations are applicable to the near-field system, each neuron can also be split into a grid of smaller areas, as can each photodiode in the imager array. Essentially, the model simply repeats the equations described in the previous section, summing up the contribution of each distributed LED point source to each neuron (or smaller neuron area), and each neuron (or smaller neuron area) to each photodetector (or smaller photodetector area). Lastly, as in OR1, the total power absorbed in each photodetector is then converted to the sample voltage signal.

The OR2 design is meant to detect action potentials, which means it has to compare recorded pixel voltages over time. Therefore, the model calculations described above are essentially modulated by neural waveforms (sine wave snippets of appropriate frequency) in order to model the response over time. This mimics the effect of the reporter proteins.

5.4.1 GCaMP3 Model

First, a model relating to a common genetically encoded calcium indicator used in neural recording, GCaMP3, is considered. GCaMP3 is used for monitoring neural activity in several different cases, including fruit flies, zebrafish, worms, and mice [95], [3], [102]. The case of recording in the mouse cortex is one where the implantable devices proposed in this thesis might be applied. GCaMP3 allows for the simultaneous chronic (several months) recording of dozens of neurons in behaving mice. Multiple recording sessions make it possible to track learning-induced changes in neural activity [95]. The authors of [102] used mice transfected with GCaMP3 in various recording experiments. Table 5.1 shows the parameter values used in the model. The GCaMP3 parameters are from [3], except for Cfl, which is from [102]. The rise time for GCaMP3 is approximately 100 ms, therefore the sine wave snippet frequency is chosen as 10 Hz. The LED parameters are for the CPR3ED (but assuming a different excitation wavelength). The LED excitation power is chosen such that the pixels would not be saturated. Table 5.2 shows the neuron locations used in the model. Note that the top surface of the LEDs is at the same height as the pixel array.

Fig. 5.4 shows the system geometry, including pixel, LED, and neuron locations. The colorbar indicates the baseline optical (fluorescent) power hitting the array. Note that the squares of the pixel array are not actually the pixels. Rather, the points of the grid are the centers of each pixel. Fig. 5.5 plots the fluorescent power emitted by each neuron over time. Each neuron has a sine wave snippet inserted at a random time. Neuron 8 (N8) has the largest fluorescent power because it is in a position that sees a lot of excitation power from the LEDs. Neuron 3 (N3) has the smallest fluorescent power because it is in a position that does not see much excitation power

Parameter	Name	Value
simulation time step	$tScale$	1 ms
sine wave snippet frequency	$Fspike$	10 Hz
imager sampling time	Ts	40 ms
neuron resolution	kN	2 μm
LED resolution	kL	20 μm
pixel fill factor	FF	0.83
CTIA feedback capacitance	Cp	4 fF
photodiode responsivity	$Resp$	0.29 A/W
fluorophore quantum yield	$QYfl$	0.65
fluorophore emission wavelength	Lfl	515 nm
fluorophore extinction coefficient	$Exfl$	37700 L/(mol cm)
fluorophore concentration	Cfl	5 μM
fluorophore sensitivity	Sfl	10 %/100mV
LED excitation power	Pex	0.5 mW
LED excitation wavelength	Lex	500 nm
fluorescent wavelength absorption coefficient	aaf	0.23/1000 μm^{-1}
fluorescent wavelength scattering coefficient	asf	1.4/1000 μm^{-1}
LED wavelength absorption coefficient	aat	0.23/1000 μm^{-1}
LED wavelength scattering coefficient	ast	1.4/1000 μm^{-1}

Table 5.1: OR2 model parameters for GCaMP3 case. The GCaMP3 parameters are from [3], except for Cfl, which is from [102]. The rise time for GCaMP3 is approximately 100 ms, therefore the sine wave snippet frequency is chosen as 10 Hz. The LED parameters are for the CPR3ED (but assuming a different excitation wavelength). The LED excitation power is chosen such that the pixels would not be saturated.

Neuron	X	Y	Z
N1	0	0	300
N2	400	0	300
N3	400	400	300
N4	0	400	200
N5	-400	400	200
N6	-400	-400	100
N7	0	0	100
N8	-400	-400	300

Table 5.2: Neuron coordinates for OR2 models.

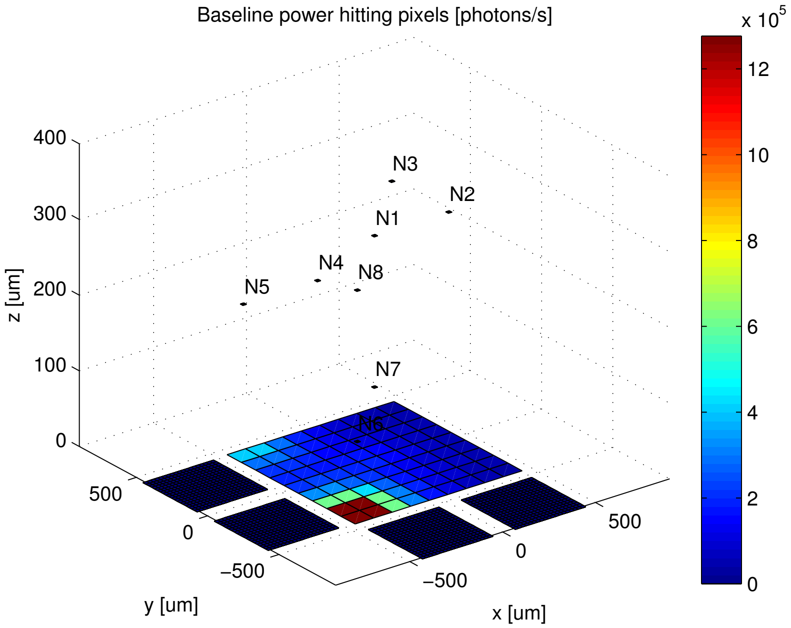


Figure 5.4: System plot of the OR2 model for GCaMP3, showing pixel, LED, and neuron locations, as well as the baseline optical (fluorescent) signal power hitting the array.

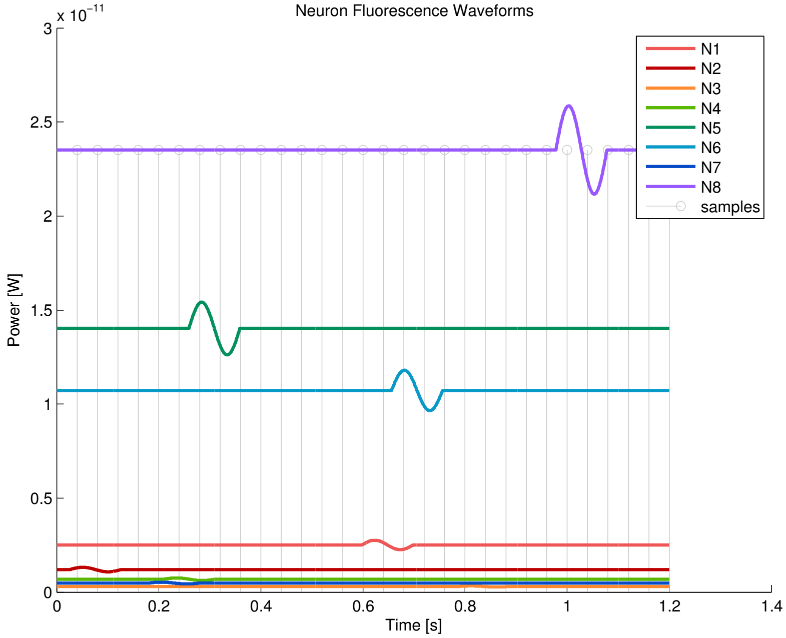


Figure 5.5: Plot of fluorescent output power for each neuron in the GCaMP3 model.

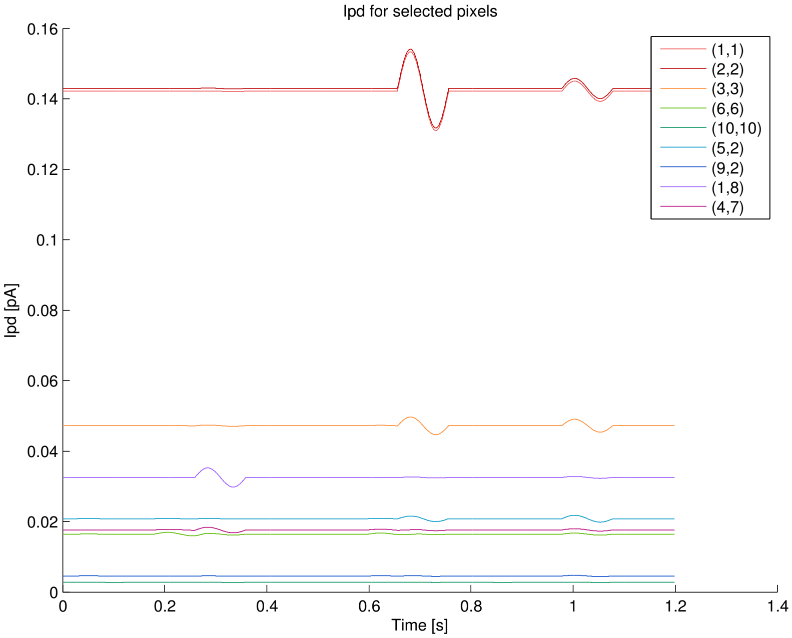


Figure 5.6: Photocurrent in selected pixels over time in the GCaMP3 model.

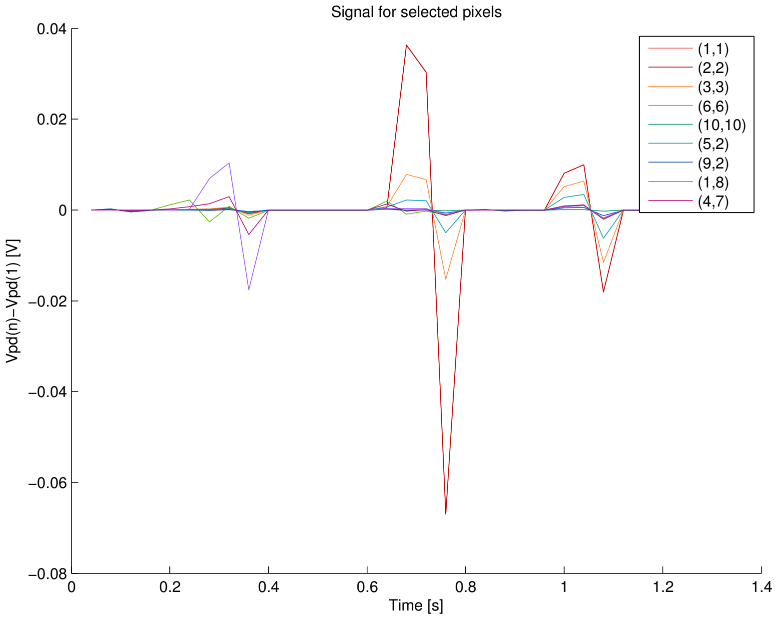


Figure 5.7: Sampled voltages in selected pixels over time in the GCaMP3 model.

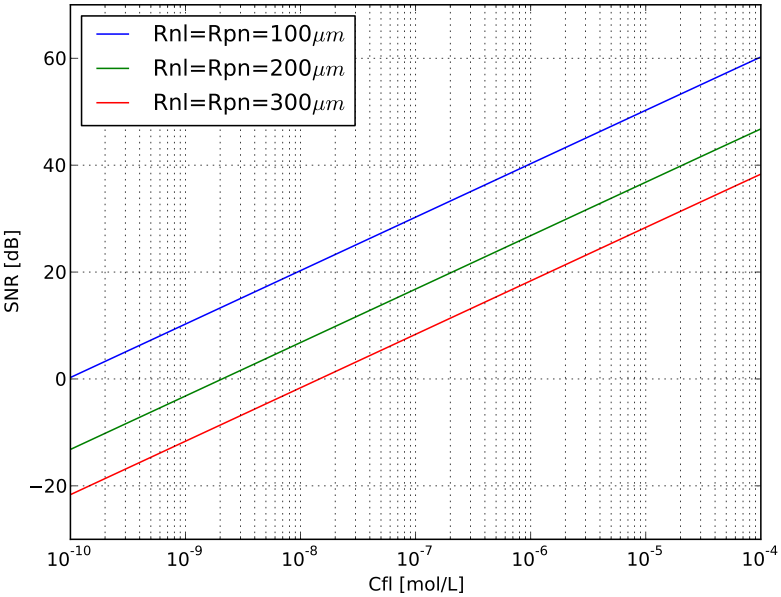


Figure 5.8: Plot of SNR vs. fluorophore concentration for the GCaMP3 model.

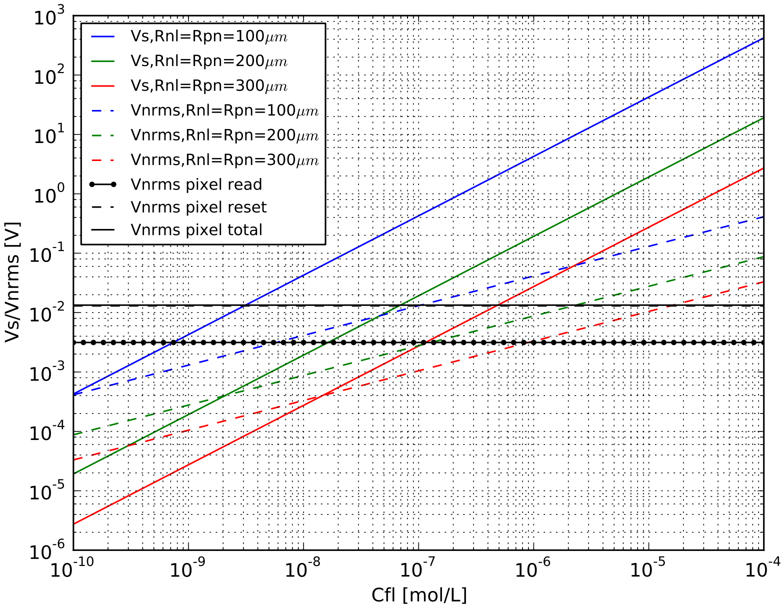


Figure 5.9: Plot of signal and noise voltages vs. fluorophore concentration for the GCaMP3 model. The signal and noise voltages due to the optical signal are Vs and Vnrms, shown for each of three distances. The read, reset and total noise voltages due to the pixel circuitry are also shown, in black.

from the LEDs. Fig. 5.6 plots the photocurrent in selected pixels over time. The currents change in response to the changing fluorescent power in the neurons. Pixels (1,1) at (-450,-450) and (2,2) at (-350,-350) have almost the exact same photocurrent, and have the largest photocurrent due to their position underneath N8, which provides the largest fluorescent signal. These pixels clearly see the spikes from N8 and N6, but do not really see any of the other spikes. Fig. 5.7 plots the signal voltage (sampled voltage minus baseline voltage when there is no spike) in selected pixels over time, at the 25Hz sampling rate. The plot is similar to the photocurrent plot, but here the actual voltage levels are shown, so it gives an idea of the resolution of the imager required to detect spikes. While some larger signals are in the 10s of mV range, resolution in the 1mV range would be good to detect spikes from some of the neurons that produce small signals. Fig. 5.8 plots the SNR (as in Eq. 5.21) vs. the fluorophore concentration, C_{fl}, for three different distance values. The SNR increases as C_{fl} increases, and increases as the distances R_{nl} and R_{pn} decrease. Fig. 5.9 plots the signal and noise voltages vs. fluorophore concentration, C_{fl}, for three different distance values. The signal and noise voltages due to the optical signal are Vs and Vnrms, shown for each of three distances. The read, reset and total noise voltages due to the pixel circuitry are also shown, in black. These noise levels are independent of the fluorophore concentration. This plot relates the circuit noise to the optical signal noise. The circuit noise begins to dominate around a concentration of $0.1\mu M$ for the case where $R_{nl} = R_{pn} = 100\mu m$. For a given distance case, as C_{fl} increases, the optical signal increases faster than the optical noise, and the SNR improves (as was seen in Fig. 5.8). This plot also gives an indication of absolute signal levels, and what concentrations can be detected before saturating the pixel circuitry. In the $100\mu m$ distance case, for example, the imager will saturate around $0.3\mu M$.

5.4.2 Arch Model

The second model presented here is geared toward the case where a fast VSFP such as Arch is available to record individual action potentials. In this case the frame rate must be increased significantly from the GCaMP3 case. Table 5.3 shows the parameter values used in the model. The Arch parameters are from [54], except for Cfl, which is assumed to be the same as in [102]. The LED parameters are for the CPR3ED (but assuming a different excitation wavelength). Table 5.2 shows the neuron locations used in the model (same as GCaMP3 model).

Figs. 5.10, 5.11, 5.12, and 5.13 show the system geometry, the fluorescent power emitted by

Parameter	Name	Value
simulation time step	$tScale$	10 μs
sine wave snippet frequency	$Fspike$	1 kHz
imager sampling time	Ts	400 μs
neuron resolution	kN	2 μm
LED resolution	kL	20 μm
pixel fill factor	FF	0.83
CTIA feedback capacitance	Cp	4 fF
photodiode responsivity	$Resp$	0.29 A/W
fluorophore quantum yield	$QYfl$	0.0009
fluorophore emission wavelength	Lfl	687 nm
fluorophore extinction coefficient	$Exfl$	6300 L/(mol cm)
fluorophore concentration	Cfl	5 μM
fluorophore sensitivity	Sfl	65 %/100mV
LED excitation power	Pex	11 mW
LED excitation wavelength	Lex	558 nm
fluorescent wavelength absorption coefficient	aaf	0.08/1000 μm^{-1}
fluorescent wavelength scattering coefficient	asf	0.8/1000 μm^{-1}
LED wavelength absorption coefficient	aat	0.25/1000 μm^{-1}
LED wavelength scattering coefficient	ast	1.25/1000 μm^{-1}

Table 5.3: OR2 model parameters for the Arch case. The Arch parameters are from [54], except for Cfl, which is assumed to be the same as in [102]. The LED parameters are for the CPR3ED (but assuming a different excitation wavelength).

each neuron over time, the photocurrent in selected pixels over time, and the signal voltage in selected pixels over time. The plots look similar to the GCaMP3 case, but the key difference is the signal amplitudes. In this case all of the final voltage signals are less than 1mV, and mostly are tens of μV. This is largely due to the low extinction coefficient of Arch. These signals are too small to be detected without significantly increasing the excitation power. Fig. 5.14 plots the SNR (as in Eq. 5.21) vs. the fluorophore concentration, C_{fl}, for three different distance values. The trends are the same as for the GCaMP3 model, but here the SNR values are lower for the same fluorophore concentration and distances. Fig. 5.15 plots the signal and noise voltages vs. fluorophore concentration, C_{fl}, for three different distance values. Here the circuit noise dominates around all concentrations shown, and only for the case where $R_{nl} = R_{pn} = 100\mu m$ could we detect any signal above the noise, at high fluorophore concentration.

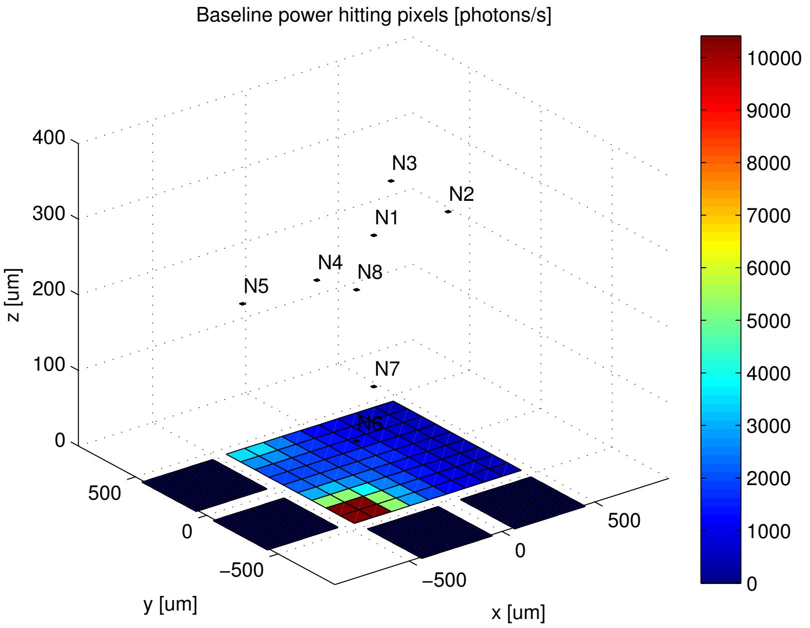


Figure 5.10: System plot of the OR2 model for Arch, showing pixel, LED, and neuron locations, as well as the baseline optical (fluorescent) signal power hitting the array.

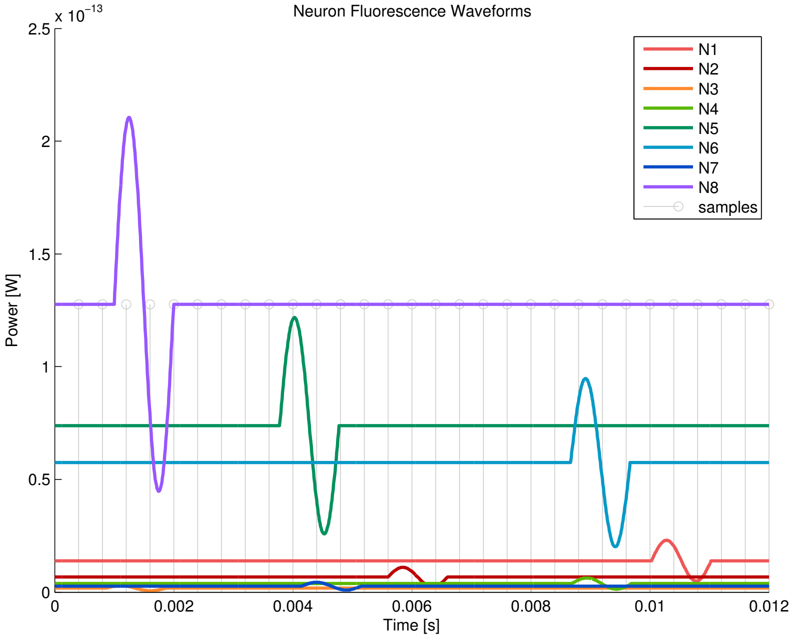


Figure 5.11: Plot of fluorescent output power for each neuron in the Arch model.

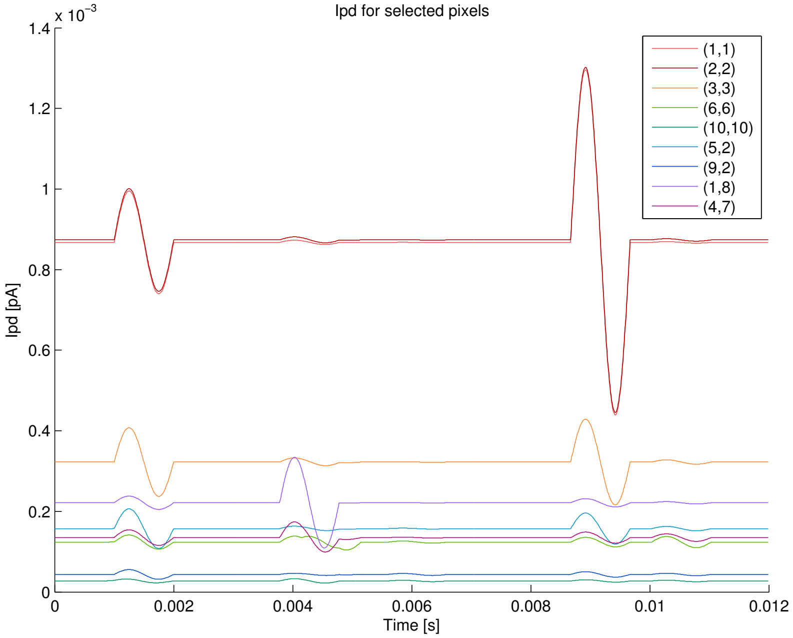


Figure 5.12: Photocurrent in selected pixels over time in the Arch model.

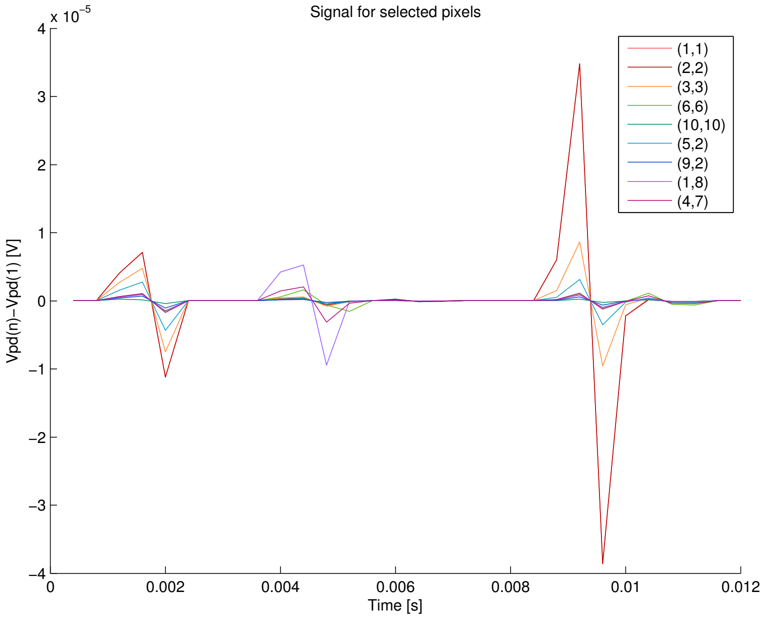


Figure 5.13: Sampled voltages in selected pixels over time in the Arch model.

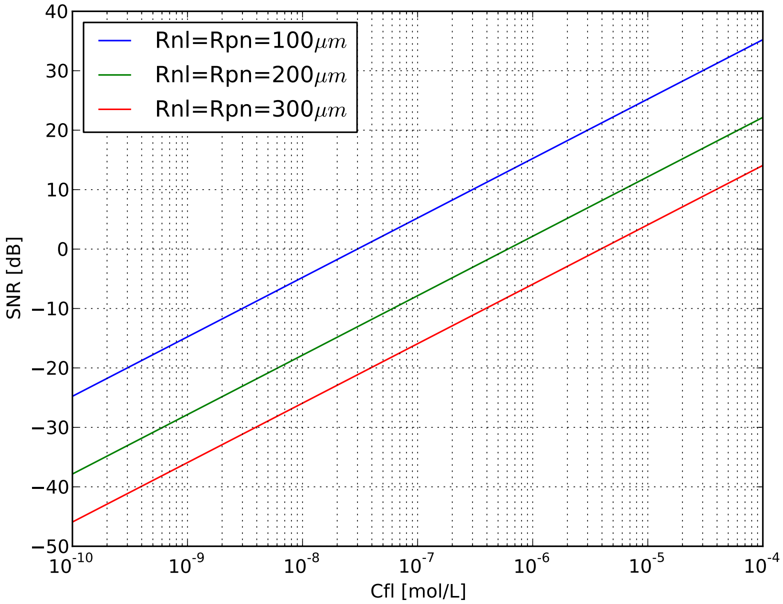


Figure 5.14: Plot of SNR vs. fluorophore concentration for the Arch model.

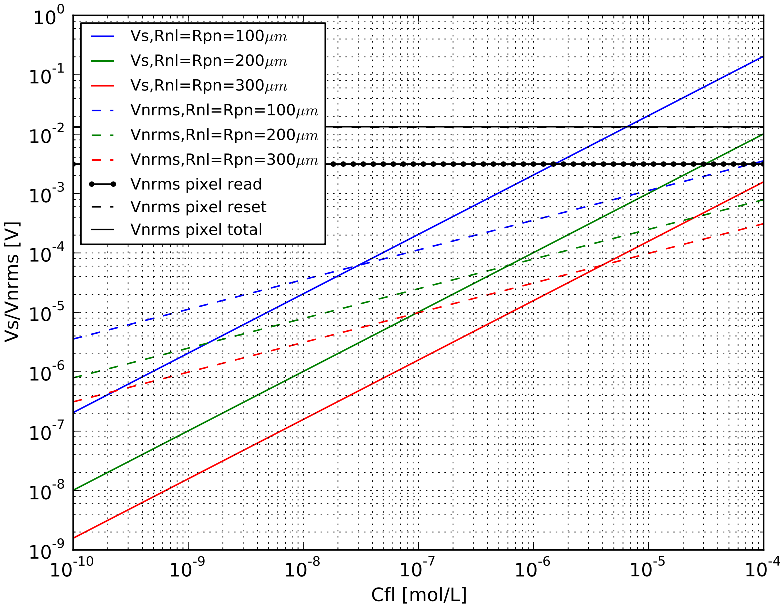


Figure 5.15: Plot of signal and noise voltages vs. fluorophore concentration for the Arch model. The signal and noise voltages due to the optical signal are Vs and Vnrms, shown for each of three distances. The read, reset and total noise voltages due to the pixel circuitry are also shown, in black.

5.4.3 Voltage Sensitive Dye Model

In order to demonstrate high frame rate operation, model results for a common voltage sensitive dye, di-4-ANEPPS, are shown. This main difference here compared to the Arch model is that the VSD has much higher quantum yield, resulting in larger signals. Table 5.4 shows the parameter values used in the model. The di-4-ANEPPS parameters are from [93] and [31], except for Cfl, which is assumed to be the same as in [102]. The quantum yield for this dye is not known exactly. A conservative value for quantum yield comparable to similar dyes as shown in [31] (with lipid binding) is used. Table 5.2 shows the neuron locations used in the model (same as GCaMP3 and

Arch models).

Parameter	Name	Value
simulation time step	$tScale$	10 μs
sine wave snippet frequency	$Fspike$	1 kHz
imager sampling time	Ts	400 μs
neuron resolution	kN	2 μm
LED resolution	kL	20 μm
pixel fill factor	FF	0.83
CTIA feedback capacitance	Cp	4 fF
photodiode responsivity	$Resp$	0.29 A/W
fluorophore quantum yield	$QYfl$	0.2
fluorophore emission wavelength	Lfl	617 nm
fluorophore extinction coefficient	$Exfl$	42000 L/(mol cm)
fluorophore concentration	Cfl	5 μM
fluorophore sensitivity	Sfl	10 %/100mV
LED excitation power	Pex	11 mW
LED excitation wavelength	Lex	475 nm
fluorescent wavelength absorption coefficient	aaf	0.13/1000 μm^{-1}
fluorescent wavelength scattering coefficient	asf	1.00/1000 μm^{-1}
LED wavelength absorption coefficient	aat	0.28/1000 μm^{-1}
LED wavelength scattering coefficient	ast	1.6/1000 μm^{-1}

Table 5.4: OR2 model parameters for di-4-ANEPPS. The di-4-ANEPPS parameters are from [93] and [31], except for Cfl, which is assumed to be the same as in [102]. The LED parameters are for the CPR3ED (but assuming a different excitation wavelength).

Figs. 5.16, 5.17, 5.18, and 5.19 show the system geometry, the fluorescent power emitted by each neuron over time, the photocurrent in selected pixels over time, and the signal voltage in selected pixels over time. The plots look similar to the Arch case, but the key difference is the signal amplitudes. In this case the final voltage signals can be higher than 10mV, which is much higher than for Arch. This is largely due to the higher quantum yield of the VSD. These signals are large enough to be detected. Fig. 5.20 plots the SNR (as in Eq. 5.21) vs. the fluorophore concentration, C_{fl}, for three different distance values. The trends are the same as for the GCaMP3 and Arch models. Here, for the same fluorophore concentration and distances, the SNR values are higher than for Arch. Fig. 5.21 plots the signal and noise voltages vs. fluorophore concentration, C_{fl}, for three different distance values. Here the circuit noise begins to dominate around a concentration of $1\mu M$ for the case where $R_{nl} = R_{pn} = 100\mu m$, and at this point the SNR is still relatively high.

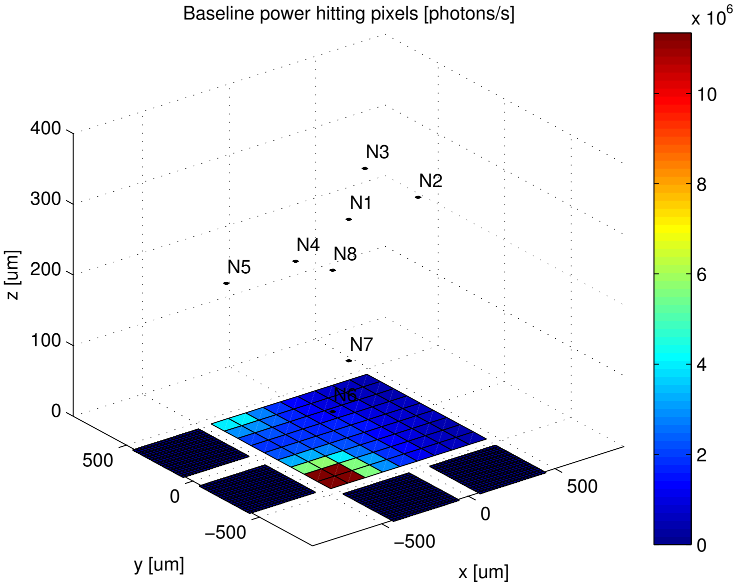


Figure 5.16: System plot of the OR2 model for di-4-ANEPPS, showing pixel, LED, and neuron locations, as well as the baseline optical (fluorescent) signal power hitting the array.

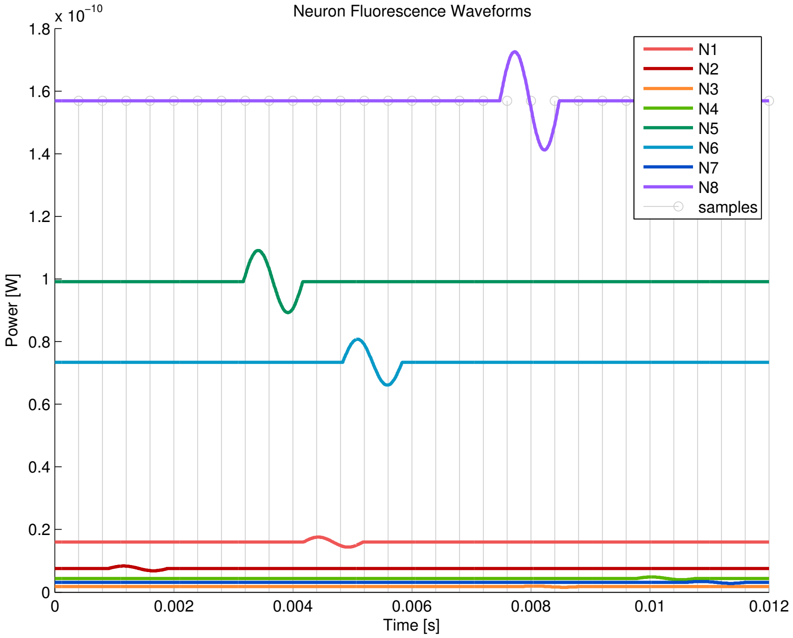


Figure 5.17: Plot of fluorescent output power for each neuron in the di-4-ANEPPS model.

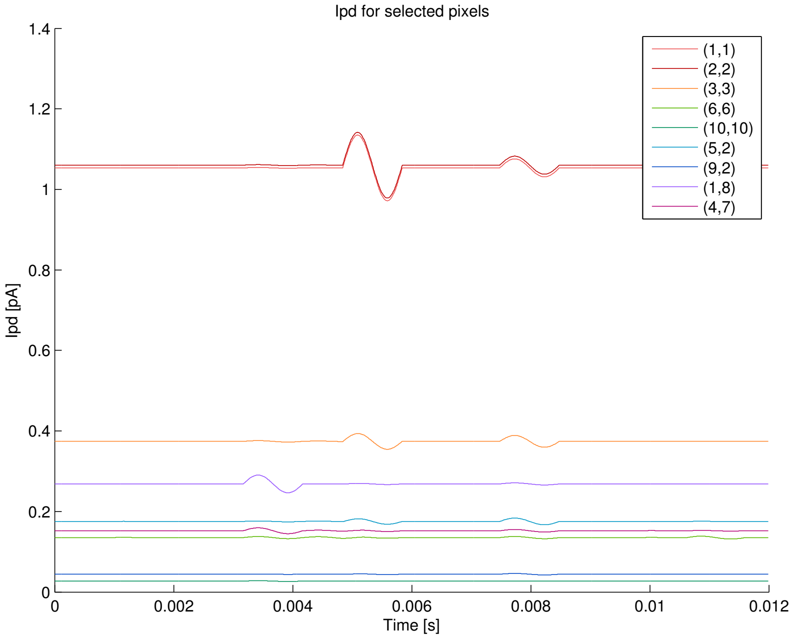


Figure 5.18: Photocurrent in selected pixels over time in the di-4-ANEPPS model.

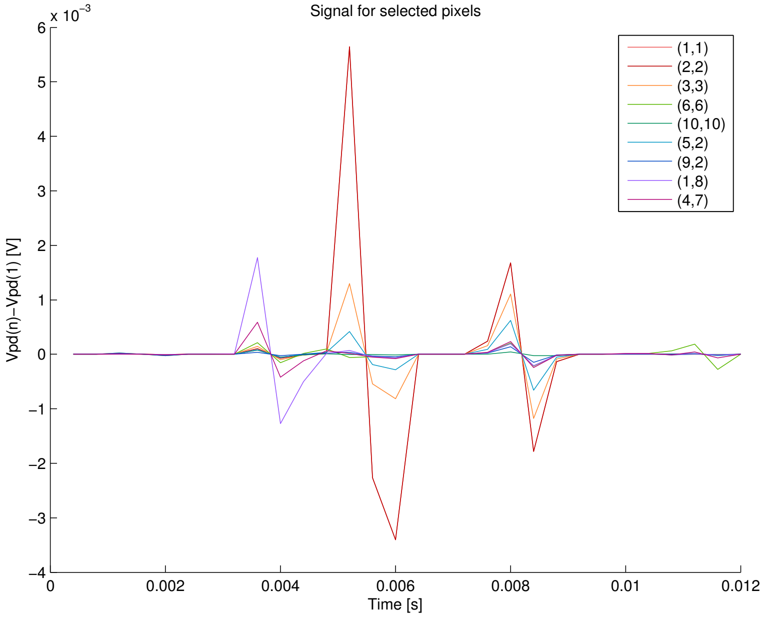


Figure 5.19: Sampled voltages in selected pixels over time in the di-4-ANEPPS model.

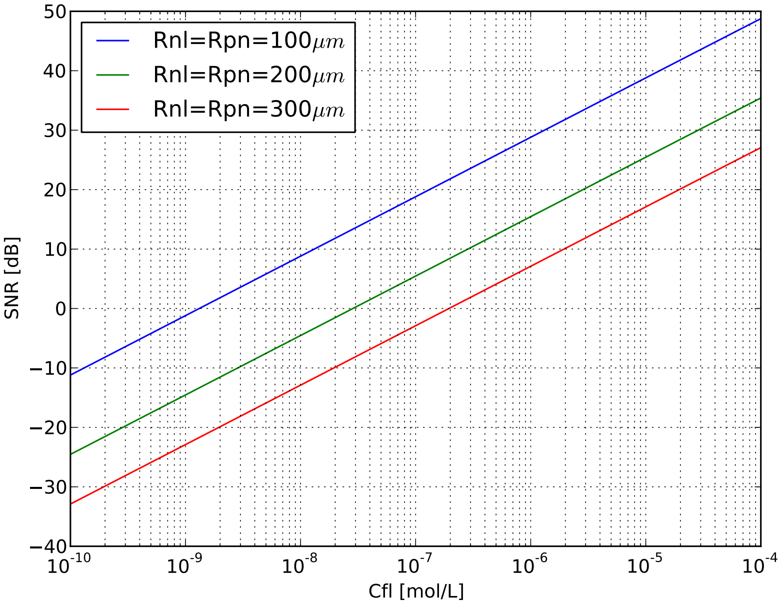


Figure 5.20: Plot of SNR vs. fluorophore concentration for the di-4-ANEPPS model.

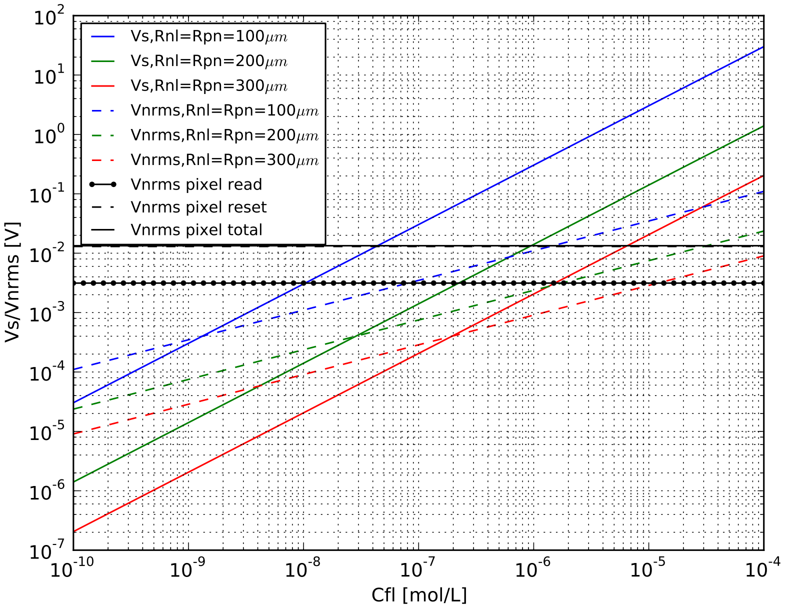


Figure 5.21: Plot of signal and noise voltages vs. fluorophore concentration for the di-4-ANEPPS model. The signal and noise voltages due to the optical signal are Vs and Vnrms, shown for each of three distances. The read, reset and total noise voltages due to the pixel circuitry are also shown, in black.

5.4.4 Discussion

In this work layer I (the top layer) of the brain is chosen as the target recording area. Layer I of the brain is approximately 300μm thick, and has on average 13000 neurons per mm^3 [99]. Therefore in a 1mm x 1mm area there would be about 4000 neurons in layer I. The characteristics of layer I neurons were studied in [43]. They found that neurons with axons descending down to layer II or deeper had somata (cell body, not including axons/dendrites) that were 14.9 x 7.9μm^2 on average. Consider the case of 100μm x 100μm pixels. That would mean there are about 40 neurons above each pixel. Given that resting neurons fire at up to 10Hz, and active neurons fire at up to 100Hz, we would expect only a couple of neurons (or fewer) above each pixel to fire at once. Now, if our VSFPs target only a fraction of the neuron population, we have a manageable number of neurons to track.

The modeling indicates that recording applications using a reporter like GCaMP3 could be immediate targets for the type of implantable recording device proposed in this work. Such devices may also work for detecting individual action potentials once a sufficient VSFP is created. Alternatively the device could be used with voltage sensitive dyes, keeping in mind that dyes are not appropriate for chronic recording. There are a few takeaways regarding choosing system design specifications:

- Pixels that are close together usually detect a given spike. For this reason in an imager that is only concerned with spike detection, smaller pixels are not particularly advantageous; in fact smaller pixels have less area to collect photons and produce lower signal levels, making spike detection more difficult. Larger pixels produce larger signal levels, but at the expense of spatial resolution. Larger pixels also allow for more circuitry to be included in the pixel. This is an important tradeoff in these systems. The general takeaway is that the physical size of neurons and neuron density can dictate the appropriate pixel size. Having pixels smaller than neuron diameters would not be useful. Having pixels the same size as neurons may also not be a good choice because a cluster of neighboring pixels will still see nearly the same signal.
- For simply detecting spikes rather than recording their shape, it is useful to reduce the frame frate from the Nyquist rate. The longer integration time allows for the collection of more photons, resulting in a larger signal. This is especially useful in applications where

the signals can be very weak.

- For the Lambertian LED dies there is a sweet spot for neuron height above the imager surface where the recorded signal amplitude is maximized. When neurons are too close to the imager surface, they see very little excitation light, and thus have a weak signal. When neurons are too high above the imager surface, they see more excitation light, but the spatial resolution decreases, since the signal can be seen across the entire array.

5.5 Summary

This chapter has added tissue attenuation and fluorescence to the optical model of Chapter 4, and applied the model to a complete implantable optical neural recording device design. The model output consists of the sampled voltage in each pixel over time, and shows the detection of spikes from different neurons. The model results are used to arrive at specifications for the device design; these are discussed in the next chapter.

Chapter 6

Second Optical Chip - OR2

The OR1 chip was designed to verify the optical modeling. A second optical chip was designed for use in an actual optical neural recording device. The second chip (as well as the overall system design) is referred to as OR2. The goals for the OR2 design were to:

- Produce a full fluorescence detection imaging system, including optical filtering, which is capable of detecting individual action potentials in the top layer (layer I) of brain tissue; this includes significantly higher sensitivity pixel circuitry
- Include on-chip analog-to-digital conversion of the imager output
- Test the system with a fluorescent tissue phantom

In this chapter the OR2 system design and packaging is discussed, followed by the circuit design, and finally the testing results.

6.1 System Design and Packaging

There are two main packaging options with regard to testability. One option is to do a system-in-package (SiP), where the chip and LED dies are bonded inside a package (in this case an 80-pin ceramic quad flat pack (CQFP80)) with the lid left off. The other option is to bond the chip and LED dies directly to a PCB. The SiP approach was taken here. The main advantage of the SiP approach is that it makes it easy to test multiple chips on a PCB, because it is easy to solder/desolder packages to a board and bonding is no longer a concern. If the device were to be tested *in vivo* then direct-to-PCB bonding may be a better option, as package walls would not prevent the chip surface from contacting the tissue.

6.1.1 System Components and Geometry

Fig. 6.1 shows a drawing of the SiP design (the chip and LEDs are close to scale, the surrounding package and bonding is not). The OR2 chip is 3mm x 3mm. The pixel array is 1mm x 1mm, and is located in the southeast corner of the chip. The goal was to have the array as close as possible to the LEDs to ensure that as much excitation light as possible could hit the tissue above the array. Therefore, no bondpads were included in that corner. Four LED dies surround the array. The original plan was to have the imaging chip 1mm x 5mm, with the pixel array in the center and two LED dies on either side of the chip. This design would have provided a much more uniform illumination across the entire array. However, it was advised in design review that such a narrow chip may have poor yield in fabrication, so the square chip was chosen. The square design also makes placement and routing, and bonding, much easier. On top of the pixel array is a thin optical filter, which is discussed in Section 6.1.5.

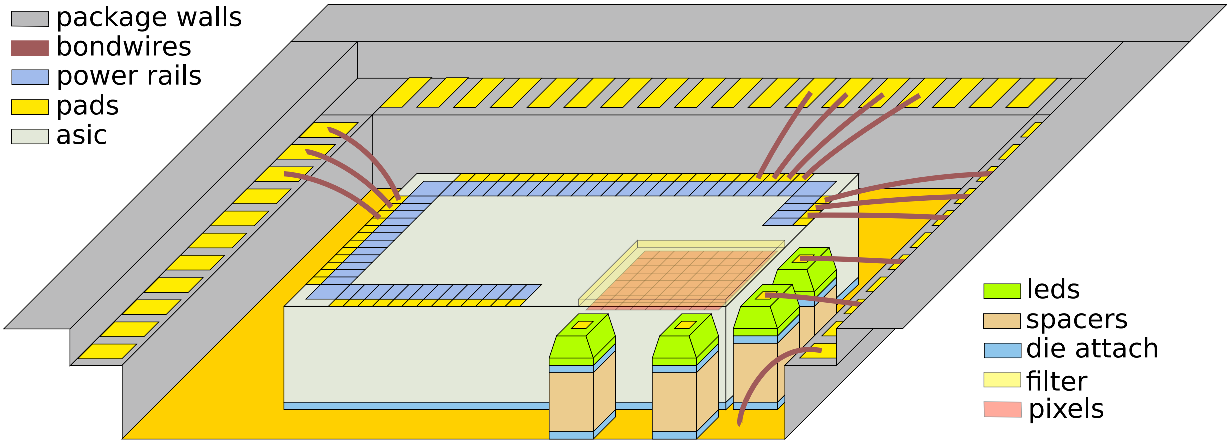


Figure 6.1: OR2 system diagram. The imager chip sits in a QFP80 package. The pixel array is in the bottom right corner of the imager. Four LED dies sit adjacent to the pixel array. The LED dies are placed on top of die spacers to match their height to the surface of the imager chip. An optical filter is attached above the pixel array.

6.1.2 Controlling Die Height

Based on the work of Chapter 4, we saw that when the LED junction is above the imager height, excitation light can directly hit the imager. A filter would limit the significance of that light. But regardless, we may want the LED junction to be below the imager height. However, a concern with having the LEDs below the surface of the imager is that LED light could still directly hit the pixels through the side of the chip where there is no filter. Therefore, the optical characteristics of silicon must be considered.

The fraction of incident light reaching a given depth into the silicon is given by the Beer-Lambert law (see Section 5.1). At 570nm, which is the high end of excitation wavelengths in this system, and the worst case for this calculation, the absorption coefficient of silicon is 751286m^{-1} [16]. Table 6.1 shows the depth of light penetration for a range of transmission (T) values. For example, 0.1% of light gets through 9.2μm. This result indicates that in terms of excitation light

T	$d\ [\mu m]$
0.100000	3.1
0.010000	6.1
0.001000	9.2
0.000100	12.3
0.000010	15.3
0.000001	18.4

Table 6.1: Depth of light penetration in silicon.

hitting the pixels, it is not an issue to have the LED below the pixel surface because any light that hits the side of the chip is absorbed before it reaches the pixels (assuming the pixels are not right up against the edge of the chip). On the other hand, this also means that light that hits the side of the imaging chip will not make it to the tissue (at least not directly) either, and is thus not useful for excitation. Therefore, this means that the height of the LED dies relative to the imager surface is a very important system parameter. Ideally the LED die junctions would be at the same height (or slightly below) as the imager chip surface.

In the case of Lambertian LED dies with restricted beam angles, there is a tradeoff between the height of the dies relative to the imager surface and the height the tissue must be at to have excitation light hit it. If the LED die height is lower than the chip surface, the widest angle excitation light is blocked from hitting the tissue by the imager chip, and thus the tissue would have to be higher. Assuming there were LED dies on both (opposite) sides of the imager, the width of the array (and the chip) will strongly impact how much light can hit tissue at a given height over the center of the array. And recall that the impact of moving tissue higher (further from the array) is reduced imaging spatial resolution and a weaker signal.

It is possible to use die spacers to adjust die height during bonding. This means there can be some control of LED die height relative to the imager height. The specification for the packaging house was to have the top surface of the LED dies (where the junction is in this case) at the same height as the imager, or slightly below (we wish to err on the lower side). They must consider their margin for error when selecting the die spacer thickness.

6.1.3 Cavity Potential

In the CPR3ED dies the bottom contact, which gets bonded to the CQFP80 package cavity (through the spacers, which must be conductive), is the positive terminal of the LED, and must be connected to VDD. However, the imager chip substrate, which gets bonded to the same cavity, must be GND. Therefore, down-bonds are used to connect the package cavity to VDD, and a non-conductive epoxy is used to bond the imager chip to the cavity.

6.1.4 Encapsulation

For an implantable device encapsulation is required to protect the device from the biological environment. In the bench testing environment encapsulation of the bondwires is still desirable to prevent any damage from accidental contact between the bondwires and the optical filter or fluorescent sample during handling.

In this case, however, encapsulation was not included for the bench testing, due to difficulties with creating a moat around the imager surface (which is lower than the bondwires), where the encapsulant would need to be very thin (to keep the tissue close to the imager). Future system designs should account for this.

6.1.5 Optical Filter

Fluorescence detection systems typically require an optical filter to prevent excitation light from hitting the photodetectors. In this system where the Lambertian LEDs (at least in theory) emit upwards, no excitation light would hit the photodetectors without reflections (off of tissue). At this stage, the optical modeling does not model reflections, so it isn't clear how much excitation light would be reflected back down to the photodetectors. The system geometry is designed to limit excitation light hitting the photodetectors, but the contribution is likely not zero. Therefore an optical filter is included in the system.

The CPR3ED datasheet does not include a plot of the emission spectrum, but it states that the dies have a full width half max of 35nm. A wavelength shift of about 20nm (extrapolated from the datasheet curve) is assumed because this system uses very low current. Therefore the LED spectrum is estimated to have a peak wavelength around 547nm ± 15nm to the half max values.

The fluorophores used in the tissue phantom (described in more detail in Section 6.3.2) are orange fluorescent microspheres from Cospheric [21]. On the webpage for the microspheres emission spectra for various excitation wavelengths are shown. The white excitation curve is assumed here. Emission wavelengths are all higher than about 570nm.

Based on the excitation and emission wavelengths, the optical filter should have a high-pass cutoff frequency of about 565nm. The optical density specification for a filter cannot be determined accurately as explained above. Higher optical density is better, as a filter with a higher optical density will reject a higher fraction of incident light. An important requirement of the filter is the thickness - it should be thin enough to keep the fluorescent sample close to the imager surface, but not too thin, otherwise fluorophores at the imager surface will see very little

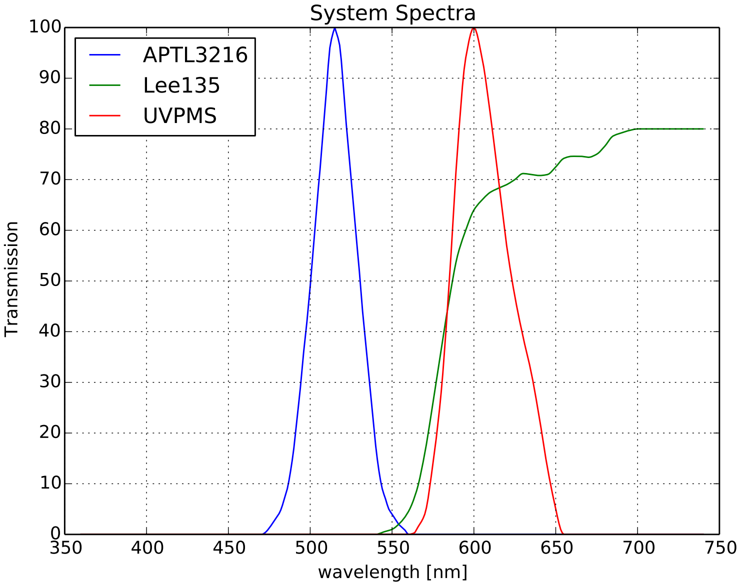


Figure 6.2: Plot of LED (APTL3216), filter, and fluorophore spectra.

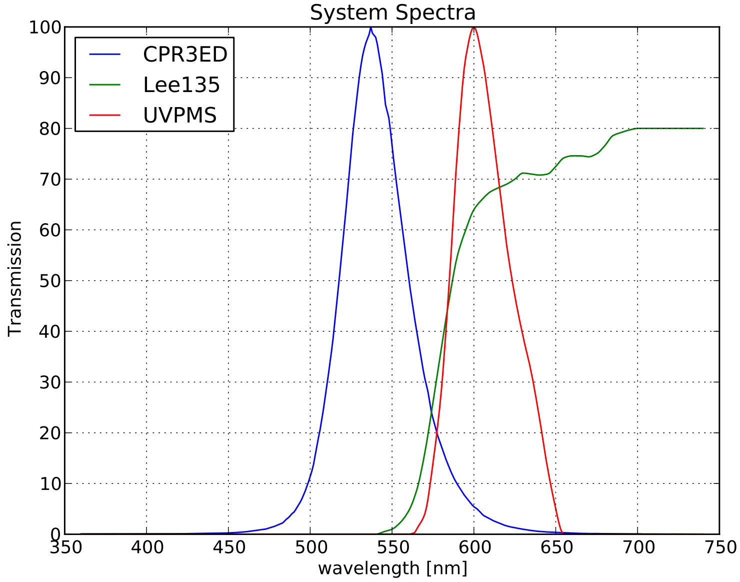


Figure 6.3: Plot of LED (CPR3ED), filter, and fluorophore spectra.

excitation light. According to the optical modeling, the desired filter thickness is around $100\mu m$.

A custom filter design was considered for this work. However, it turns out that a much cheaper alternative is to use “theatre gels”, or “color gels”. These are color filters that are placed over lighting in theatre productions to get desired lighting effects. These filters are sold by several companies, and a single sheet close to 2ft x 2ft can be purchased for less than \$10. The filter chosen for this work is 135 Deep Golden Amber from Lee Filters [30]. It is an absorption filter, made from a thin surface coated PET (Polyester Film), and is $80\mu m$ thick.

Fig. 6.2 shows a plot of the filter transmission spectrum (Lee135), along with the fluorescent microsphere emission spectrum (UVPMS) and a test LED emission spectrum (APTL3216). Fig. 6.3 shows a plot of the filter transmission spectrum (Lee135), along with the fluorescent microsphere emission spectrum (UVPMS) and the LED emission spectrum for the CPR3ED, which is used in the actual system for fluorescence testing (not for imager characterization).

6.1.6 Final Packaged Device

The final packaging of the device, including die-attach, bonding, and filter attach, was done by Quik Pak [73], facilitated by CMC Microsystems. Fig. 6.4 shows one of the final packaged devices. Fig. 6.5 shows a zoomed view of the pixel array with filter and LEDs. From these pictures a fabrication/packaging issue is apparent: the OR2 chips were diced larger than they should have been. The chip edge should have been at the edge of the pad frame, but instead is around $400\mu m$ away. The result of this is that the LED dies are about $400\mu m$ further away from the pixel array than they should be. This will be discussed further in Sec. 6.5.

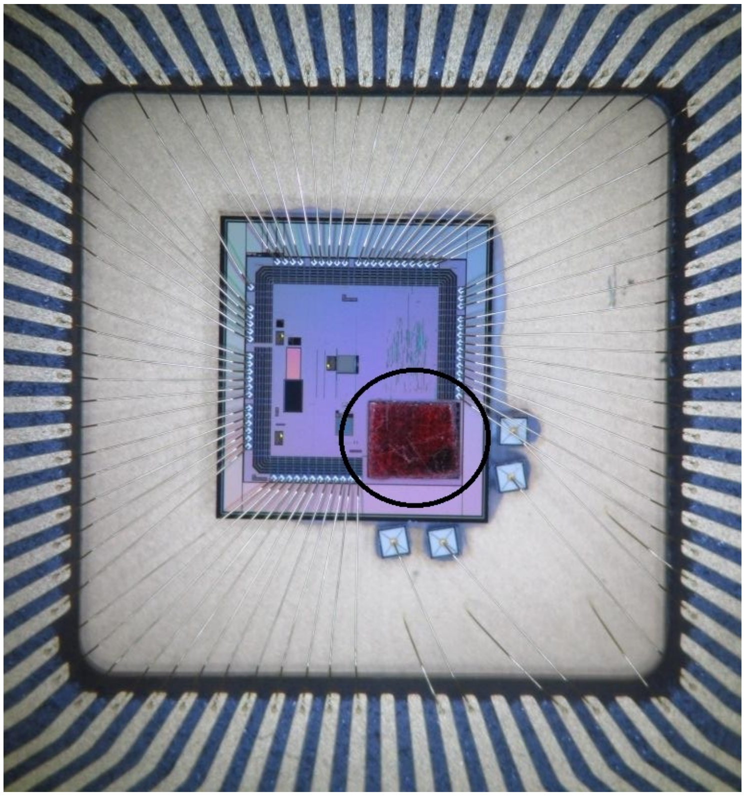


Figure 6.4: Photo of final packaged device. The red absorption filter sits on top of the pixel array. The LED dies sit adjacent to the array. Four downbonds for VDD are seen in the bottom right corner of the package.

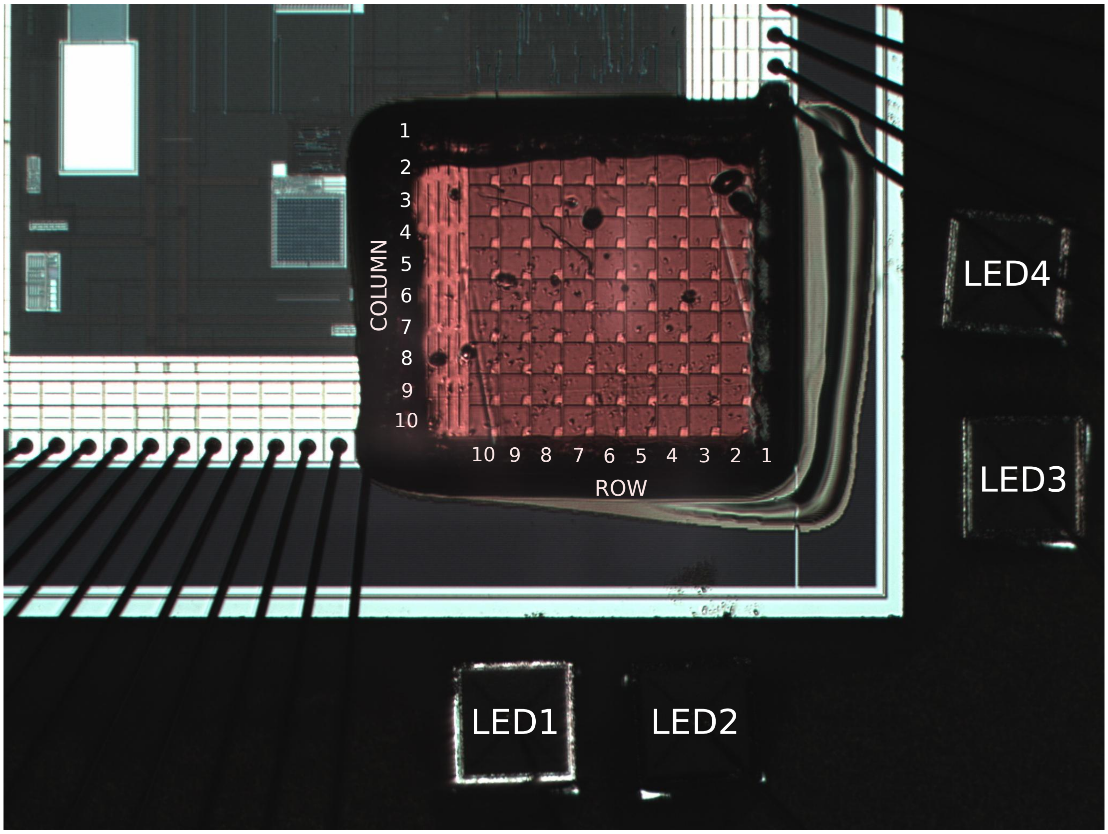


Figure 6.5: OR2 microphotograph zoomed in to the imager and LED dies. The imager rows and columns are labeled. The wide dicing of the chip is apparent. The edges of the LED dies (closest to the chip) are still about 150-200μm from the chip.

6.2 Imaging Circuitry Design

As discussed in the modeling section 5.4, the chip is designed based on a new system design where LED dies are bonded beside the chip instead of bonded on top of the chip. Thus, OR2 does not include the large bondpads for bonding LED dies to the top of the chip. Compared to OR1, OR2 includes much higher sensistivity imaging circuitry, with in-pixel CTIAs, and programmable gain (individually controlled for each pixel).

6.2.1 Imager Architecture and Pixel Circuitry

Based on the MATLAB model results, the following main imaging specifications were chosen:

- 10x10 pixel array with 100x100μm^2 pixels; adjacent smaller pixels see nearly the same signal, so it is worthwhile to increase pixel size, which allows each pixel to collect more photons and produce a larger signal

- Up to 10kHz frame rate (but also programmable); 2.5kHz is the ideal target speed based on modeling of Arch and di-4-ANEPPS, from the point of view of maximizing peak signal levels. Frame rates lower than 2.5kHz may miss spikes, while higher frame rates reduce signal strength. In order to work with slower responding reporters like GCaMP3, the frame rate can be significantly reduced.

- 1mV resolution; This strikes a balance between higher resolution required to detect weaker signals (or signals deeper in tissue) and the technical difficulties with designing high resolution data converters.

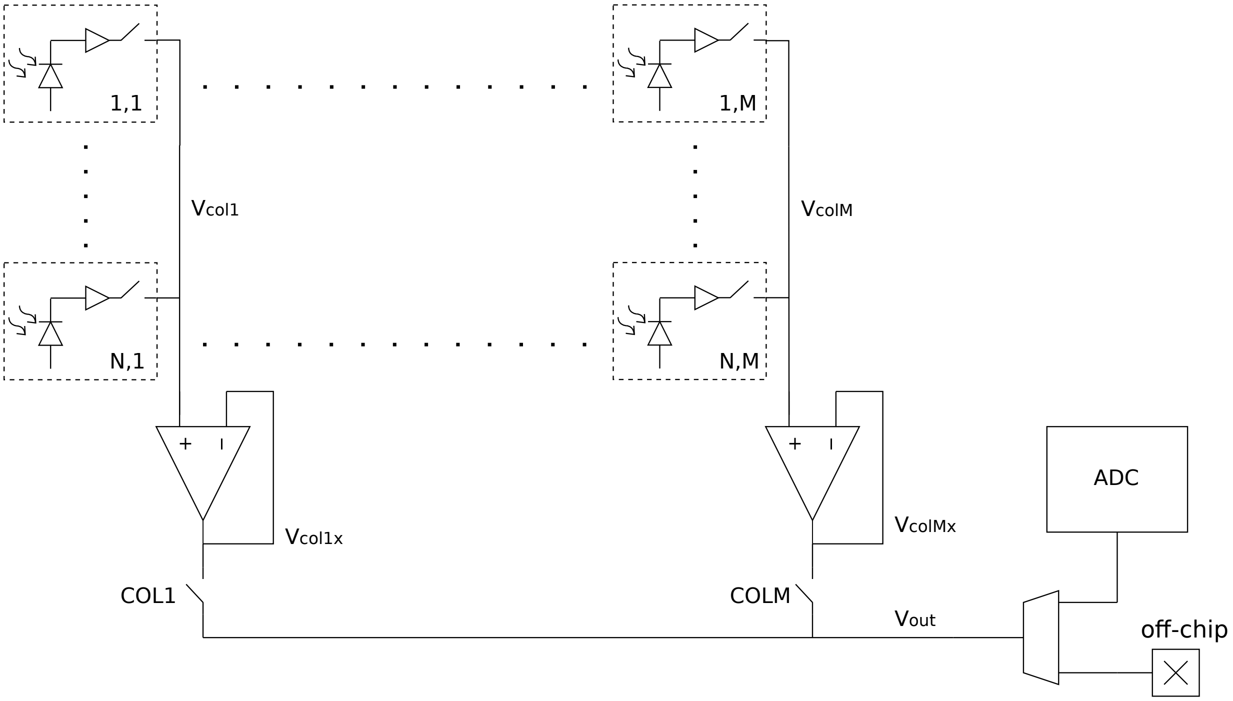


Figure 6.6: Architecture of the image sensor. Pixel outputs connect to the column line through row select switches. Each column line is buffered with an op-amp. The op-amp outputs are multiplexed to the on-chip ADC or off chip.

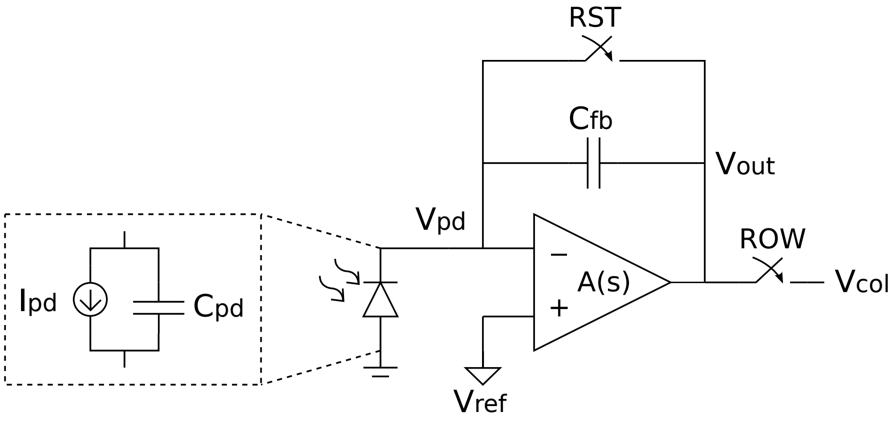


Figure 6.7: CTIA pixel circuit. The photocurrent is integrated onto the feedback capacitor once the RST switch is opened.

- Programmable gain; This is desirable because signal strength can vary greatly across the pixel array. Neurons that are close to a light source and close to the array could saturate pixels. Therefore it is useful to be able to reduce the gain in individual pixels.

The overall architecture of the imager is shown in Fig. 6.6. Each column has an op-amp buffer before the column MUX switches. These op-amps are strong enough to drive the signal off chip. The column MUX drives the output signal through the imager MUX, which sends it either off chip or to the on-chip ADC. The ADC block includes the sample and hold circuit.

Fig. 6.7 shows the basic pixel circuit, which is a CTIA. The transfer function of the circuit is given by:

$$V_{out}(s) = \frac{I_{pd}(s)}{sC_{fb}\left(1 - \frac{1}{A(s)} - \frac{C_{pd}}{C_{fb}A(s)}\right)} \tag{6.1}$$

or

$$V_{out}(s) = \frac{I_{pd}(s)}{sC_{fb}\left(1 - \frac{1}{A(s)}\left(1 - \frac{C_{pd}}{C_{fb}}\right)\right)} \tag{6.2}$$

Assuming $A \gg Cpd/Cfb$ and $Cpd \gg Cfb$, the transfer function of the circuit is

$$V_{out}(s) = \frac{I_{pd}}{sC_{fb}} \ , \tag{6.3}$$

which in the time domain is

$$v_{out}(t) = \frac{1}{Cfb}\int i_{pd}dt \ . \tag{6.4}$$

The actual implemented pixel circuit is shown in Fig. 6.8. Table 6.3 shows device sizes for the pixel transistors (other than the SRAM, which is shown in Fig. 6.10, and the amplifier

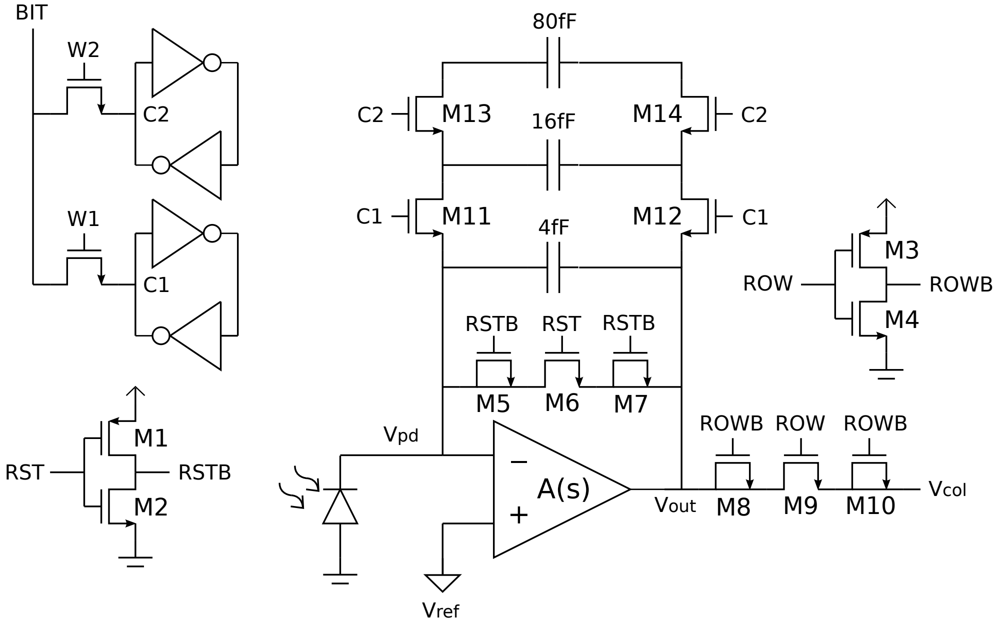


Figure 6.8: Final CTIA pixel schematic. The reset and row select switches are NMOS devices with a dummy device on each side. The main feedback capacitor is 4fF. Additional 16fF and 80fF capacitors can be switched in as desired. The capacitor switches C1 and C2 are programmed by the two SRAM cells that are included in the pixel.

	M1	M2	M3	M4	M5	M6	M7
W	0.8	0.4	0.8	0.4	0.55	1	0.45
L	0.35	0.35	0.35	0.35	0.35	0.35	0.35
	M8	M9	M10	M11	M12	M13	M14
W	0.5	1	0.5	0.5	0.5	0.5	0.5
L	0.35	0.35	0.35	0.35	0.35	0.35	0.35

Table 6.2: Pixel external device sizing (all sizes in μm).

internals).

The pixel amplifier is designed to meet the following key specifications (assuming one LSB is 1mV and the supply voltage is 3.3V):

- The rising slew rate and settling time must ensure that the output has settled to within 0.5 LSB within the sample time (400μs for the target frame rate). The settling is related to the gain-bandwidth product (GBW) of the amplifier. Assuming a single-pole step response, the amplifier output voltage can be written as

  $$V_{out}(t) = V_0(1 - e^{t/\tau}) \ , \tag{6.5}$$

  where $\tau = 1/GBW$. Assuming 90% of the sample time is allocated to linear settling, we have:

  $$\frac{3.3 - 0.0005}{3.3} = 1 - e^{-0.9*400*10^{-6}*GBW} \ . \tag{6.6}$$

  This results in a minimum GBW of 24.5kHz. Assuming 10% of the sample time is allocated to slewing, we have:

  $$i = C\frac{dV}{dt} = 1 * 10^{-12} * \frac{3.3}{0.1 * 400 * 10^{-6}} \ , \tag{6.7}$$

  resulting in a required output current of $0.083\mu A$.

- The falling slew rate and settling time must ensure the output has settled to within 0.5 LSB of the reset level within the reset time (assume 1% of the target frame rate). Assuming a single-pole step response, the amplifier output voltage can be written as

$$V_{out}(t) = V_0 e^{t/\tau} \ , \tag{6.8}$$

where $\tau = 1/GBW$. Assuming 90% of the reset time is allocated to linear settling, we have:

$$0.7 = 3.3e^{-0.9*4*10^{-6}*GBW} \ . \tag{6.9}$$

This results in a minimum GBW of 431kHz. During reset, the output of the pixel is disconnected from the column line, and the output load is much smaller (0.1pF). Assuming 10% of the sample time is allocated to slewing, we have:

$$i = C\frac{dV}{dt} = 0.1 * 10^{-12} * \frac{3.3}{0.1 * 4 * 10^{-6}} \ , \tag{6.10}$$

resulting in a required output current of $0.83\mu A$.

- DC gain must be large enough such that the amplifier achieves a steady-state error less than 0.5 LSB. During reset, the amplifier is in unity-gain configuration. We require the error voltage to be less than 0.5 LSB:

  $$\frac{1}{1 + A_0} < 0.0005 \tag{6.11}$$

  resulting in a minimum DC gain of $1999V/V$.

Note that these numbers are conservative since they assume that the output swings rail to rail. In this case, the DC gain is the hardest to achieve. A single transistor (differential pair) would not provide the necessary gain. Therefore, the differential pair was cascoded. This came at the cost of a reasonably small amount of pixel area (which isn't much of an issue given the large pixel size), but no increase in noise (since the noise contribution of cascode devices is negligible). In order to reduce noise, the input pair device sizes were increased.

Fig. 6.9 shows the transistor-level telescopic amplifier. Table 6.2 shows the amplifier device sizes. Table 6.4 shows the simulated characteristics of the telescopic amplifier alone (without the rest of the pixel circuitry), with a 1 pF load (approximate column capacitance and column amplifier input capacitance). Transient characteristics are found from the step response (step in output swing range) in the unity-gain configuration, with 1 pF load.

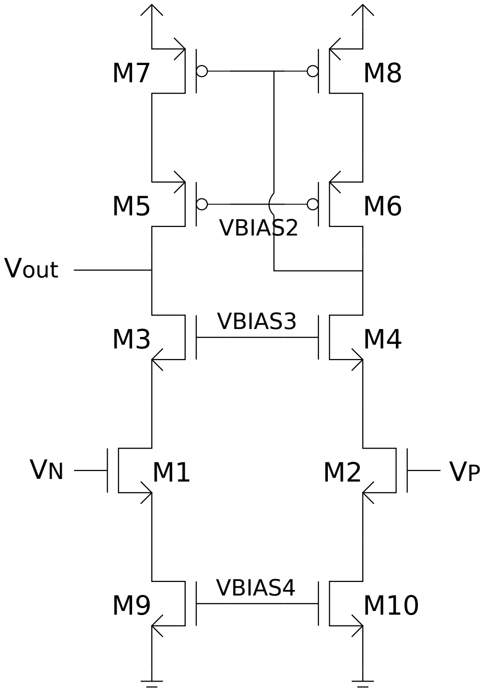


Figure 6.9: Pixel telescopic amplifier schematic.

	M1	M3	M5	M7	M9
W	32	1	2	2	1.4
L	1	1	1	1	1

Table 6.3: Pixel amplifier device sizing (all sizes in μm).

Spec.	Value	Units
DC Gain	94.4	dB
3-dB Freq.	101	Hz
Unity-gain Freq.	5.5	MHz
Phase Margin	86	degrees
Input Capacitance	80	fF
Input Referred Noise[a]	4.8	μV_{rms}
Output Resistance[b]	47.3	MΩ
Input Offset[c]	320	μV
Static Power[d]	9.4	μW
$V_{out,min}$	0.7	V
$V_{out,max}$	2.4	V
Slew rate (rising)	2.891	$V/\mu s$
Slew rate (falling)	2.985	$V/\mu s$
Settling time (rising)	209	μs
Settling time (falling)	0.645	μs

[a] Integrated from 0.1 to 158 Hz (noise equivalent bandwidth).

[b] At 10 kHz, a 20 mV_{pp} voltage into the output produced $0.423 nA_{pp}$ current.

[c] Systematic offset.

[d] 2.852 μA at 3.3 V.

Table 6.4: Simulated characteristics of the pixel telescopic amplifier, with 1pF load and 0.7V DC input level. Transient characteristics are found from the step response (step in output swing range) in the unity-gain configuration, with 1 pF load.

Photodiodes

The photodiodes are NW-PSUB diodes (same as OR1).

Reset Switch

The non-ideal effects of the reset switch can have a significant impact on the design. The high-impedance Vpd node is very sensitive, and any change there gets amplified at the output. Thus any charge injection can cause a significant jump in the output voltage when the switch is turned off. Adding dummy switches to collect the charge helps. The downside is that complementary clocks are now needed, which requires an inverter and extra area in the pixel.

Another option would be to use transmission gate switches instead of dummy switches. However, a transmission gate requires a PMOS device, which requires an N-well in the layout, and thus a lot more area. Also, PMOS and NMOS devices do not track with process variation; the dummy switch uses all NMOS devices.

Variable C_{fb}

The imager circuitry must be very sensitive to detect weak signals from neurons farther away from the device. This means the feedback capacitance of the CTIA must be very small to produce a larger voltage from small photocurrents. However, for neurons close to the device the signals could be significantly stronger, and having such a small feedback capacitor might mean that the pixel would saturate for closer neurons. For this reason, each pixel has a variable C_{fb}, which is individually controllable for each pixel, as shown in Fig. 6.8.

The feedback capacitors are poly capacitors; in the $0.35\mu m$ technology, cpoly has a minimum width of $0.8\mu m$, and leakage of $1aA/\mu m^2$. Because of the requirements for small C_{fb}, the capacitors will be very small, which means capacitor matching is not very good. In the chosen technology, based on the provided matching information, $1\mu m$ x $1\mu m$ capacitors that were close together would have a 0.45% standard deviation. The value for $10\mu m$ x $10\mu m$ capacitors would be to 0.045%. However, according to the plots in the documentation, the smallest capacitors that were measured for mismatch were $5\mu m$ x $5\mu m$, and at that small size it looked like the matching was worse than what the equations indicate. In this design the devices are not close together (matching between pixels) so matching may be quite poor regardless. However, in this application matching between pixels is not particularly important. Because the device is trying to simply detect action potentials, the important output is the change in voltage in a given pixel from frame to frame. The three capacitors used here are shown in Table 6.5. These values were chosen such that switching in each larger capacitor increases C_{fb} by approximately 5x (total C_{fb} settings of approximately C_{min} = 4fF, C_{mid} = 20fF, C_{max} = 100fF). C_{min} is close to the smallest capacitor that can be designed in the technology.

Name	W [μm]	L [μm]	C [fF]
C_{min}	2	2	4.128
C_{mid}	2	8	15.480
C_{max}	2	42.5	80.754

Table 6.5: Pixel feedback capacitor values. Switching in the next largest capacitor increases the total capacitance by a factor of 5.

Programming the Variable C_{fb}

Considering the pixel layout, it would be relatively easy to make the capacitor variable if all pixels were to have the same C_{fb} value (at the same time); it would be a matter of adding two more wires to the pixel - basically an extra “row select" line for each set of switches between adjacent capacitors. However, because different pixels could see significantly different light levels, depending on neuron depths, each pixel should have independently controlled C_{fb}. Having separate wires going to each pixel to control each individual set of switches would require too much area. Instead, each pixel needs to store its own bit values. One way to do this would be to have flip-flops in the pixels. All of the flip-flops in the array could be connected as a serial shift register so that only one wire line would be needed in the pixel. The problem with this method is that flip-flops are quite large in the $0.35\mu m$ technology and would reduce the fill factor significantly. The other way to store the information is to have RAM (random access memory) cells in the pixels. Single-transistor dynamic RAM (DRAM) would take up very little area, but it must be refreshed. Static RAM (SRAM) is non-volatile and doesn’t have dynamic power consumption. For these reasons SRAM was chosen in this case. Once the bit values are set they will rarely, if ever, have to change, so the dynamic power consumption of DRAM would be especially wasteful. It is also best to avoid digital signals switching in the pixels, as coupling of such signals can corrupt the analog signal of interest.

SRAM Design

There are two SRAM cells in each pixel - one for each extra feedback capacitor. Typical SRAM cells have six transistors (6T). It is not actually necessary to have two bitlines. Two bitlines would improve noise margins, and make reads easier, but here there is no need for read operations. Therefore, one bitline is used, which means only 5 transistors per cell are needed, as shown in Fig. 6.10. This saves layout area. All minimum-sized devices (NMOS and PMOS) were used for the inverters (0.4μm width). The key is that the pass transistor and drive strength on the bitline must be able to overpower the inverters. Monte Carlo simulation shows that with a 1.2μm pass transistor, 100 of 100 cells functioned correctly.

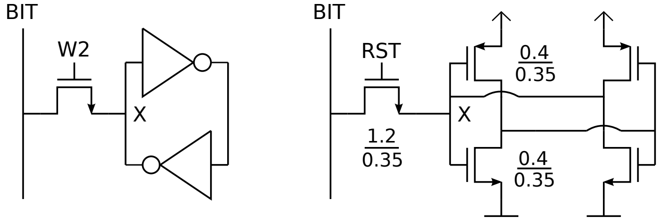


Figure 6.10: 5T SRAM schematic. The pass transistor is sized strong enough to overpower the inverters.

Layout

Fig. 6.11 shows the full pixel layout. Fig. 6.12 zooms in on the main circuitry (top metal (MET4) is removed from the picture). Note that the feedback capacitors extend to the left underneath the wires along the top of the pixel. The fill factor is 83.3%. The following are considerations for the pixel layout:

- All of the circuitry besides the photodiode was covered in MET4 to ensure that any light hitting the pixel circuitry does not affect the biasing.
- Power rails of 1μm width were used. Given the pixel bias current, this results in approximately a 20mV voltage drop from the first to last row of the array.

- Double-vias were used for reliability, with minimal impact on fill factor.

- The poly capacitors used for the feedback do not have guard rings. This should be acceptable because the pixel circuitry should be very quiet during integration, when no digital signals in the array are switching.

- Transistors were positioned such that N-wells were kept away from the perimeter of the pixel (on the right and top sides) to minimize how much they cut into the photodiode of the adjacent pixel.

- Even very small parasitic capacitances have to considered in the layout, especially parasitic capacitance to the sensitive Vpd node. Care was taken to minimize capacitance between Vpd and the digital nets, and to try to equalize the capacitance between Vpd and complementary digital nets (like RST and RSTB). Crosstalk due to a parasitic capacitance can be described by:

  $$Vx(t) = \left(\frac{C_c}{C_c + C_{tot}}\right) V_{DD} e^{-t/\tau} u(t) \tag{6.12}$$

  where C_c is the coupling capacitance between the agressor net (such as a digital net), C_{tot} is the total capacitance on the victim net (such as the sensitive Vpd node), and $u(t)$ is the unit step function. If the only concern is the final voltage, then the time constant and step function can be neglected. For example, a 10fF parasitic capacitance between a digital net and Vpd (assuming 1pF on Vpd) gives:

  $$\frac{0.10e^{-15}}{0.10e^{-15} + 1e^{-12}} * 3.3V = 0.33mV \tag{6.13}$$

  This 0.33mV is amplified by the amplifier to produce a shift in the output voltage of around

150mV. To fix such an issue, the layout can be modified, either to completely eliminate the coupling capacitance, or to cancel the effect by adding extra coupling capacitance between Vpd and the complementary net. To add extra capacitance, wires can be extended. However, when doing this, it is important to keep in mind that during fabrication metal layers can shift relative to one another, so the layout should be robust to such shifts.

- The pixels are designed to tile together perfectly.
- It is important to know the total capacitance (including parasitics) on each input, so that information can be given to Synopsys, which will ensure the digital nets are properly driven.

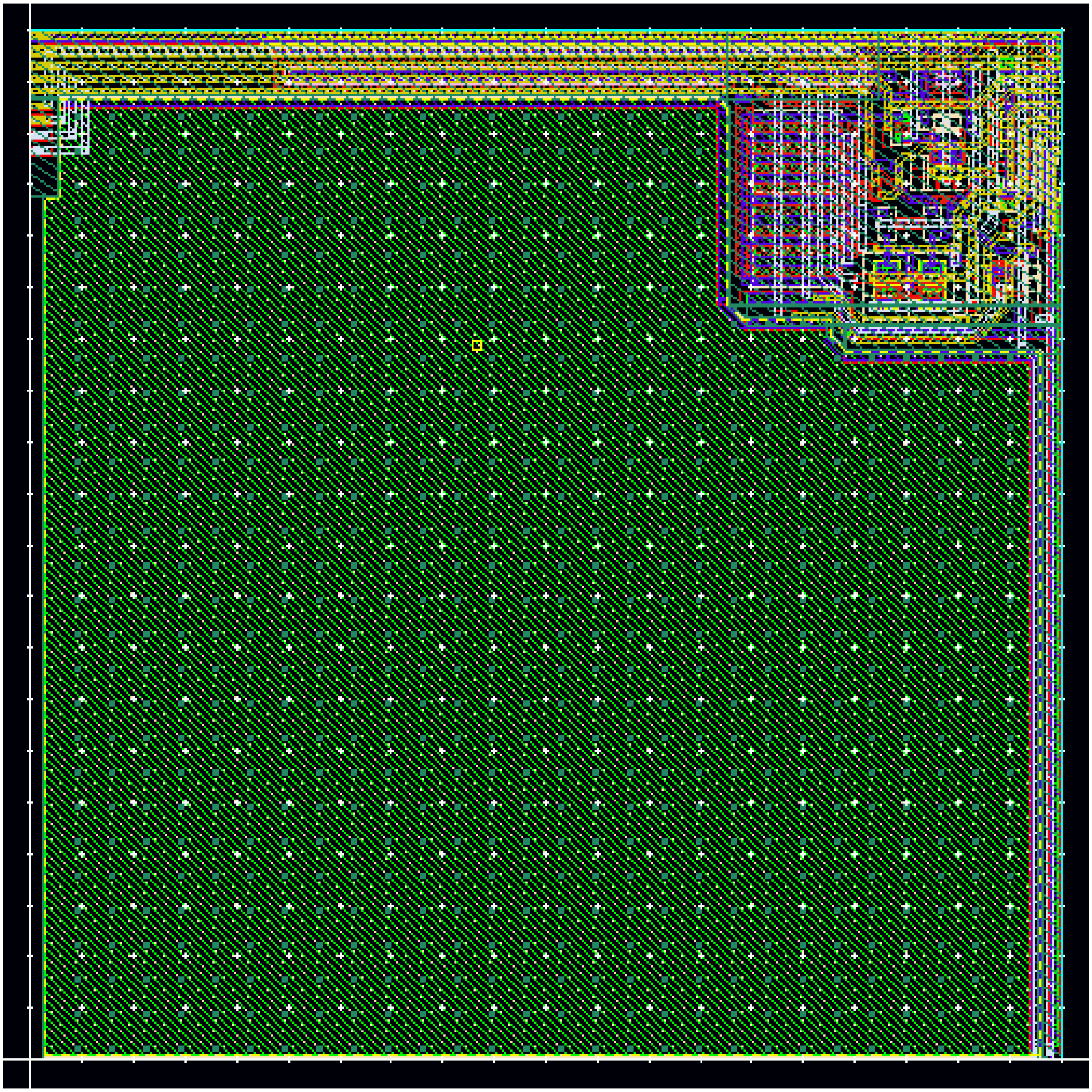


Figure 6.11: Full pixel layout. All of the circuitry other than the photodiode is covered in top metal. The feedback capacitors extend along the top edge underneath the wire lines.

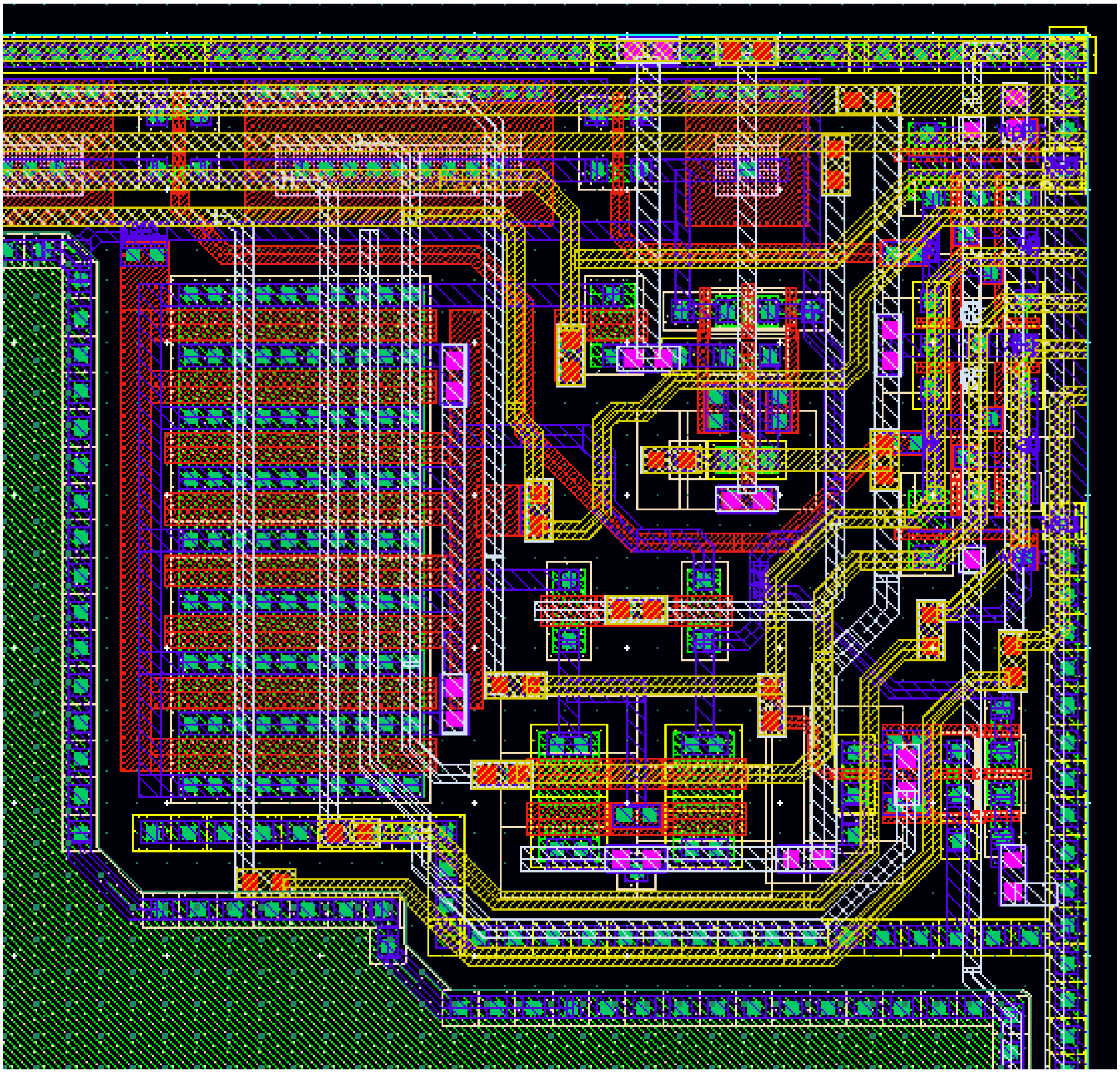


Figure 6.12: Zoomed pixel layout. Here the top metal is removed. Double-vias are used for reliability. Certain wires are extended over the large input device transistors to balance coupling from digital nets.

6.3 Testing Apparatus

6.3.1 Circuit Boards

As with OR1, a custom printed circuit board (OR2PCB) was designed with Altium Designer. OR2PCB is similar to OR1PCB, in that it interfaces with the Opal Kelly XEM6010 FPGA board. OR2PCB includes a separate constant current source driver circuit for each LED die bonded beside the imager. The circuit is shown in Fig. 6.13. The op-amp forces the DAC (AD5060) voltage across the resistor R_e such that the current through the LED is $I_{LED} = V_{dac}/R_e$. The R_{small} resistor allows for an ammeter to be inserted in the circuit to measure the collector current (the true LED current) if desired. Alternatively the emitter current can easily be calculated by measuring V_r/R_e. The PCB includes a 16-bit ADC (AD7985) for digitizing the imager analog output. This off-chip ADC is an alternative option to the on-chip ADC. All measurement results

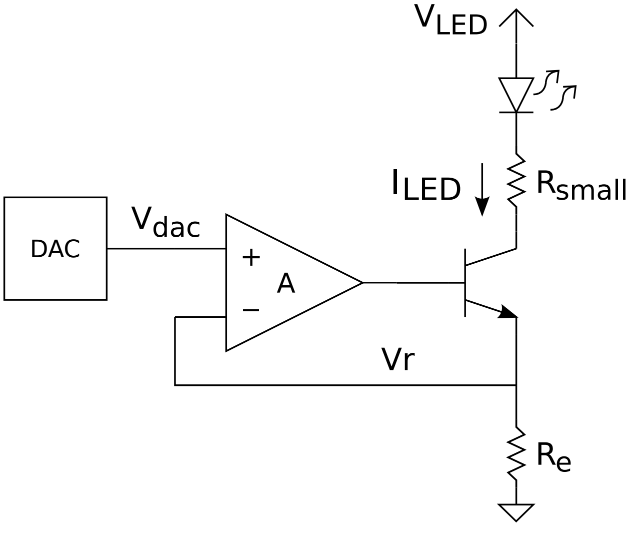


Figure 6.13: Constant current LED driver circuit. The amplifier feedback forces $V_r = V_{dac}$. V_r/R_e then defines the current.

in this section used the higher-resolution off-chip ADC.

A second printed circuit board (LEDPCB) was also designed. It is mounted on top of OR2PCB. The bottom side of LEDPCB has three surface-mounted LEDs (Kingbright APTL3216: 525nm, 601nm, 630nm) that are used to illuminate the imager chip during imager performance characterization. A similar constant current source circuit to the one described above is included on this board to drive the LED. Fig. 6.14 shows a picture of OR2PCB with and without LEDPCB sitting on top of it.

Testing is controlled through verilog HDL code and Python scripts that use the Opal Kelly API. The scripts configure the FPGA, program the imager chip and LEDs, read out the sampled ADC data and process it.

6.3.2 Tissue Phantom

A tissue phantom allows for repeatable experiments that do not require a wet lab or any special handling protocols, and is appropriate for testing initial device prototypes. The tissue phantom is a simple design consisting of fluorescent microspheres mixed into a clear PDMS (polydimethylsiloxane) gel. The fluorescent microspheres are UVPMS-BO-1.00 (27-32μm diameter) from Cospheric [21]. The PDMS is MED 6215 from Nusil [66]. The density of the fluorophores in the gel is representative of targeted neuron density in the brain. Fig. 6.15 shows a picture of the tissue phantom that was used in this work, illuminated with UV light. Sec. 6.5.1 describes in more detail how the tissue phantom is used.

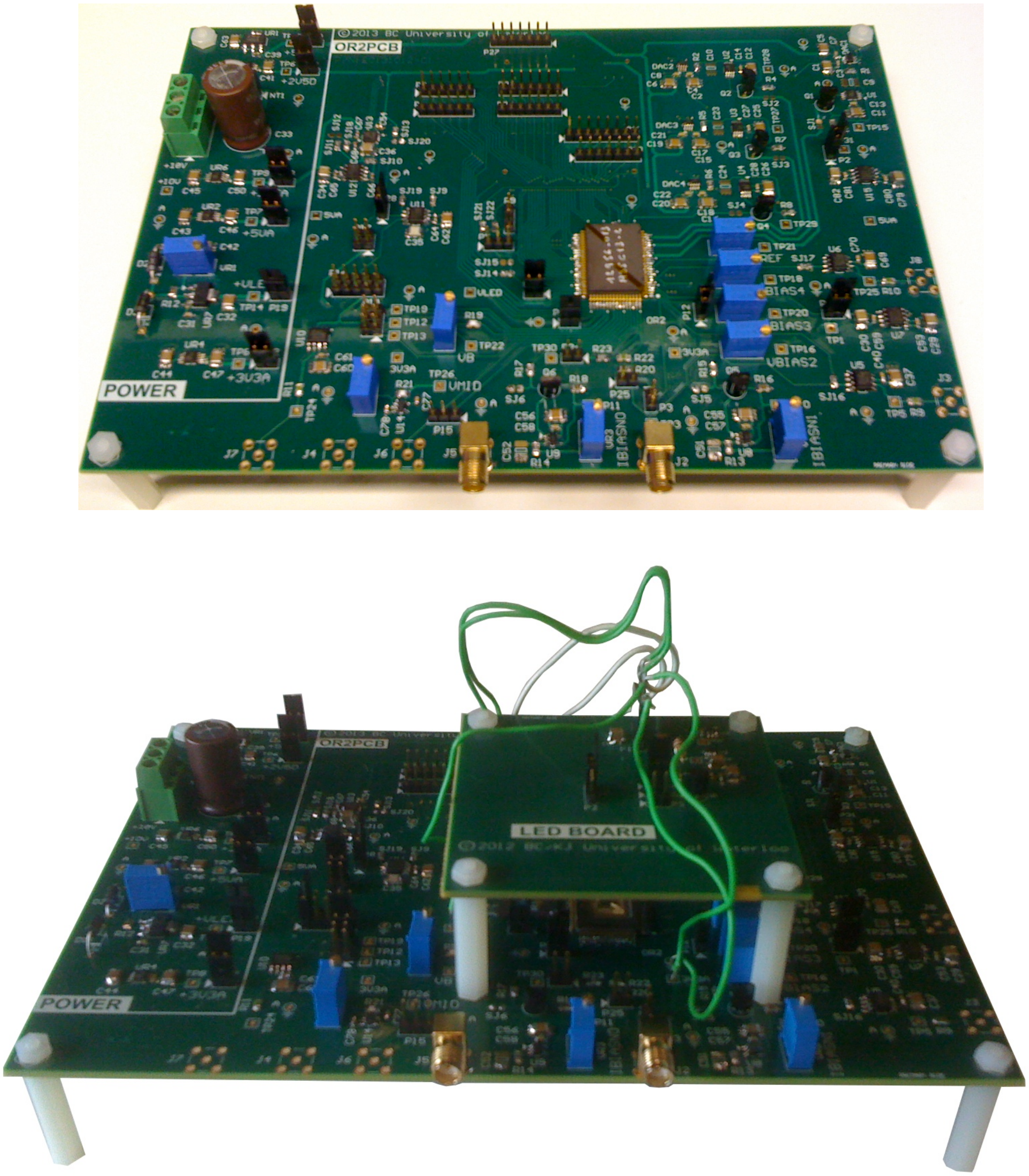


Figure 6.14: Picture of the OR2 testing setup. LEDPCB sits on top of OR2PCB. Green, orange and red LEDs on the bottom side of LED BOARD are used to illuminate the imager for characterization.

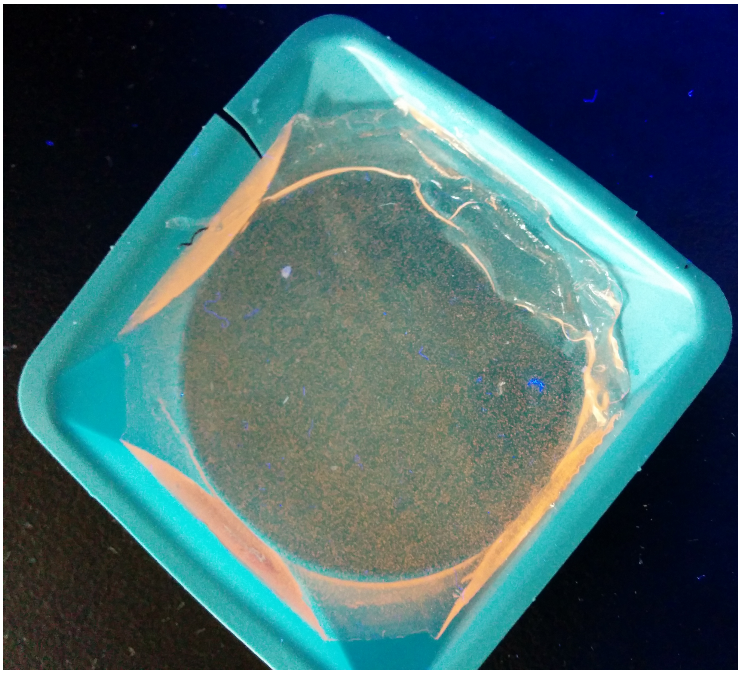


Figure 6.15: Picture of the tissue phantom under UV light (the blue container is about 4cm on a side).

6.4 Imager Performance Characterization

As discussed above, the LEDs on LEDPCB are used to characterize the imager. Several of the plots in the following sections plot some performance metric vs. incident light power, or irradiance (W/cm^2). Therefore, the first step in the testing process is to map the LED current (which is controlled) to LED irradiance. This is done by setting the current and then measuring the LED light output with a radiometer. See Appendix C for the radiometer measurement procedure.

Fig. 6.16 shows a plot of irradiance vs. LED current for each of the three LEDPCB LEDs. Similarly, the irradiance vs. current curve is found for the LED dies that are bonded beside the imager, as shown in Fig. 6.17. Unfortunately LED4 was not operational (determined to be a

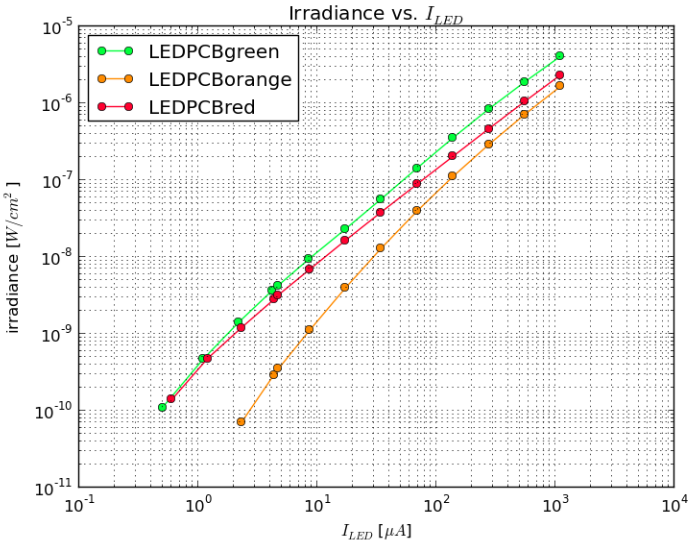


Figure 6.16: Irradiance as a function of current for the LEDPCB LEDs.

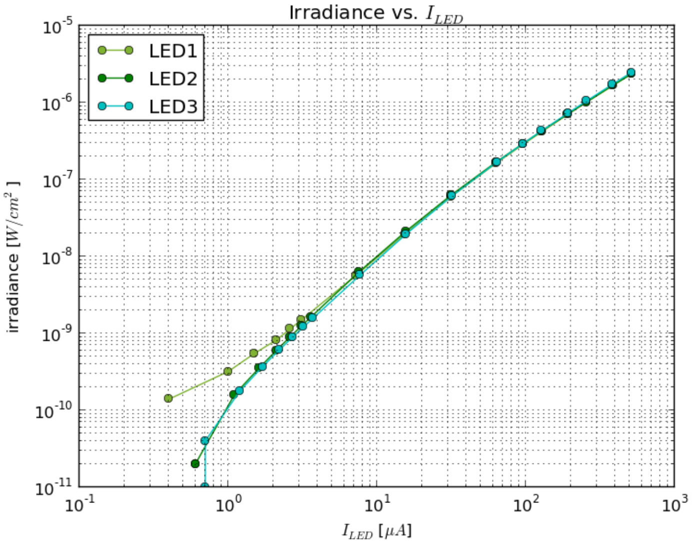


Figure 6.17: Irradiance as a function of current for the LED dies.

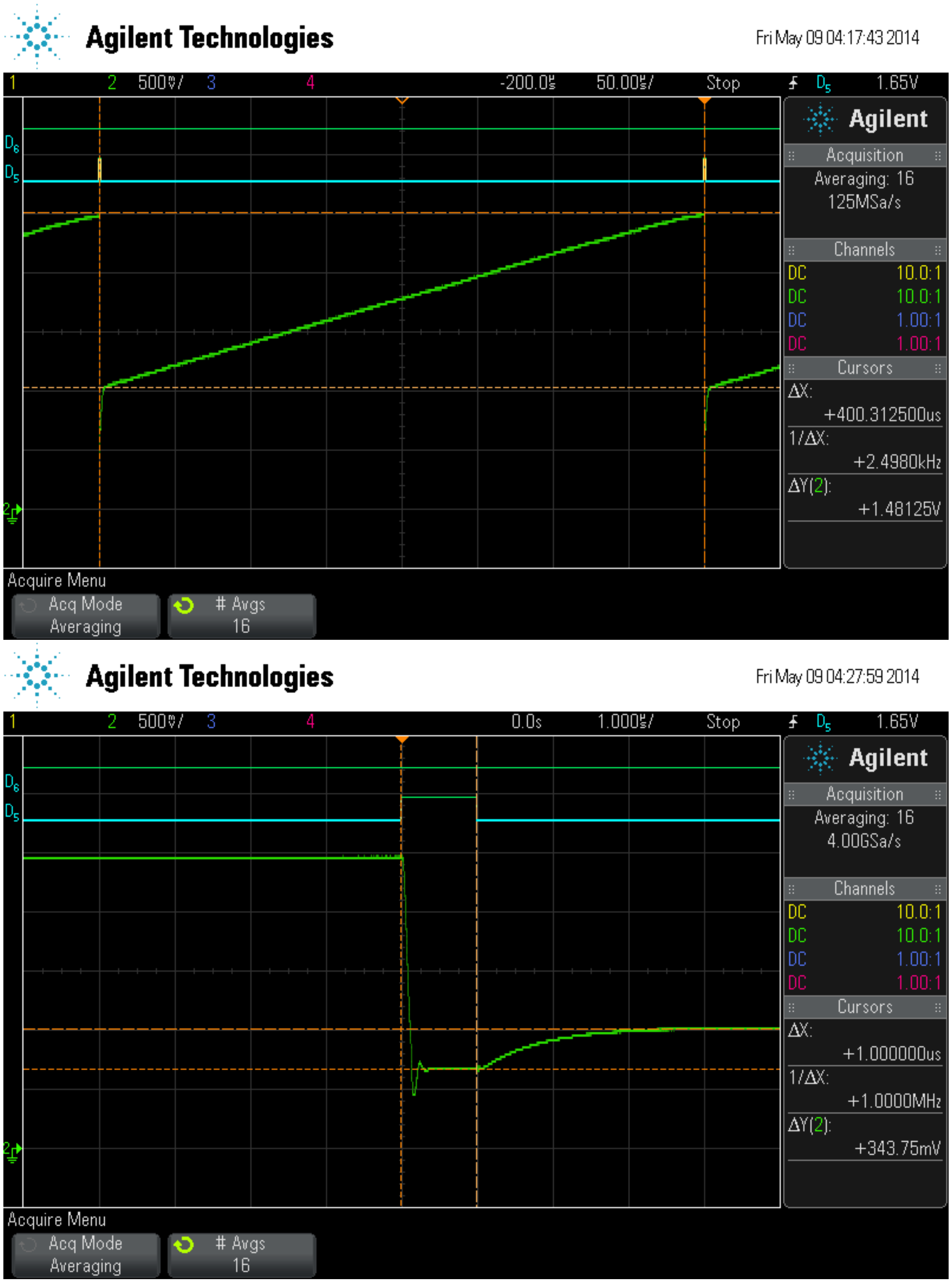
bonding issue) and is not included here.

The imager is characterized at each value of integration capacitance: $C_{min} = 4$fF, $C_{mid} = 20$fF, and $C_{max} = 100$fF. Unless otherwise stated, the results are for the target frame rate of 2500fps. For performance metrics where the statistics are averaged over all of the pixels in the array, such as SNR or fixed pattern noise, only the core 8x8 pixels are included in the calculations. The reason for this is that the pixels around the outer edge of the array see a different environment than the core pixels (which are surrounded on all sides by other pixels), and their outputs are typically much different than the core pixels, which produce relatively similar results. In some instances where frame images are shown, all 10x10 pixels are included; in other instances only the core 8x8 pixels are included.

6.4.1 Pixel Operation

A test pixel was included on the chip for testing purposes. It provides direct control over all pixel signals. This is in contrast to the pixel array which is controlled by on-chip logic only. The main advantage of the test pixel is that its output voltage can been seen over all time.

Fig. 6.18 (top) shows an oscilloscope capture of the test pixel output over time, at 2500fps, and with C_{min}. D6 (green) is the row select signal (which is just left high). D5 (blue) is the reset signal, which pulses high to reset the pixel. In this capture the incident light is just strong enough to saturate the output. Here 16 averages are used on the oscilloscope to eliminate the effect of reset noise. Fig. 6.18 (bottom) zooms in on the output voltage around the reset pulse. When reset goes low, the output quickly jumps up about 350mV. This is due to parasitic coupling into the sensitive CTIA input node, as was discussed previously, and is also seen in simulation.



Fri May 09 04:17:43 2014

2 500mV/ -200.0µs 50.00µs/ Stop D5 1.65V

Averaging: 16
125MSa/s

DC 10.0:1
DC 10.0:1
DC 1.00:1
DC 1.00:1

ΔX: +400.312500us
1/ΔX: +2.4980kHz
ΔY(2): +1.48125V

Acq Mode: Averaging
Avgs: 16

Fri May 09 04:27:59 2014

2 500mV/ 0.0s 1.000µs/ Stop D5 1.65V

Averaging: 16
4.00GSa/s

DC 10.0:1
DC 10.0:1
DC 1.00:1
DC 1.00:1

ΔX: +1.000000us
1/ΔX: +1.0000MHz
ΔY(2): +343.75mV

Acq Mode: Averaging
Avgs: 16

Figure 6.18: Top: Pixel output voltage with pixel just saturated, with C_{min} and 2500fps (16 averages). Bottom: Pixel output voltage zoomed in on the reset pulse (16 averages).

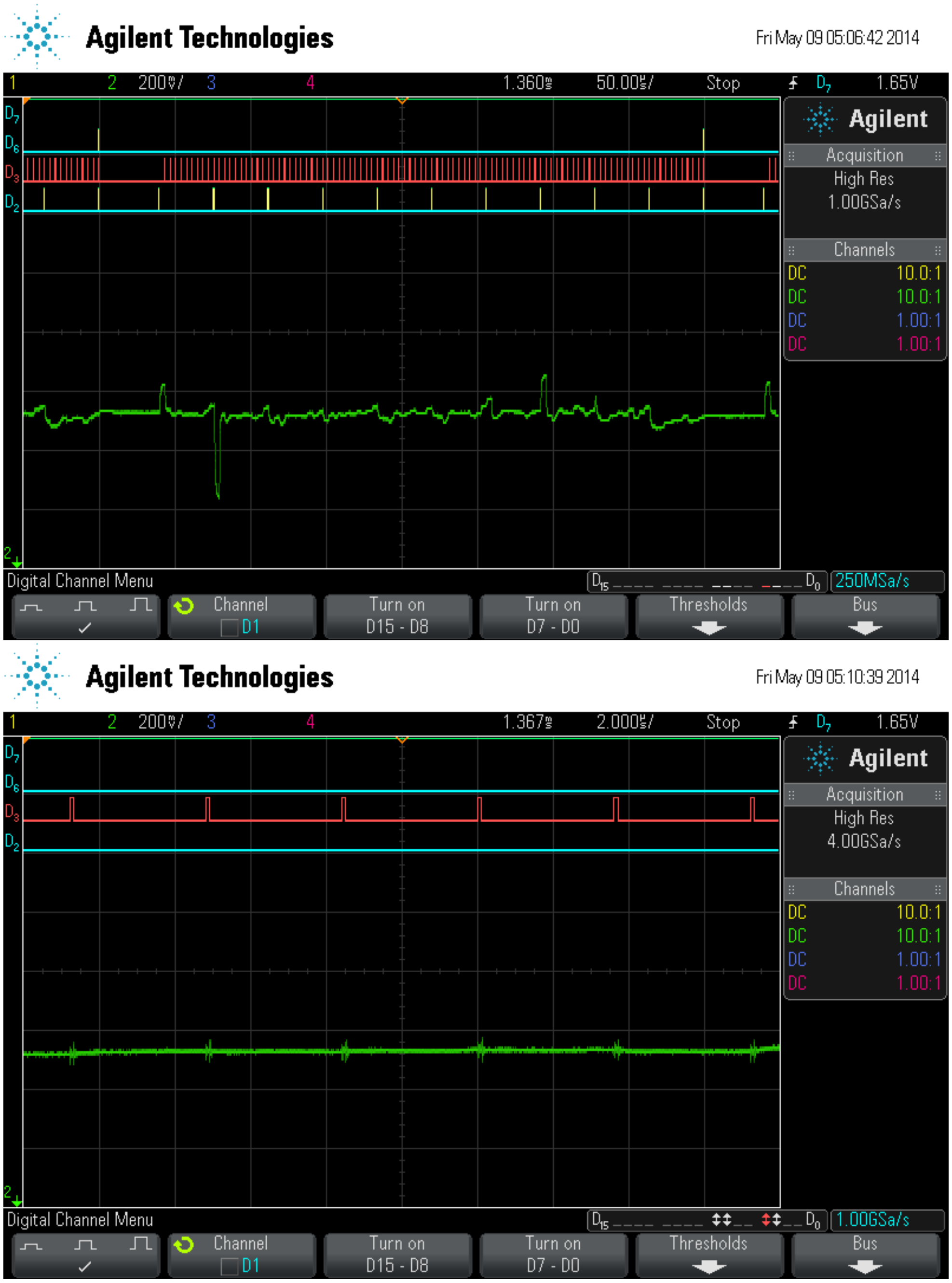


Figure 6.19: Single-shot oscilloscope capture of the imager analog output voltage for all 100 pixels in a frame (top), and for a few pixels (bottom), for C_{min} and 2500fps. A combination of fixed pattern and reset noise is seen.

Fig. 6.19 (top) shows a single-shot (no averaging) oscilloscope capture of the imager analog output voltage for an entire frame, for C_{min} and 2500fps. Fig. 6.19 (bottom) zooms in on a few pixels. D7 is the CAPTURE (imager enable) signal. D6 is the FRAMEDONE signal, which pulses high at the end of each frame capture. D3 is the SHIFTCOL signal, which pulses to select the next column output. D2 is the SHIFTROW signal, which pulses to select the next row. Ideally the output should be a constant value, but we see steps in the waveform due to noise in the imager (fixed pattern and random).

Fig. 6.20 shows a single-shot oscilloscope capture of the imager analog output voltage where C_{fb} changes for adjacent pixels. The high, middle, and low output voltage levels correspond to C_{min}, C_{mid} and C_{max}, respectively.

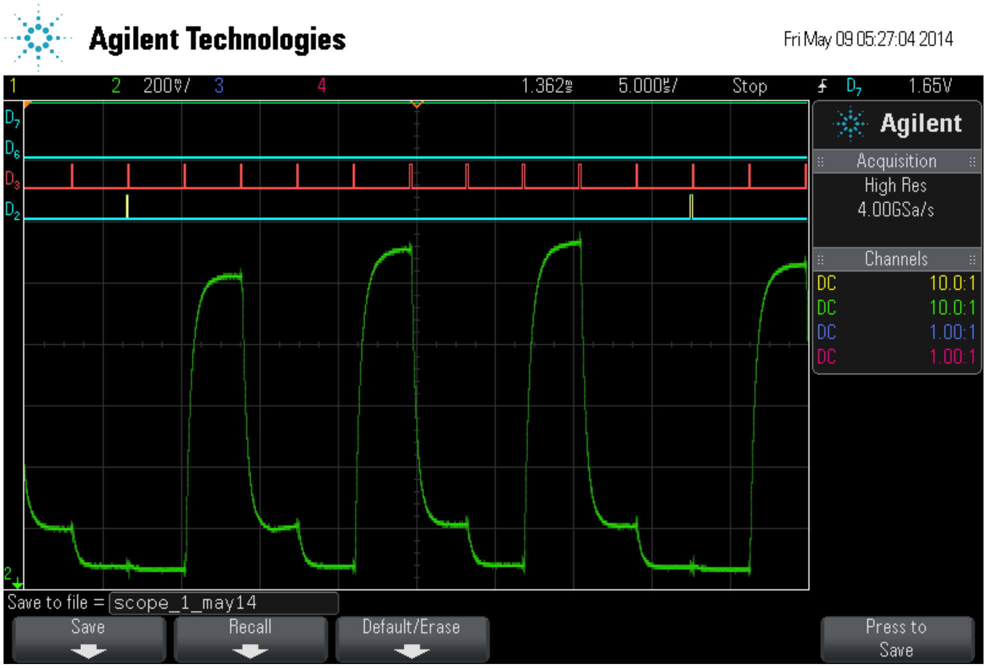
6.4.2 Dark Signal

The dark signal is measured using the test pixel. It is important to see the pixel voltage over the entire integration time because the reset noise of the pixel turned out to be significant, and causes the reset level (the output voltage immediately after reset, at the start of the integration period) to vary by tens of mV, as is shown later in Fig. 6.28.

The dark current is calculated with

$$i_{dark} = C_{fb}\frac{V_{samp} - V_{reset}}{t_{int}}[A] \ . \tag{6.14}$$

In order to get reliable dark current measurements, a long integration time of 50ms is used such that the output voltage is significant. 1024 averages on the scope are also used to eliminate the



Acquisition
High Res
4.00GSa/s

DC 10.0:1
DC 10.0:1
DC 1.00:1
DC 1.00:1

Save to file = scope_1_may14

Figure 6.20: Single-shot oscilloscope capture of the imager analog output voltage where C_{fb} changes for adjacent pixels. The high, middle, and low output voltage levels correspond to C_{min}, C_{mid} and C_{max}, respectively.

reset noise, which effectively DC-shifts the output. Cursors are then used to find V_{samp} and V_{reset}.

According to design kit documentation, the dark current should be around 8fA for this pixel size. Table 6.6 shows the dark current with each C_{fb} setting. The larger measured dark current is likely due to another leakage component. The dark current of 38fA corresponds to 0.45nA/cm^2

C_{fb}	i_{dark}
C_{min}	38fA
C_{mid}	39fA
C_{max}	38fA

Table 6.6: Measured pixel dark current.

when dividing by the photodiode area. The dark signal is 9.5V/s (with C_{min}).

The dark current level can be converted to an equivalent incident light power as follows:

$$I = \frac{i_{dark}}{Resp} * \frac{1}{A_{px} * FF} * \frac{1}{(10^{-4})^2}\ [W/cm^2]\ , \tag{6.15}$$

where $Resp$ is the photodiode responsivity (0.29A/W), A_{px} is the pixel area in μm^2 (10000μm^2), and FF is the fill factor (83.3%). The following equations can be used to convert to a photon count:

$$q_{dark} = C_{fb}(V_{samp} - V_{reset}) = i_{dark}t_{int}\ [Coulombs] \tag{6.16}$$

$$N_{e^-} = \frac{q_{dark}}{1.6 * 10^{-19}}\ [electrons] \tag{6.17}$$

$$N_{ph} = N_{e^-} * QE_{pd}\ [photons] \tag{6.18}$$

where QE_{pd} is the quantum efficiency of the photodiode (0.6321). Table 6.7 shows each of these calculated values.

i_{dark}	irradiance	electrons	photons
38fA	1.57nW/cm^2	95	60

Table 6.7: Dark current equivalents (2.5kHz frame rate).

6.4.3 Imager Average Reset Level and Fixed Pattern Noise

In this work, the reset level of the pixel is defined as the pixel output voltage under zero-light conditions (i.e. due to dark current). This reset level is subtracted from the output voltage samples to get the voltage signal of interest (i.e. the voltage due to light integration).

The fixed pattern noise (FPN) of the imager is defined as the standard deviation of the average pixel values (over 1000 frames). FPN will vary for different lighting and circuit conditions. Figs. 6.21, 6.22 and 6.23 plot the FPN vs. irradiance at 2500fps for red, orange and green illumination, respectively.

Fig. 6.24 (a-c) shows the average reset level of each pixel for each C_{fb} value. (d-f) shows the average sample level of each pixel for orange illumination at 2e-6 W/cm^2, for each C_{fb} value. From looking at the images it is clear that the core pixels (8x8 in the center) in the array have relatively little variation in their response, whereas many of the outer pixels of the array (rows and columns 1 and 10) deviate from the average response of the core pixels. Column 10 and row 1 tend to have higher voltages. These pixels are closer to the edge of the chip. Row 10 tends to have lower voltages; it is closer to the center of the chip. In general the outer pixels do not see the same environment as the inner pixels, because they are not entirely surrounded by other pixels like the inner pixels are. In future designs it may be prudent to include an extra ring of pixels to improve matching, but not necessarily connect the pixels to the control or readout circuitry.

6.4.4 Output-Referred Read and Reset Noise

The total output noise of the imager, including reset and read noise, can be calculated from the sampled data by finding the standard deviations of each pixel over 1000 frames, and averaging the results. Figs. 6.25, 6.26 and 6.27 plot the output noise vs. irradiance for red, orange and green illumination, respectively. For a given C_{fb}, the noise is essentially constant and independent of wavelength, and a larger C_{fb} results in lower noise.

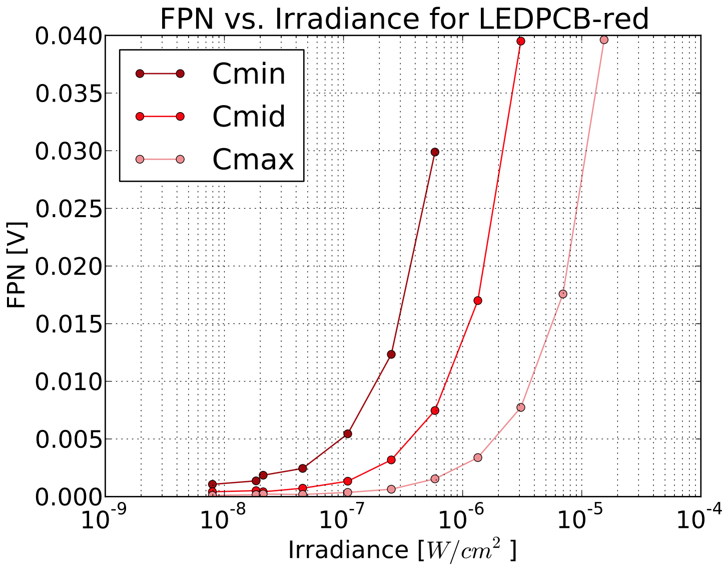


Figure 6.21: Imager FPN vs. irradiance for red illumination.

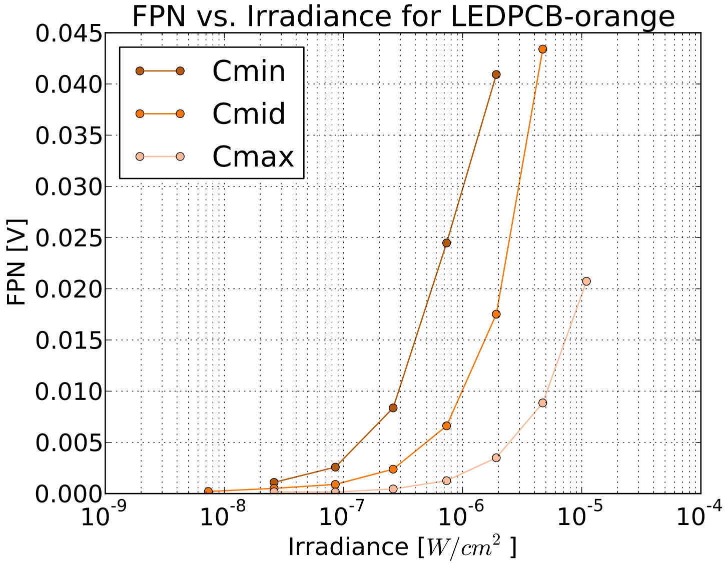


Figure 6.22: Imager FPN vs. irradiance for orange illumination.

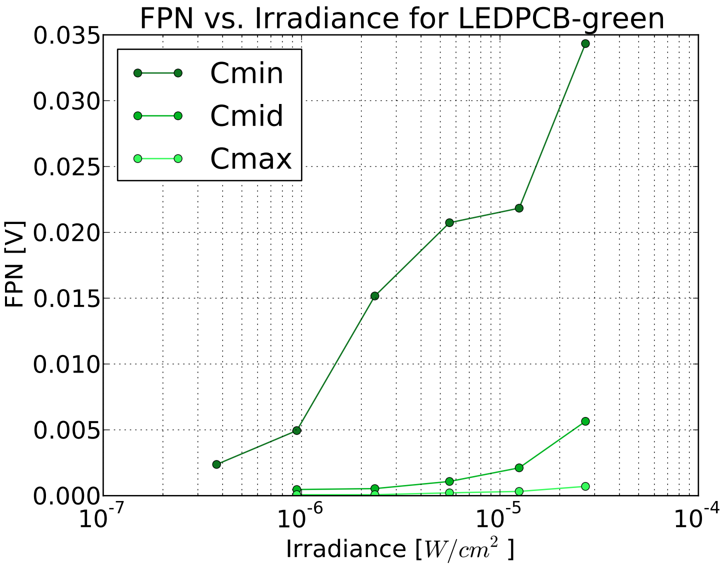


Figure 6.23: Imager FPN vs. irradiance for green illumination.

The pixel reset noise is found from estimating the peak-to-peak variation in reset level of the pixel output voltage. This is done using the *persistance* functionality of the oscilloscope. A common way to estimate RMS noise is to equate the peak-to-peak noise to a 6σ or 6.6σ range. 6.6σ variation corresponds to 1/1000 samples being outside that range. Therefore, if we collect 1000 samples and find the peak-to-peak value, we can divide the result by 6.6 to obtain a value for the RMS noise (1σ). 1000 data points are captured several times and the results are averaged to find the final estimate. To estimate the peak-to-peak noise using the oscilloscope, the oscilloscope persistance is set to a time frame corresponding to the desired number of samples (the oscilloscope update rate must be known). Then for one trial we stop the oscilloscope and use cursors to measure the peak-to-peak voltage range at a given instant in time (here, after reset). This is repeated 10 times and the results are averaged. Fig. 6.28 shows an example of this. The time instant at 50us is chosen. From this method, the RMS reset noise was found to be 24.06mV

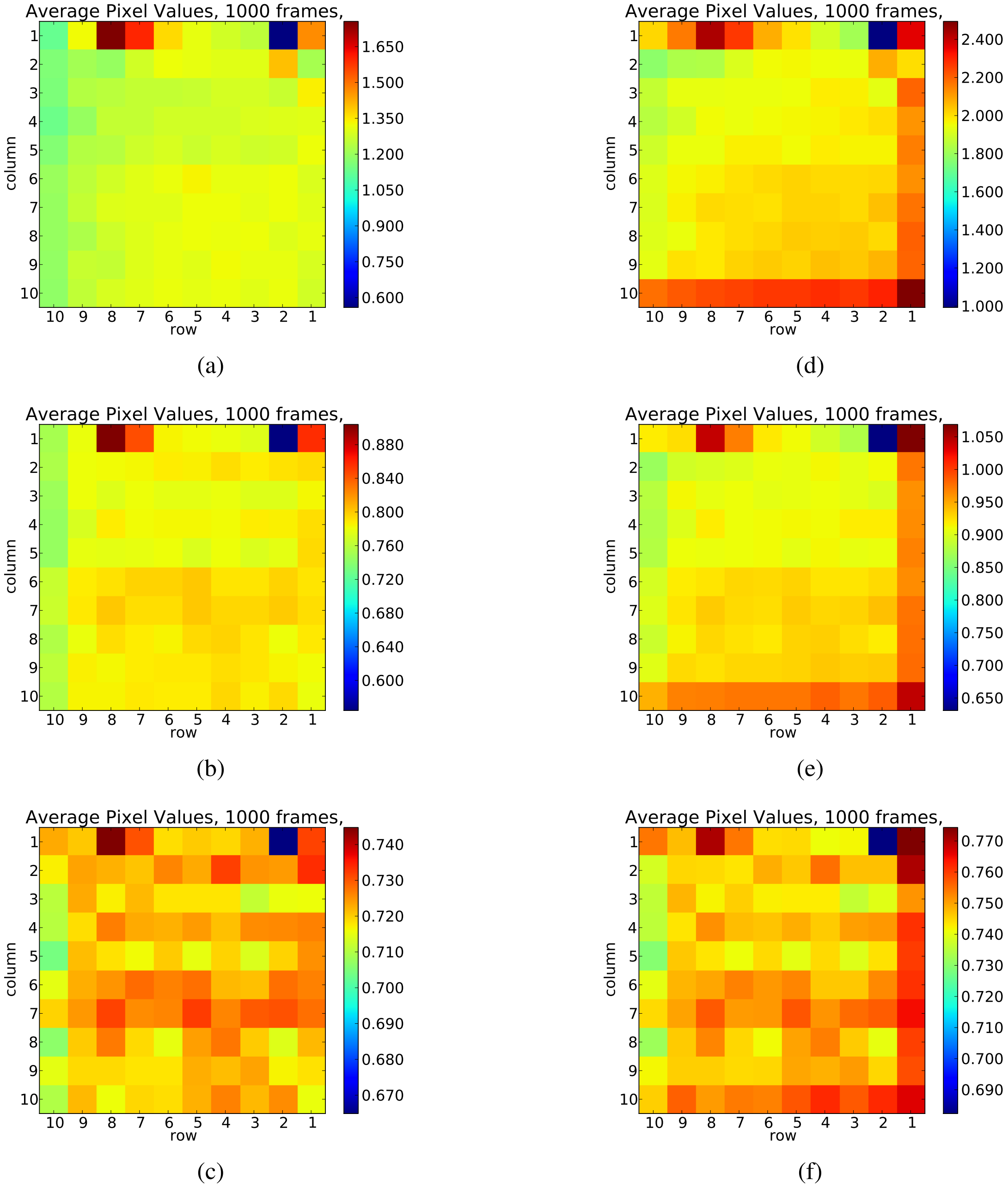


Figure 6.24: (a-c) Average reset level of each pixel for C_{min}, C_{mid}, and C_{max}. (d-f) Average voltage level near saturation (for C_{min}) for C_{min}, C_{mid}, and C_{max}. Here is is clear that the pixels around the outside of the imager have significant variation compared with the core pixels.

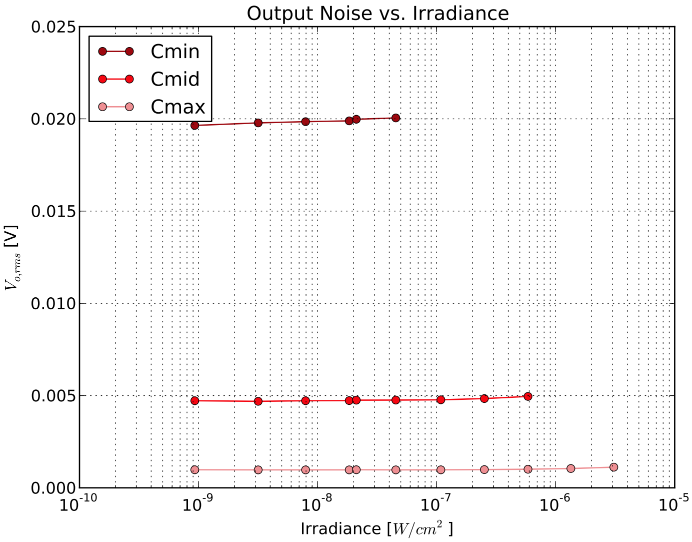


Figure 6.25: Imager output noise vs. irradiance for red illumination.

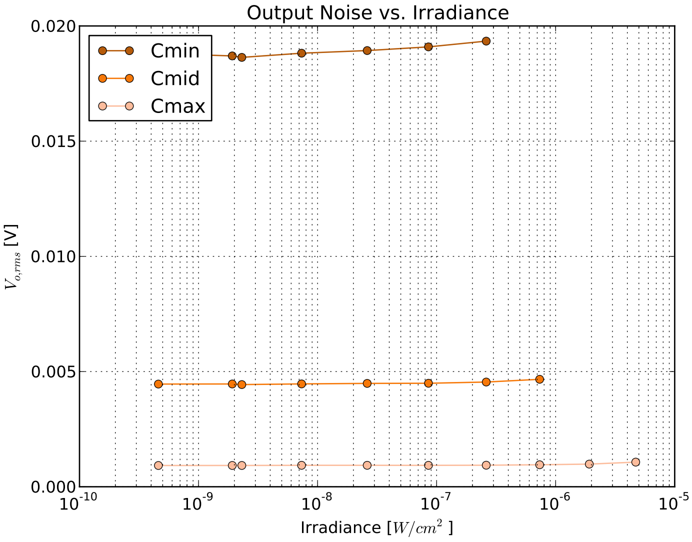


Figure 6.26: Imager output noise vs. irradiance for orange illumination.

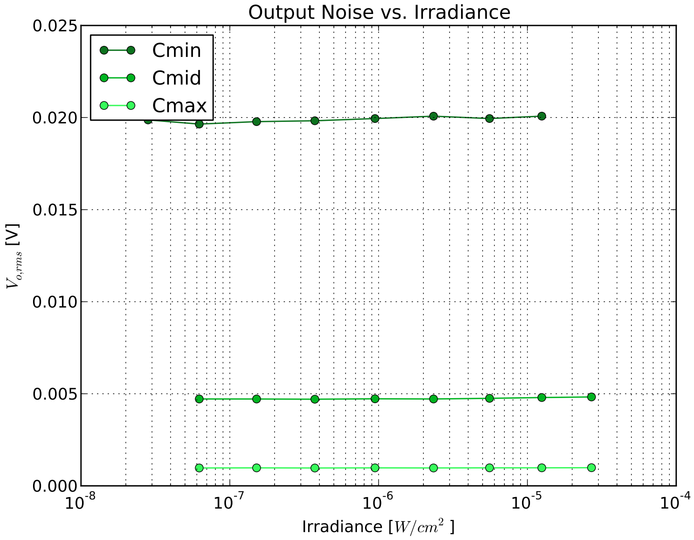
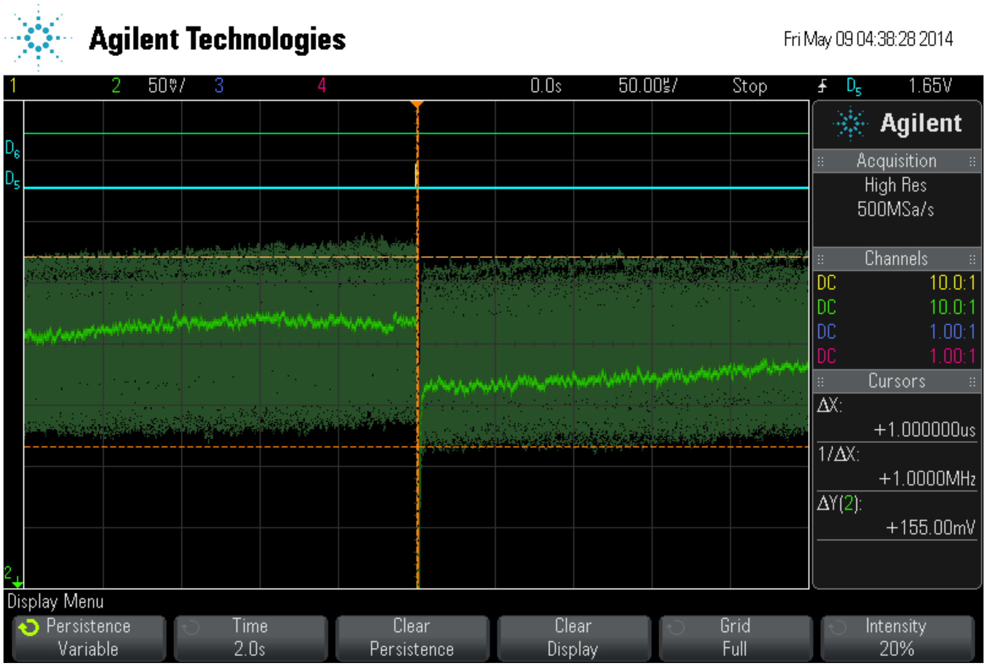


Figure 6.27: Imager output noise vs. irradiance for green illumination.



Acquisition
High Res
500MSa/s

Channels
DC 10.0:1
DC 10.0:1
DC 1.00:1
DC 1.00:1

Cursors
ΔX: +1.000000us
1/ΔX: +1.0000MHz
ΔY(2): +155.00mV

Figure 6.28: Oscilloscope capture of the test pixel output voltage with persistance, used to find the pixel reset noise.

(note this is for C_{min} and 2500fps).

The read noise can then be found using the following equation:

$$V_{n,read}^2 = V_{n,total}^2 - V_{n,reset}^2 \; . \tag{6.19}$$

In this case, the estimated reset noise from the test pixel is actually larger than the total noise calculated from the sampled data from the imager. The likely cause of this is that capacitive loading of the test pixel is lower than that of the pixels in the array, which all see an extra 9 pixels of column capacitance. The noise analysis in Appendix A indicates that the reset noise is much more significant than the read noise. Therefore, the read noise is assumed to be small and is not calculated here.

6.4.5 Imager Average Output

The imager average output is a plot of the average signal voltage for all pixels in the array at a given incident light power and given wavelength. The average signal voltage is found by subtracting the average dark signal voltage from the average light signal voltage (both over 1000 frames). Figs. 6.29, 6.30 and 6.31 plot the average output vs. irradiance at 2500fps for red, orange and green illumination, respectively.

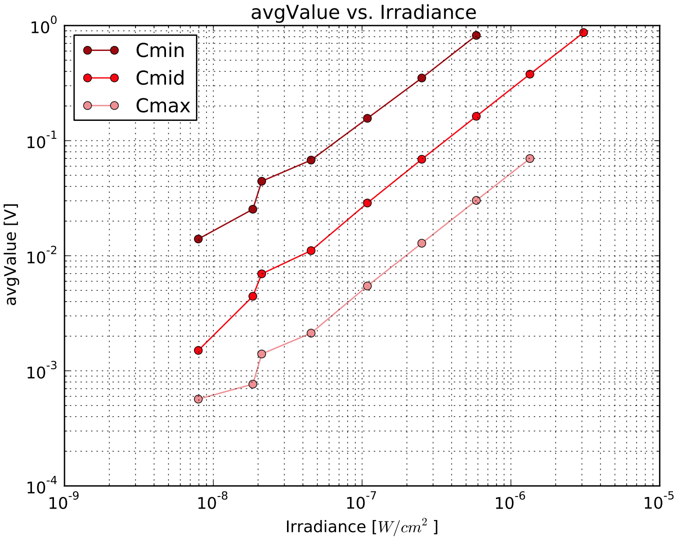


Figure 6.29: Imager average signal vs. irradiance for red illumination.

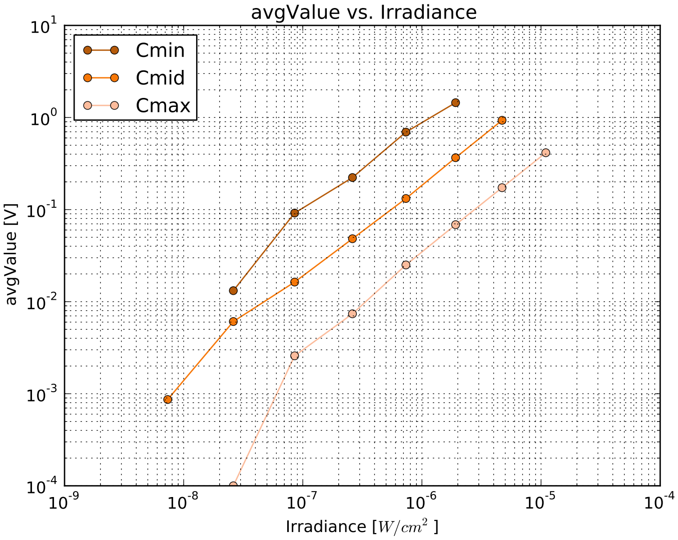


Figure 6.30: Imager average signal vs. irradiance for orange illumination.

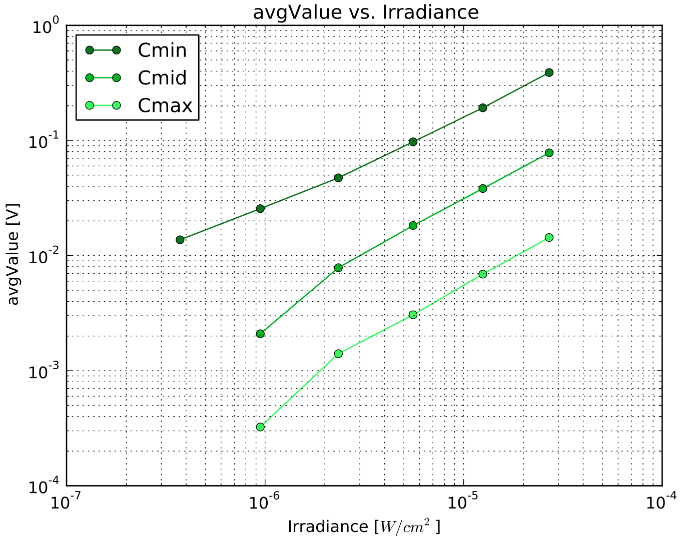


Figure 6.31: Imager average signal vs. irradiance for green illumination.

6.4.6 Limit of Detection

The limit of detection (LOD) of the imager is defined here as the incident light power that gives an SNR of 0dB. The average SNR of the imager is defined as

$$SNR = \frac{\left(\sum_1^N \mu_i\right)/N}{\left(\sqrt{\sum_1^N \sigma_i^2}\right)/N} \tag{6.20}$$

where N is the number of pixels, μ_i and σ_i are the mean and standard deviation of the output of the ith pixel calculated over 1000 frames [60]. Note that here the output is defined as the average sampled output voltage of the pixel minus the average reset level of the pixel. The average reset level is found by averaging the pixel dark current output voltage over 1000 frames. Note this “reset” level is not the pixel voltage immediately after the pixel is reset.

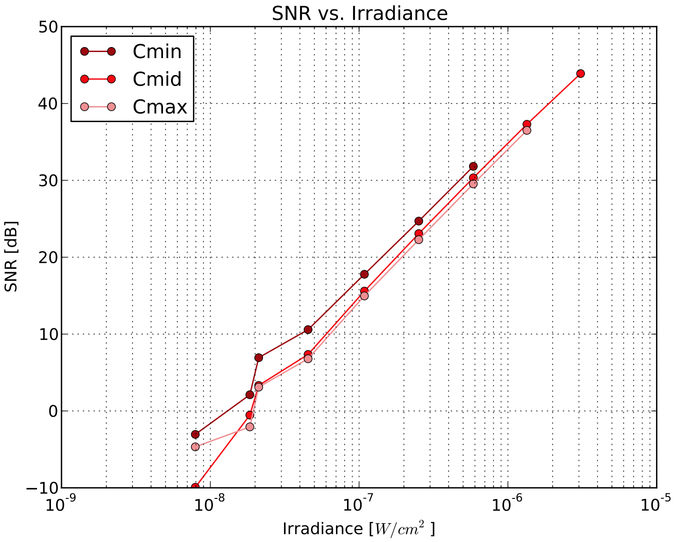


Figure 6.32: Imager SNR vs. irradiance for red illumination.

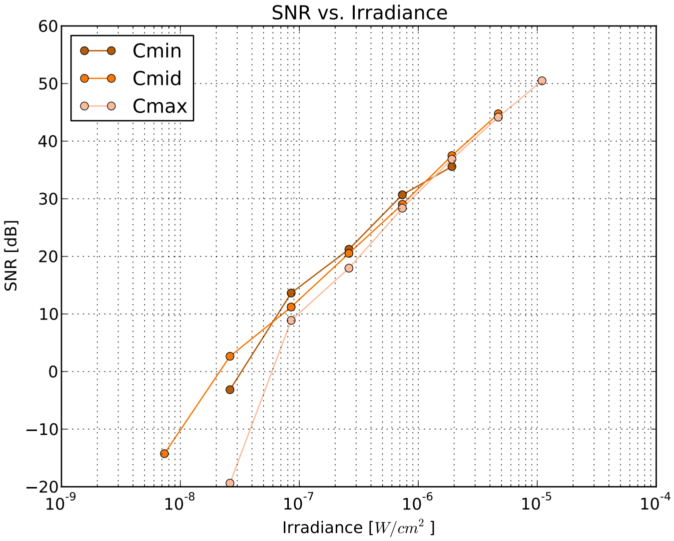


Figure 6.33: Imager SNR vs. irradiance for orange illumination.

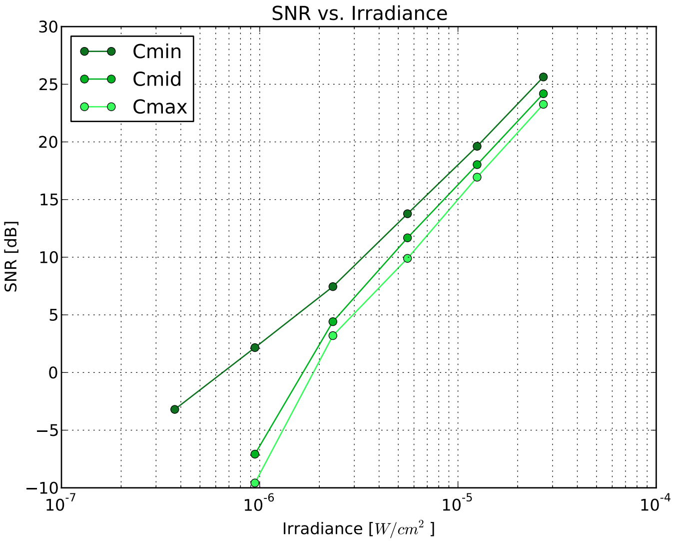


Figure 6.34: Imager SNR vs. irradiance for green illumination.

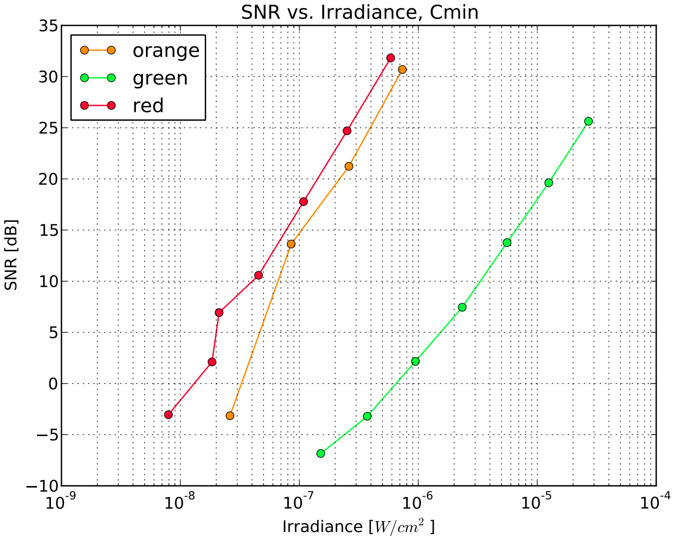


Figure 6.35: Imager SNR vs. irradiance for each wavelength and C_{min}.

Figs. 6.32, 6.33 and 6.34 plot the SNR vs. irradiance at 2500fps for red, orange and green illumination, respectively. Fig. 6.35 compares the SNR for each incident wavelength at the minimum feedback capacitance. We can see that, as expected, for the green light that is blocked by the optical filter, a higher irradiance is required to achieve the same SNR as the higher orange and red wavelengths.

Fig. 6.36 shows the SNR vs. irradiance curves for a chip with no optical filter included, with green excitation. We can see here that a lower LOD is achievable with no filter. This indicates that the filter absorbs some light even at wavelengths that it should pass.

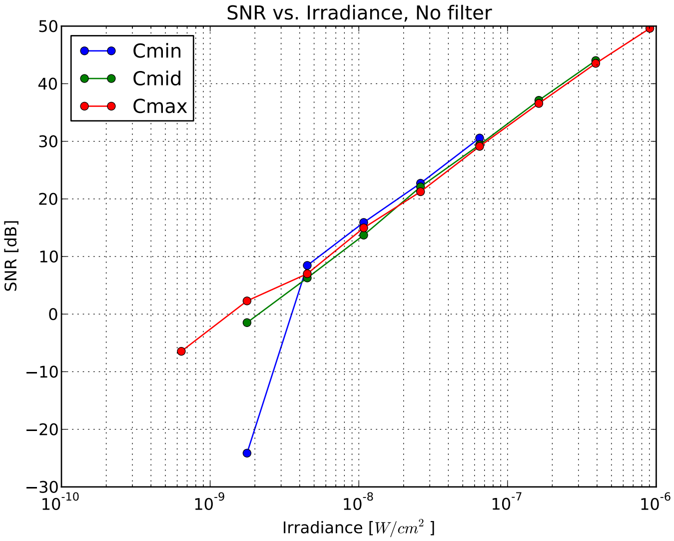


Figure 6.36: Imager SNR vs. irradiance for green excitation and no filter.

Table 6.8 summarizes the LOD in W/cm^2 for each feedback capacitance and incident wavelength, and for the case with no filter and green illumination. The LOD values are estimated from the SNR curves.

LOD	Green	Orange	Red	No Filter
C_{min}	6.4e-7	3.1e-8	1.2e-8	3.3e-9
C_{mid}	1.5e-6	2.0e-8	2.0e-8	2.0e-9
C_{max}	1.8e-6	6.0e-8	2.1e-8	1.3e-9

Table 6.8: Limit of detection (W/cm^2) for each C_{fb} and LED color, and for green illumination with no filter.

6.4.7 Comparison With the Literature

Table 6.9 compares the imager designed in this work with several low-intensity CMOS imagers from the literature. For this work the imager characterization results for two cases are shown: 1) C_{min} and the 601nm orange light with the optical filter included, and 2) C_{min} and the 525nm green light without filter. A simple figure of merit (FOM) for imager performance comparison is proposed here:

$$FOM = \frac{FR}{A_{px} I_{LOD}} \left[\frac{Hz}{nW}\right], \tag{6.21}$$

where FR is the frame rate in frames per second, A_{px} is the pixel area [cm^2] and I_{LOD} is the limit of detection at an SNR of 0dB [nW/cm^2]. Here a larger FOM (higher frame rate, smaller pixel area, lower LOD) indicates higher performance. Case 2 with no filter achieves the highest FOM among devices where the FOM could be calculated.

Reference	[60]	[63]	[25]	[9]	[75]	[78]	[77]	[89]	Case 1	Case 2
Imager Size	132x124	176x144	8x16	128x128	32x32	32x32	100x100	64x128	10x10	10x10
Pixel Size [μm^2]	20.1x20.1	7.5x7.5	240x210	7x7	75x75	75x75	30x30	19x19	100x100	100x100
Fill Factor [%]	42	-	44.6	30	45	50	44.4	27	83.3	83.3
CMOS Technology [μm]	0.5	0.35	0.18	0.18	0.5	0.5[b]	0.35	0.35	0.35	0.35
Frame Rate [fps]	70	0.3	0.03	30	40	500	500	0.075	2500	2500
Dark Current [fA]	0.237	-	0.39[g]	1.47[c]	-	162	-	3.6	38	38
Dark Current [nA/cm^2]	0.14	-	0.0017[g]	10[c]	-	6.4	-	3.62	0.45	0.45
FPN [%][j]	0.84	-	-	0.16[i]	0.02	1.5[i]	0.4	-	2.6	2.6
Well size [e^-]	-	-	-	-	700000	2.1M	253000	1222800	45036	45036
Conv. Gain [$\mu V/e^-$]	32.15	-	-	27.5	0.8	0.52	4.9	1.6	34.26	34.26
Reset Noise (op) [μV]	192.9	-	-	273	-	115	-	239	24060	24060
Reset Noise (ip)[e^-]	6	-	-	9.9	-	221	-	149	-	-
Read Noise (op) [μV]	824	-	-	-	54	260	-	52	-	-
Read Noise (ip) [e^-]	26	-	-	-	67	500	-	33	-	-
SNR max. [dB]	44	-	44[e]	47[h]	76	61	54	60.1	36	40
SNR [dB]	0	0	0	15	35.2	-	16.3[e]	-	0	0
Wavelength [nm]	450	470	562	450	-	-	-	-	601	525
Intensity [nW/cm^2]	4	100	0.0001	400	2900	-	-	-	30	2
Intensity [photons/s/μm^2]	90.6	2366	0.0028	9061	-	-	14300	-	908	52.8
Photon Count	520	443625	4753	14800	1.1e7	-	25740	-	3631	212
FOM [Hz/nW]	4.33e6	5.33e3[f]	0.59e6	-[a]	-[a]	-[a]	-[a]	-[a]	0.833e6	12.5e6

[a] Irradiance not reported at SNR = 0dB.

[b] Assumed based on older paper.

[c] Best case dark current, at low reset voltage of 0.6V.

[d] FPN calculated over 64 frames.

[e] Eyeballed from plot.

[f] Fill factor not reported, assumed to be 1.0 (ideal best case).

[g] Calculated assuming charge integration on 150fF and 0.03fps.

[h] Assumed from plot limit.

[i] Saturation level for FPN measurement not specified.

[j] FPN calculated at or near saturation.

Table 6.9: Comparison of low-intensity CMOS imagers. Read and reset noise is output-referred. The incident intensities and the photon count are the values required to achieve the reported SNR. The intensity at SNR = 0dB is considered the limit-of-detection (not reported for all imagers).

6.5 System Testing

As was discussed in Sec. 6.1.6, the OR2 dies were diced larger than they should have been, resulting in the LED-to-imager distance being about 400μm larger than desired. This larger distance has some impact on the testing results. In theory, if zero excitation light is hitting the imager, then simply increasing the LED current can compensate for the larger distance to the fluorophores. However, if excitation light does hit the imager, then increasing the LED irradiance may result in greater pixel saturation and lower SNR. Fortunately, it turned out that the problem was not too severe.

In the particular chip used for this testing, LED4 was not functional. The results shown here are for LEDs 1, 2 and 3.

6.5.1 Adding the Tissue Phantom

An x-acto knife was used to cut a piece from the tissue phantom. In order to cut the sample to the same size as the imager, a 1mm x 1mm line grid was printed on paper. Placing the clear tissue phantom on the paper, the grid lines are visible and can be used as a guide for cutting.

Before placing the sample on the chip, the sample was cleaned with alcohol and a cloth. Tweezers were used to pick up the sample and carefully place it on the chip. The handling must be gentle to avoid damaging any bondwires, which were not encapsulated.

Figs. 6.37 and 6.38 show microphotographs of the tissue sample sitting on the chip above the pixel array. The dark round shapes are microspheres. In this particular sample, there are no microspheres over the lower half of the pixel array.

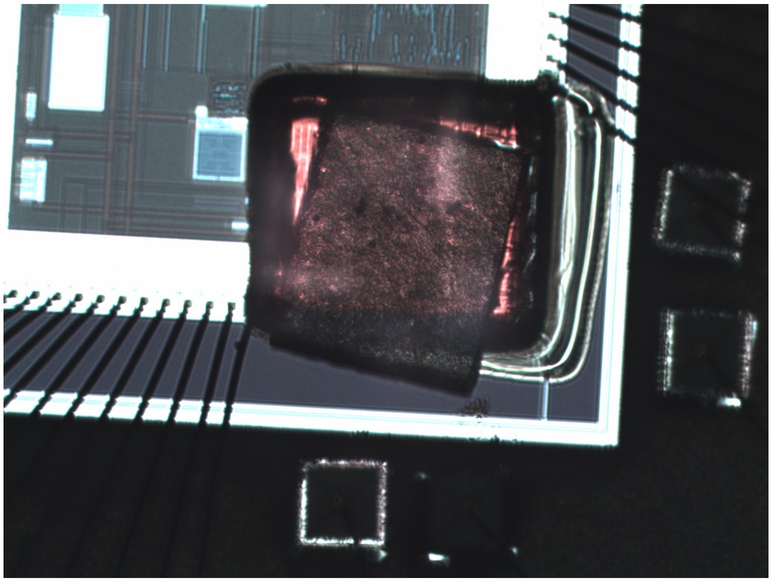


Figure 6.37: Microphotograph of the tissue sample sitting on the chip.

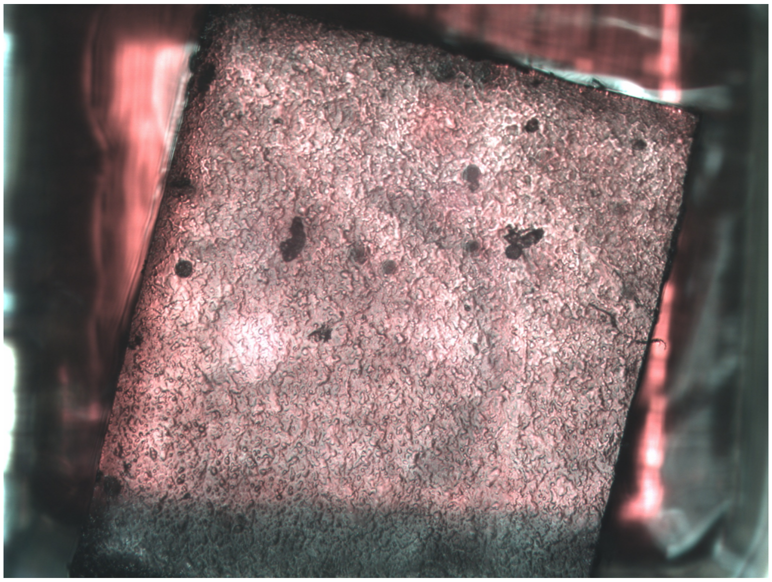


Figure 6.38: Close-up microphotograph of the tissue sample sitting on the chip. The black dots are the microspheres. Microspheres are generally located along the left edge, across the middle and along the top edge to the right, and not in the bottom portion of the phantom.

6.5.2 Baseline and Fluorescent Signal

The left side of Fig. 6.39 shows the baseline signal in the imager due to each adjacent LED die (by itself) without the tissue sample. For this plot the LED currents were set to $24\mu A$, $24\mu A$ and $48\mu A$. C_{min} and 2500fps are used. The plots are an average signal level over 1000 frames. The baseline signal is near zero, indicating that almost no excitation light is being collected. The right side of the figure shows the signal with the tissue phantom placed on top of the chip, at the same settings. We see that the fluorescent signal is significant, up to the 1V range, despite the LED intensity amplitude being less than 100% of the baseline level. This indicates that the signal is indeed due to fluorescence. The apparent fluorophore locations match the fluorophore locations seen in the microphotograph.

6.5.3 Fluorescent Signal Over Time

The key test for this device is whether or not it can distinguish a spike from the baseline signal. In order to emulate a spike, the LED current was modulated with 1ms sine wave snippets. Sine wave data with the desired DC level and amplitude was loaded into block RAM (BRAM) on the FPGA. The BRAM data was then periodically written to the DAC controlling the LED current, at the required speed. Fig. 6.40 shows the resulting signal from pixel (7,2) due to the modulated light from LED1 (the 1ms sine wave snippet is inserted approximately every 5ms). The signal is strong enough to overcome the noise level.

Fig. 6.41 shows a series of consecutive frame images (like a video) around a sine wave snippet. Each frame is a plot of the current frame minus the previous frame; i.e. the difference

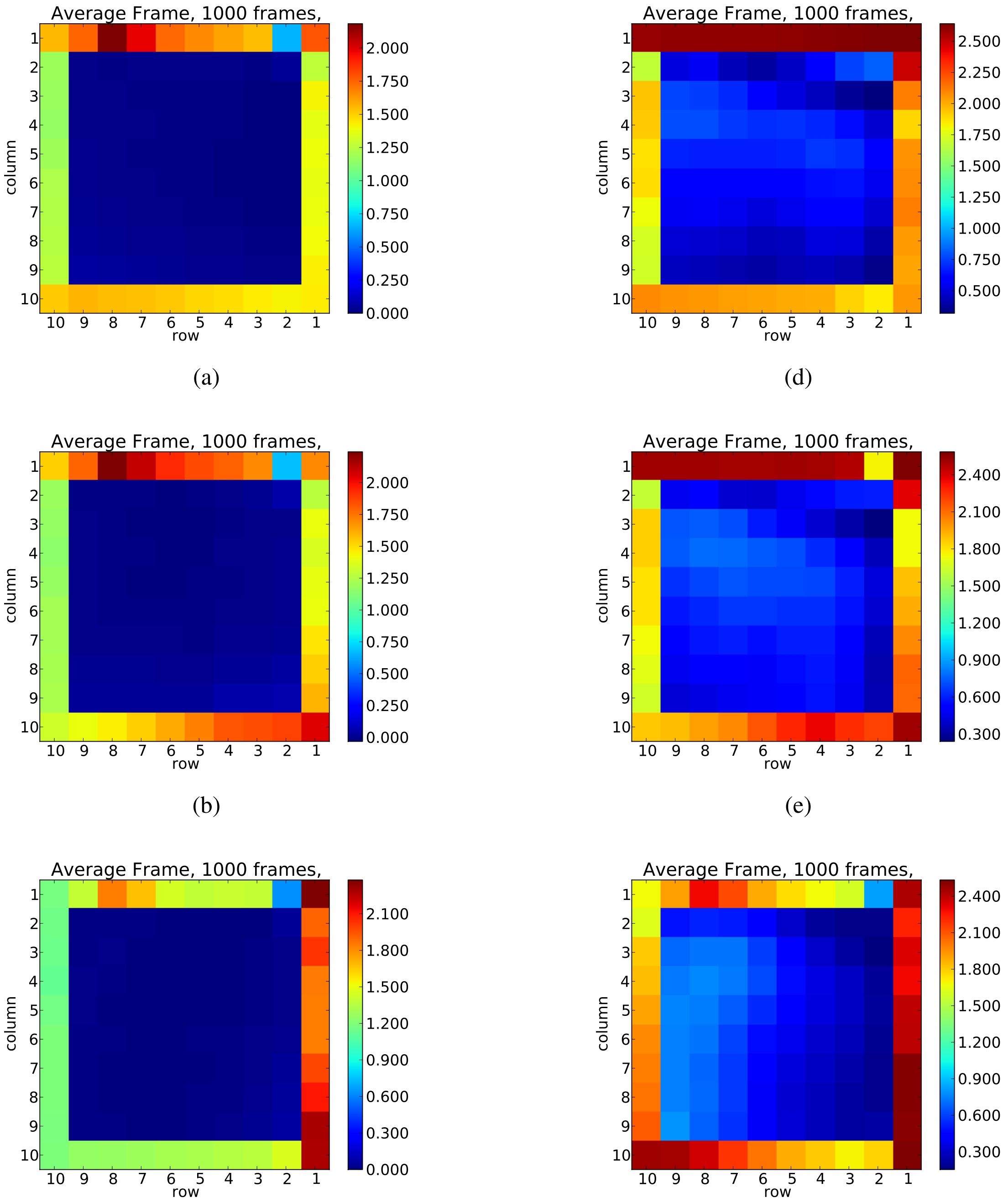


Figure 6.39: Left: Baseline excitation signal for LEDs 1, 2, 3 with 24μA, 24μA, and 48μA current, respectively, C_{min}, and 2500fps. Right: Baseline plus fluorescence signal for LEDs 1, 2, 3 with the same settings.

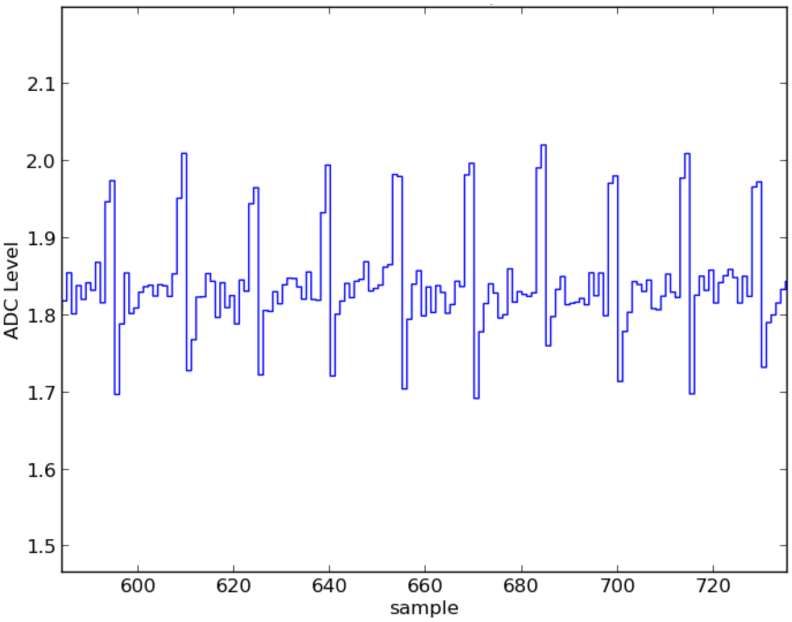
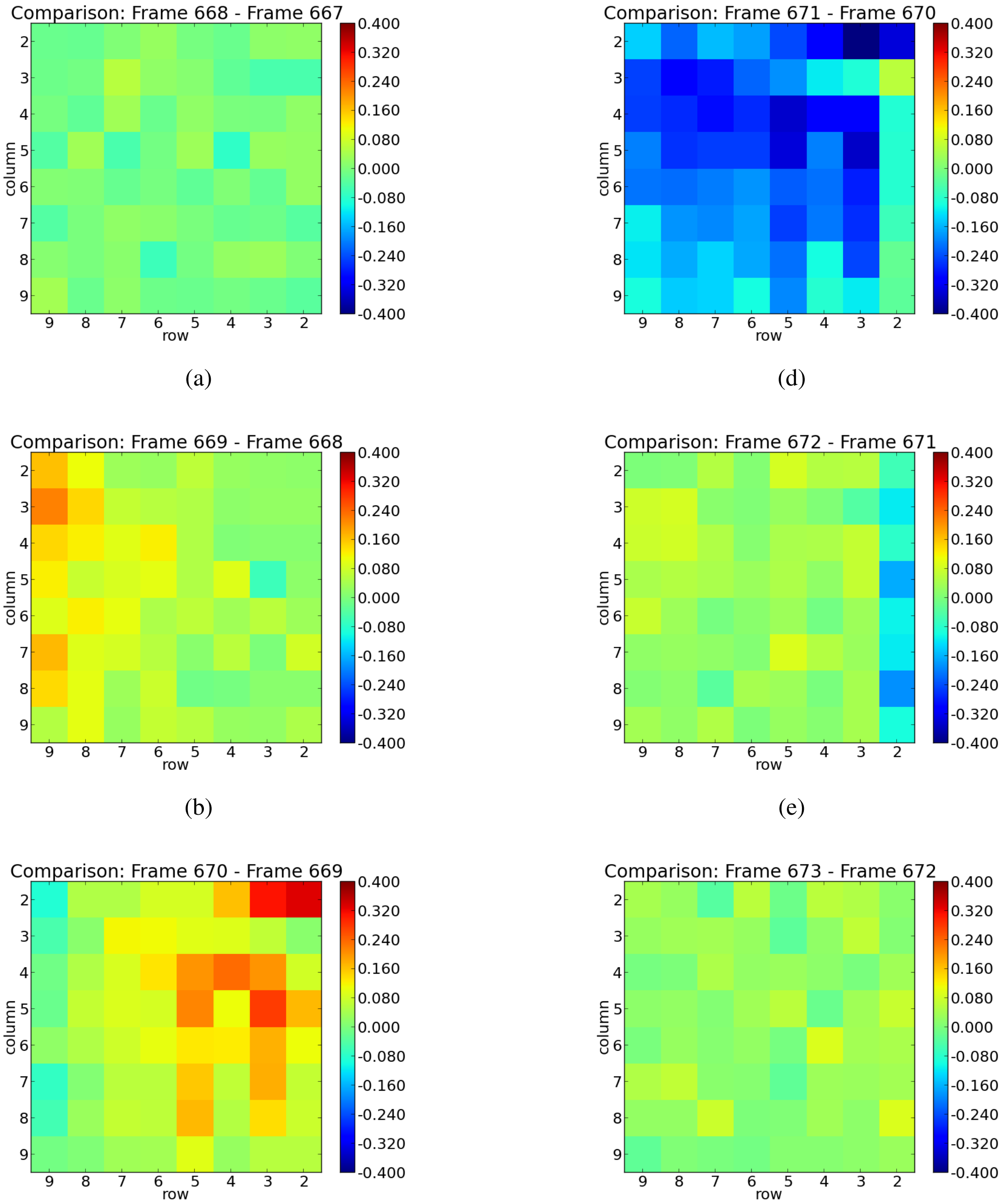


Figure 6.40: Signal from pixel (8,3) due to LED1 light modulated with 1ms sine wave snippets every 5ms.

between frames. Dark red and dark blue indicate large positive and negative signal swings between frames, which indicates the presence of a spike. We see that pixels on the left side of the array detected the spike in (b), while pixels on the right side were detected in frame (c). This is due to the timing of the imager circuitry, which has a rolling shutter. The negative half of the spike is detected in (d). Frames (a), (e), and (f) represent the “resting state”, where the change in signal between frames is due to noise, which results in neutral colors in the plots.



(a)

(d)

(b)

(e)

(c)

(f)

Figure 6.41: Consecutive frame images around a sine wave snippet, where each image is a plot of the current frame minus the previous frame, with C_{min} and 2500fps, for LED1.

6.6 Summary

This chapter detailed the design of a high-sensitivity CMOS imaging chip, as well as an optical neural recording system that included the imaging chip, LED dies and an optical excitation filter packaged together. The system was designed with the speed and sensitivity to detect individual spikes from neurons in layer I of the brain. The performance characterization of the imager chip was presented, followed by full system fluorescence testing with a tissue phantom. The test results indicate that the system functions as intended, and that optical spike detection is indeed feasible.

Chapter 7

Conclusion

In summary, this thesis work was concerned with advancing the state-of-the-art in neural recording devices. The main contributions of this thesis are the modeling, design and testing of an implantable optical neural recording device capable of detecting individual action potentials. The following sections discuss the conclusions for each of these contributions and improvements that could be made through future work.

7.1 Optical Modeling

7.1.1 Contributions

OR1 Model

From the distributed point source optical model introduced in the OR1 work, the following was learned:

- Typically LED radiation patterns are considered from a far-field perspective. Prior to this work it was not clear what kind of light distribution would be seen in a system with dimensions on par with the LED die dimensions, and in particular, below the bottom surface of an unpackaged LED die. It was determined that significant amounts of light from an LED die can exit the die in the downwards direction.
- The testing results for the ED-113SG LED die showed that the distributed isotropic point source MATLAB model was useful for predicting the light distribution in the system.
- In general, such a model can be used to predict the excitation light power incident at different locations within this type of integrated device with near-field lighting conditions. The model can be extended to systems with multiple LEDs, and to predict fluorescent light power as well. The model can provide valuable insight into the design of these systems and guide circuit and system design specifications for fluorescence-detection systems.

OR2 Model

The basic distributed point source model was extended to include multiple LED sources, light attenuation in tissue, fluorescence from multiple neurons and imager response over time. The new model was used to guide design choices for the OR2 chip. The following was learned from the OR2 modeling:

- Most typical imager designs aim to reduce the pixel size as much as possible to increase the imaging resolution. However, in this case larger pixels improve the imaging depth. Additionally, smaller pixels are not necessarily advantageous in this type of contact imaging application, as they may contain a lot of redundant information if they see the same signal from a given neuron. Therefore, pixel size presents a very important tradeoff in these devices.
- For LED dies with a Lambertian light distribution there is a sweet spot for neuron height above the imager surface where the recorded signal is maximized; if the neurons are too close to the imager or too far away, they will not see much excitation light. This sweet spot depends on the lateral distance between the LED and neuron.
- There exists a design tradeoff related to frame rate. At faster frame rates (such as the Nyquist rate, $\sim$15kHz), the received signal is smaller, since there is less time for photocurrent to charge the pixel capacitor. By slowing the frame rate down and allowing a longer integration time (in this case about 2.5kHz), larger signals, or alternatively neurons deeper in tissue, can be seen. Thus, for simply detecting spikes rather than recording their shape, relaxing the requirement of the Nyquist sampling rate is beneficial.

- It is important to note that the optical modeling could be considered pessimistic in terms of light collection efficiency, since the model does not consider what happens to light that does not directly hit a neuron. In reality, scattered light could continue on to excite adjacent neurons, which would help the device performance.

Modeling results indicate that the type of device presented in this work could be used in recording applications that use a reporter like GCaMP3. Such devices may also work for detecting individual action potentials once a sufficient VSFP is created. Alternatively the device could be used with voltage sensitive dyes, keeping in mind that dyes are not appropriate for chronic recording.

7.1.2 Future Work

While the MATLAB modeling presented in this work was useful as a starting point for design, it is possible to create much more accurate optical models by using Monte Carlo-based ray-tracing software, such as FRED, by Photon Engineering [26], or TracePro, by Lambda Research Corporation [19]. With this type of software, a physical system geometry can be defined, using different materials with different types of properties, including biological tissue. The software computes the random paths of individual photons based on their interaction with the system components. Creating such a model should give a much better idea of how much fluorescent light and excitation light (including due to reflections) would arrive at the photodetectors. This would allow for more accurate design specifications.

7.2 CMOS Imaging Chip

7.2.1 Contributions

The key aspect that differentiates the OR2 imager design from the existing literature is that the device is designed specifically for neural recording at frame rates sufficient to detect individual spikes, and not for imaging (i.e. visualizing) the tissue. This resulted in design choices to maximize sensitivity (and imaging depth) to such signals while maintaining the required frame rates.

The pixel pitch in this imager is 4x smaller than that of a typical MEA (400μm). However, a large 100μm x 100μm pixel could see signals from many neurons (based on total neuron density in the brain) if all neurons were fluorescently labeled. The appropriate pixel size would depend on the particular application, and the fraction of neurons that would be labeled by a given VSFP. If only a small fraction of neurons are labeled, then larger pixels would be more appropriate; if a large number of neurons are labeled, then smaller pixels would provide better resolution, though at the expense of imaging depth.

7.2.2 Future Work

There are several changes and/or improvements that could be made to imager.

- Design parameters could likely be improved if a more complete optical model was created, as discussed above. Certainly the pixel size (spatial resolution) would be reconsidered for different system geometries and LED locations, and different LED emission patterns. One way to address the issues discussed above might be to design the imager with smaller pixels, but also provide the capability to combine adjacent pixels into larger pixels. Such programmability would make it possible to tailor the imager to the specific application.

- The reset noise is significant and needs to be significantly reduced in future designs in order to have decent resolution. Future imager circuitry should include some form of double-sampling to help eliminate the reset noise.
- In-pixel (perhaps at the column level) analog circuitry to subtract adjacent samples and find the difference signal may be beneficial.
- Improvements to the imaging circuitry, such as lower noise, higher SNR, and increased fill factor, will always help.
- It is worth investigating methods of asynchronous imaging, which may have the advantage of reduced data rates and reduced power consumption.
- For characterization purposes, it would be useful if the output voltage waveform for the entire integration time, and not just the final sampled voltages, could be seen for all pixels in the array. This would require a change in the digital control circuitry.
- The dark current readings were about 5x higher than expected, likely due to other leakage current in the pixels. Reducing such leakage current is another way to improve sensitivity.

7.3 Optical Neural Recording Device

7.3.1 Contributions

The key contribution of this work is a functioning optical neural recording device based on fluorescence detection. To the best of the author's knowledge, this is the first integrated device designed specifically for detecting individual spikes in an implant.

An absorption filter made from theatre gels was used to prevent excitation light from hitting the imager. An interference filter could achieve a higher optical density, but only for light incident perpendicular to the filter surface. The absorption filter used here was sufficient, and seems to be a good choice for future designs.

The fabricated device did not match the intended system geometry due to an error in chip dicing that resulted in the LED dies being located about $400\mu m$ further from the imager than they should have been. However, the device was still functional. When tissue attentuation is introduced, this could be more of an issue.

7.3.2 Future Work

- In this type of system, uniform delivery of excitation light to all areas of the tissue is desirable, as is targeted excitation. Micro-LED arrays may be a nice solution for this. One method would be to flip-chip bond a micro-LED array, such as those used in [36], to the imager chip. Another option would be to deposit organic LEDs on the imager chip through CMOS post-processing steps, as was done in [81] or [82]. This type of integration may make the inclusion of the optical filter more difficult, and that would need to be considered.

- Improved optical modeling would provide much better information on how much excitation light is expected to hit the photodetectors, and at what angle of incidence the light will hit. Such information would be very useful for specifying the requirements for the excitation filter, including whether or not an interference filter is appropriate, and what the optical density should be. A dichroic (interference) filter may be a good choice because it would reflect incident excitation light back into tissue, so that the excitation light could be "re-used" to excite fluorophores, rather than absorbed and converted to heat, as in an absorption filter.

- Another improvement would be to deposit the filter directly on the imager chip during the fabrication process, which would likely reduce the device thickness, eliminate extra packaging costs, and eliminate defects that may result from the filter attachment process.

- Longer wavelengths, into the near-infrared region, for both excitation and emission light may be advantageous in making deeper imaging possible, though increased scattering would have to be considered. A key in this regard is VSFP design. Improved VSFPs (ideal wavelengths, higher sensitivity, etc.) could go a long way to making these devices even more feasible.

- Signal processing such as that done with compressed sensing/sampling algorithms as in [20] may be beneficial in improving spatial resolution.

- The details of encapsulation must be considered in future device designs.

7.4 Device Testing

7.4.1 Contributions

An important contribution of this work is the use of a fluorescent tissue phantom for device testing. The reasons to use a tissue phantom are: 1) it alleviates the need for a wet lab and all of the procedures required to test with real tissue, and 2) it allows for repeatable experiments where the system parameters are known. This is appropriate for initial testing of new prototypes. Such a tissue phantom can be created in a matter of hours, and the properties of the tissue phantom, such as thickness, optical properties, and fluorophore concentration can be tuned as needed for the application.

Because relatively little excitation light hits the imager due to system geometry and the optical filtering, LED intensity can be modulated with sine/spike waveforms to emulate the action of VSFPs for temporal device testing.

7.4.2 Future Work

Future tissue phantoms that incorporate actual optical properties of tissue would be a very useful step up. It may also be useful to incorporate fluorescent microspheres of different colors.

Once an optimized prototype has been tested with a tissue phantom, then further testing with real tissue would be appropriate.

7.5 Feasibility of Fully Implantable Optical Recording

The results of this work indicate that fully implantable optical neural recording devices could be feasible, and should be investigated further. The device in this work was able to detect spike shapes from the fluorescent signal, as designed. Future work to further answer this question must include improved device design and testing with real tissue with VSFPs. The achievable imaging depth in an optimized device may dictate the applications where the device can be used. The target recording site in this work was layer I of the brain. This was chosen as a first step in exploring the feasibility of these devices. It should be noted that there are applications where for the human brain recording in deeper layers is desired. However, there are also other potential recording sites, such as the spinal cord, where these devices might be useful. The baseline results in this thesis can be used as a starting point for investigation of specific applications to target in future work.

7.6 Other Work

There are other design areas that must be tackled to produce a fully implantable device:

- Integration with wireless data transmission and power sources.
- LED power consumption and tissue heating effects must be investigated further.
- For a real implant it may be advantageous to use a flexible PCB instead of a conventionally packaged devices. Eliminating the chip package would allow the tissue to be as close as possible to imager (though LED emission patterns would dictate how important that is).

7.7 Other Applications

Fluorescence detection is used in many other biomedical applications, such as in lab-on-a-chip systems. Lab-on-a-chip systems also require a high level of integration in order to reduce their cost. Integration of LED dies with an imager chip, as well as the optical modeling presented in this work could be applied to these types of applications.

APPENDICES

Appendix A

Pixel Noise Analysis

The noise analysis for the pixel is split into two phases: the reset phase and the integration phase. Figs. A.1 and A.2 show the circuits used to calculate the noise of the integration phase and reset phase, respectively. In the integration phase, each noise source is referred to the pixel output (V_{out}). In the reset phase, each noise source is referred to the input node (V_{pd}), and then multiplied by the gain to get the equivalent output noise.

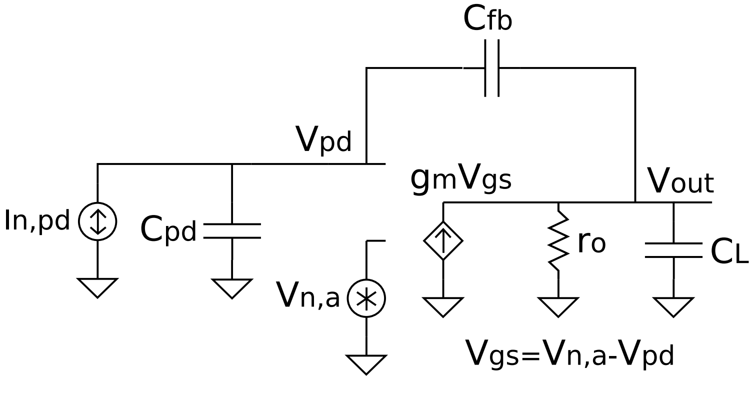


Figure A.1: Pixel integration noise circuit.

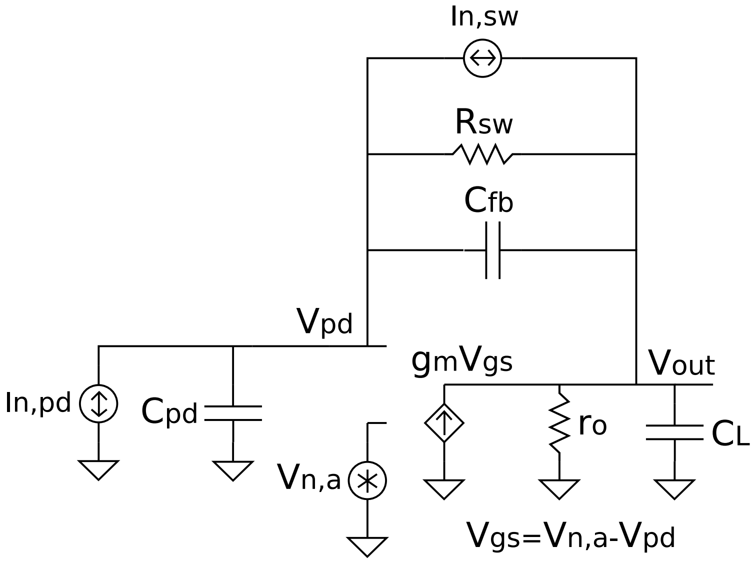


Figure A.2: Pixel reset noise circuit.

A.1 Amplifier Input-Referred Noise

Fig. A.3 shows the telescopic amplifier circuit used to find the input-referred noise (the amplifier noise is referred to the non-inverting input).

In general, a MOSFET has the following noise power spectral density (PSD) for thermal and flicker noise:

$$V_i^2(f) = 4kT\frac{2}{3}\frac{1}{g_m} + \frac{K}{WLC_{ox}f}\left[\frac{V^2}{Hz}\right], \tag{A.1}$$

$$I_i^2(f) = 4kT\frac{2}{3}g_m + \frac{K}{WLC_{ox}f}g_m^2\left[\frac{A^2}{Hz}\right]. \tag{A.2}$$

The cascode devices and the tail current device do not contribute significant noise to the amplifier. The input devices and the load current source devices are the significant noise contrib-

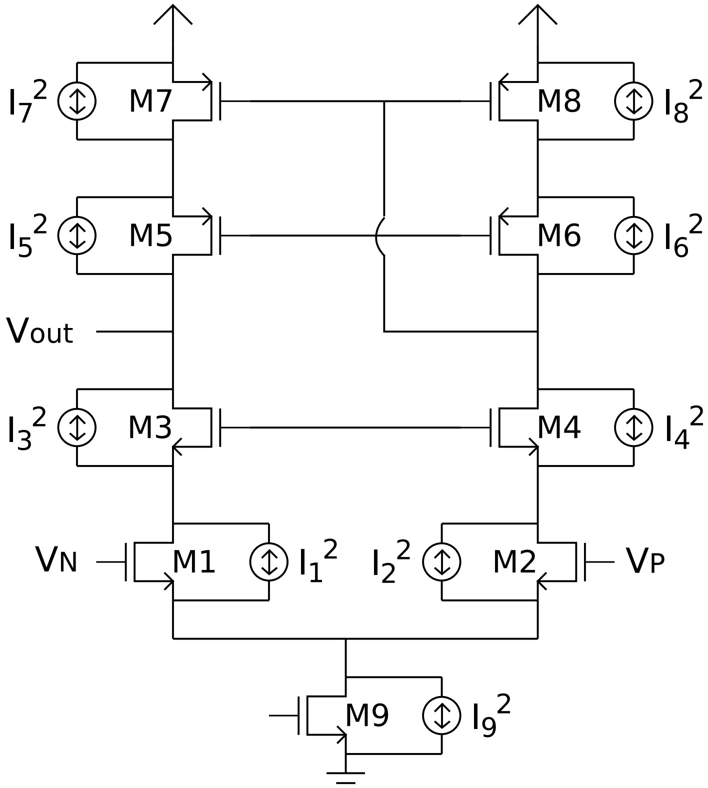


Figure A.3: Telescopic amplifier noise circuit.

utors. The noise current of each of these devices is in parallel with each device (drain to source). These noise currents flow directly to the output. Therefore, the output noise current is given by

$$I_{o,n}^2 = I_1^2 - I_2^2 + I_7^2 - I_8^2 \ . \tag{A.3}$$

We can refer the output noise current to a corresponding input voltage by dividing by the transconductance of the amplifier:

$$V_{i,n}^2 = \frac{I_{o,n}^2}{g_{m1}^2} = \left(\frac{8kT}{3g_{m1}} + \frac{K_1}{W_1 L_1 C_{ox} f}\right) + \left(\frac{8kT}{3g_{m1}} + \frac{K_2}{W_2 L_2 C_{ox} f}\right) + \left(\frac{8kT g_{m7}}{3g_{m1}^2} + \frac{K_7 g_{m7}^2}{W_7 L_7 C_{ox} f g_{m1}^2}\right) + \left(\frac{8kT g_{m8}}{3g_{m1}^2} + \frac{K_8 g_{m8}^2}{W_8 L_8 C_{ox} f g_{m1}^2}\right) . \tag{A.4}$$

Assuming M7 = M8 and M1 = M2,

$$V_{i,n}^2 = \left(\frac{16kT}{3g_{m1}} + \frac{2K_1}{W_1 L_1 C_{ox} f}\right) + \left(\frac{16kT g_{m7}}{3g_{m1}^2} + \frac{2K_7 g_{m7}^2}{W_7 L_7 C_{ox} f g_{m1}^2}\right) . \tag{A.5}$$

Now given that

$$g_m \propto \frac{\sqrt{I_D}\sqrt{W}}{\sqrt{L}} \ , \tag{A.6}$$

and neglecting constants, we can write an equation that is a function of only parameters that we directly control:

$$V_{i,n}^2 \propto \frac{\sqrt{L_1}}{\sqrt{I_{D1}}\sqrt{W_1}} + \frac{1}{W_1 L_1} + \frac{\sqrt{I_{D7}}\sqrt{W_7}}{\sqrt{L_7}} \frac{L_1}{I_{D1} W_1} + \frac{W_7 L_1}{L_7 W_1} \ . \tag{A.7}$$

This equation tells us that to minimize the noise we should do the following:

- increase current (but the effect on the noise bandwidth must also be considered)
- at a given current, scale up W_1, scale down L_1 (if thermal noise is greater than flicker noise)
- at a given current, scale down W_7, scale up L_7

The transfer function for the integration (read) noise of the amplifier is

$$\frac{V_{out}}{V_{n,a}} = \frac{-gm(C_{pd} + C_{fb})r_o}{C_{fb}(sC_{fb} - gm)r_o - (1 + sr_oC_L - sr_oC_{fb})(C_{pd} + C_{fb})} \; . \tag{A.8}$$

The transfer function for the reset noise of the amplifier is

$$\frac{V_{pd}}{V_{n,a}} = \frac{g_m}{g_m + sC_{pd} + \left(\frac{1}{r_o} + sC_L\right)\left(\frac{sC_{pd}R_{sw}}{1+sC_{fb}R_{sw}} + 1\right)} \; . \tag{A.9}$$

A.2 Reset Switch Noise

The transfer function for the reset switch noise to the input (V_{pd}) during reset is

$$\frac{V_{pd}}{I_{n,rst}} = \frac{-sC_L - 1/r_o}{s^2C_{pd}C_L + s(C_{fb}g_m + \frac{C_{pd}+CL}{R_{sw}}) + (\frac{1}{r_o} + g_m)\frac{1}{R_{sw}}} \; . \tag{A.10}$$

A.3 Photocurrent and Dark Current Read Shot Noise

The transfer function for all types of current shot noise at the input node (V_{pd}) during integration is:

$$\frac{V_{out}}{I_n} = \frac{r_o(g_m - sC_{fb})}{(1 + sr_o(C_L + C_{fb}))s(C_{pd} + C_{fb}) - sr_oC_{fb}(sC_{fb} - g_m)} \; . \tag{A.11}$$

The sources of current shot noise are dark current, excitation light current and signal (fluorescent) current. The noise spectral density due to these sources is given by:

$$i_{shot}^2 = 2q(I_{dark} + I_{ex} + I_{fl}) \ , \tag{A.12}$$

where q is the charge of an electron, $1.6 * 10^{-19}C$.

The dark current shot noise is due to the dark current in the photodiode, which depends on the area of the photodiode and temperature. According to design kit documentation, the dark current at $27°C$ for a 100x$100\mu m^2$ photodiode (opto process option) is 8fA, regardless of the reverse bias voltage on the photodiode. The excitation light current refers to excitation light from the LEDs that manages to make it through the optical filter and hit the photodiodes. This is expected to be very low. The signal current refers to the fluorescent light from fluorophores that reaches the photodiode, and is the signal of interest. Note that the shot noise values related to the fluorescence and excitation signals will be different for different pixels, because different pixels will be seeing different amounts of light.

A.4 Photodiode Thermal Noise

Photodiode thermal noise is the noise generated by the shunt resistance [51], and is given by:

$$i_{pd}^2 = \frac{4kT}{R_{pd}} \ . \tag{A.13}$$

The photodiode shunt resistance R_{pd} “is defined to be the inverse slope of the I-V curve at the 0V point" [94]. “The generally accepted practice in industry is to measure the current I at

V = 10mV, and apply Ohm's law to find the shunt resistance" [94]. R_{pd} can be found through Ohm's law by taking the reverse-bias diode voltage and dividing by the leakage current. But a better way is to average the results for -10mV and +10mV. However, in simulation a DC sweep of the reverse-bias voltage can be performed, and the current plotted. Because the current is so small, the simulator parameters gmin and iabstol are reduced to 1e-15, and the voltage is swept over a small range so there are enough points around 0V to get a nice curve. For the nwd (N-well diode) model in the design kit (90μm x 90μm photodiode) this gives:

$$R_{pd} = \frac{1mV - (-1mV)}{1.298fA - (-1.354fA)} = 7.54e11\Omega = 754G\Omega \ . \tag{A.14}$$

This gives i_n^2 = 2.196e-32 A^2/Hz, and $V_{o,nrms} = 706\mu V$.

A.5 Output-Referred Noise Spectral Density

Fig. A.4 shows a plot of the output-referred noise spectral density of each noise source, with the values from Table A.1. Note that in this plot only the dark current is considered (I_{ex} and I_{fl} are set to 0, as they could vary widely).

Parameter	Value
g_m	$28\mu A/V$
r_o	$26.6M\Omega$
R_{sw}	3726Ω
C_{pd}	7.3182pF
C_{fb}	4.6688fF
C_L	0.95353pF
T	300K
I_{dark}	38fA

Table A.1: Parameters used to plot output-referred noise spectral density of Fig. A.4.

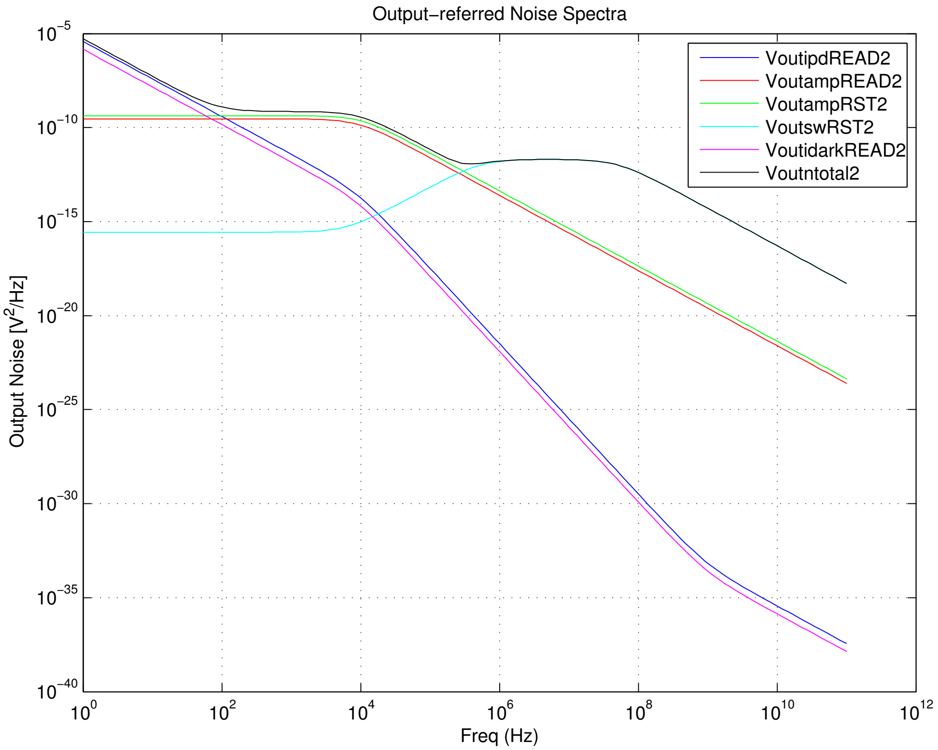


Figure A.4: Output-referred noise spectra of the pixel.

A.6 Integrating Noise Over the Bandwidth of Interest

Table A.2 shows the output RMS noise voltage of each pixel noise source, integrated from 0 to ∞, for each value of C_{fb}. The photocurrent shot noise is calculated assuming a photocurrent of 0.1pA, which is the minimum detectable signal with C_{min} and 2.5kSps frame rate; this minimum detectable current gives the worst case SNR. Summing all of the output noise PSDs and integrating from 0 to ∞, the worst-case ($C_{fb,min}$) total RMS output noise due to all sources is 13.3mV. The majority of this noise is due to the switch reset noise, which is 12.7mV$_{rms}$.

However, the signal of interest in this work is the difference in output voltage in a given pixel between adjacent samples. Because the sampling frequency is 2.5kHz, lower frequency noise would not show up in these difference calculations. Therefore, the lower limit of integration can be higher than 0Hz. If a value of 1kHz is used, the RMS output noise is 13.1mV. This value is nearly the same as before, because the switch thermal noise is dominating, and eliminating 1kHz of white noise from a wide-band noise spectrum is insignificant.

Noise Source	$C_{fb,min}$	$C_{fb,mid}$	$C_{fb,max}$
VonswRST	12.7mV	3.7mV	0.7mV
VoampRST	2.7mV	1.6mV	0.8mV
VoutampREAD	2.1mV	1.6mV	0.9mV
VoutipdREAD	2.0mV	1.1mV	0.3mV
VoutidarkREAD	1.2mV	0.7mV	0.2mV
Voutntotal	13.3mV	4.5mV	1.4mV

Table A.2: Output-referred RMS noise voltage for each pixel noise source.

In order to reduce the switch reset noise, according to Eq. A.10, we can try increasing g_m, or r_o, or reducing R_{sw} or C_L (assuming we cannot change C_{pd} or C_{fb}). However, changing these parameters affects the other noise contributors as well. According to the models, reducing C_L

is the most effective way to reduce the noise. However, C_L is largely defined by the column amplifier input capacitance. The column amplifier input devices are sized relatively large for noise and offset performance, and can't really be reduced. Reducing R_{sw} actually has minimal impact on the switch reset noise, but a significant negative impact on the other noise sources. Increasing r_o and/or g_m also does not help. Thus, there is not much that can be done to improve the noise performance of the pixel simply by adjusting device sizes (and not significantly affecting the fill factor). Other methods for reducing the effects of reset noise, such as correlated double-sampling, could be used.

Appendix B

SAR ADC Design

An 8-bit SAR ADC was designed to be used in the OR2 imager chip, as well as in future electrical recording chips. The requirements for the imager chip were more stringent, so they were used for the design specifications.

Given a 100 pixel imager sampled at 10kSps, the required ADC sampling rate is 1MSps. The ADC uses 10 cycles per conversion/sample, so the required clock rate is 10MHz. The target sampling rate is 2.5kSps, so this 10kSps specification should be more than sufficient.

The ADC uses a binary-weighted capacitive charge-sharing DAC, a basic sample and hold circuit, and a clocked comparator. Fig. B.1 shows a block diagram of the ADC design. The ADC is designed to operate with a rail-to-rail reference range; operating over 3.3V gives $V_{LSB} = 12.89$mV.

At the start of the conversion cycle, DACRESET is high, and the DAC is reset. V_{in} is sampled by the S/H in the sample phase, and the S<7:0> signal is set to 01111111. DACRESET then

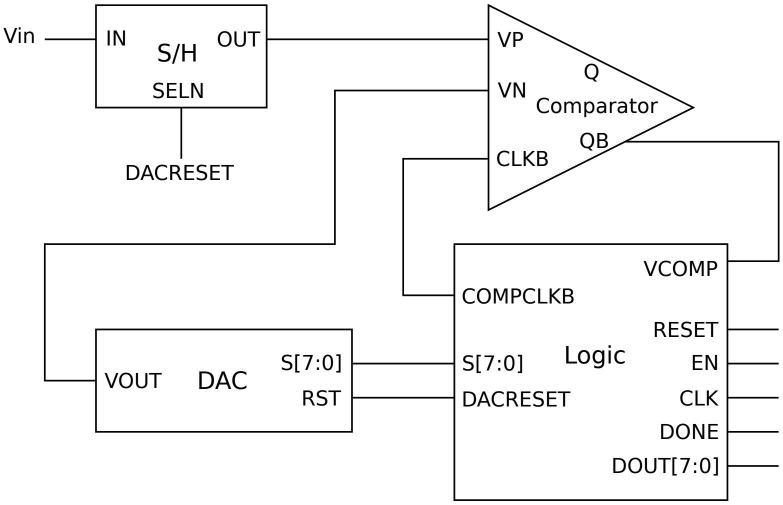


Figure B.1: SAR ADC block diagram.

goes low and V_{in} is held. For the 8 comparison cycles DACRESET stays low, and the value of VCOMP determines the next S setting. When the conversion is complete, DONE goes high, the result is written to DOUT<7:0>, and DACRESET goes high.

DACRESET (and DACRESETB, its complement, not shown here) is used as the sampling signal for the S/H switch. DACRESETB is also used to gate the clock signal going into the comparator (not shown here). VCOMPB, the complemented output from the comparator, is actually the signal used by the SAR logic (VCOMP input).

An important consideration with the timing is that the comparator output is properly read by the SAR logic before the comparator is reset. Therefore, the comparator clock signal is a slightly delayed version of the main clock. Some delay cells were also added at the comparator outputs to help with this.

B.1 DAC

B.1.1 Operation

Fig. B.2 shows an example 3-bit binary-weighted capacitor DAC. C_p is any parasitic and load capacitance (such as input capacitance of the comparator) on the DAC output. C is the unit capacitance.

In the RESET phase, the RST switch is closed, grounding the output (to V_{refL}), and the other switches are switched to V_{refL} as well. Both terminals of all capacitors are grounded, discharging all capacitors and V_{out}.

In the OUTPUT phase, the RST switch is opened, allowing the output to float. The capacitor switches are set according to the digital input values. If the input is a 1, the switch connects the capacitor to V_{refH}.

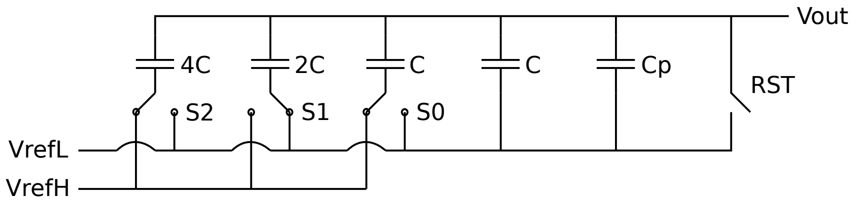


Figure B.2: Example 3-bit binary-weighted capacitor DAC.

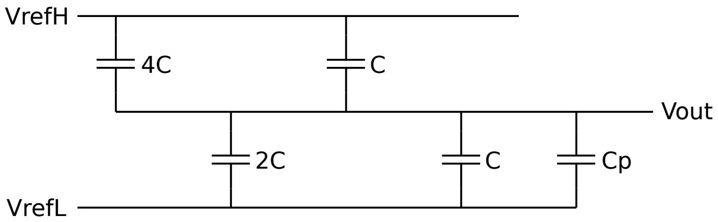


Figure B.3: Example 3-bit DAC redrawn with switches set as in previous figure.

As becomes clear in the the equivalent circuit of Fig. B.3, with the digital inputs set to 101, a capacitive divider is formed between V_{refH} and V_{refL}, and the output voltage is given by:

$$V_{out} = \frac{(4C + C)(V_{refH} - V_{refL})}{(4C + C) + (2C + C) + C_p} = \frac{C_{eq}(V_{refH} - V_{refL})}{C_{eq} + (2^N C - C_{eq}) + C_p} = \frac{C_{eq}(V_{refH} - V_{refL})}{C_{total} + C_p} \tag{B.1}$$

$$V_{out} = \frac{5C}{8C + C_p}(V_{refH} - V_{refL}) \tag{B.2}$$

Note that the digital inputs can change and the output voltage will change without the need for reset. i.e. in the SAR ADC the DAC is only reset at the start of a conversion, not in every cycle.

B.1.2 Design

Fig. B.4 shows the full 8-bit DAC schematic. The unit capacitance value is C = 15.14fF. In each switch, a single NMOS is used to connect VrefL to the capacitor, and a single PMOS is used to connect VrefH to the capacitor. Both transistor gates are controlled by the same switch signal. Note that for S6 each transistor width is 2x unit size (2 fingers), and for S7 each transistor width is 4x unit size (4 fingers). This was required to speed up the response for the most significant bits with the largest changes in voltage.

B.1.3 Layout

In schematic simulation the DAC works beautifully, because the ratios of the capacitors are perfect (disregarding load capacitance, which introduces gain error). However, when simulating an

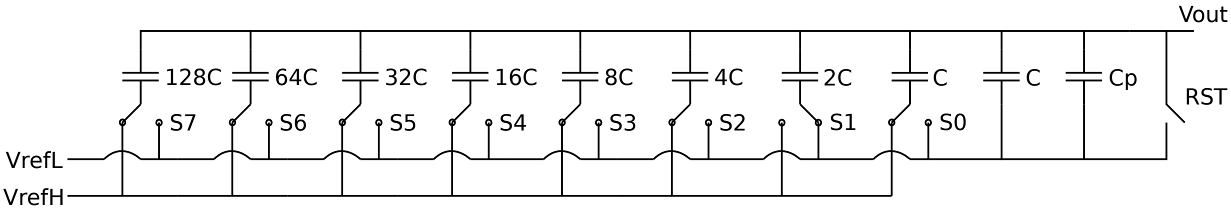


Figure B.4: 8-bit charge-sharing DAC schematic.

extracted version of the DAC layout, it is clear that parasitics can cause significant non-linearities. Therefore, an iterative process of adjusting the layout and simulating the DAC performance was undertaken in order to reduce the non-linearities.

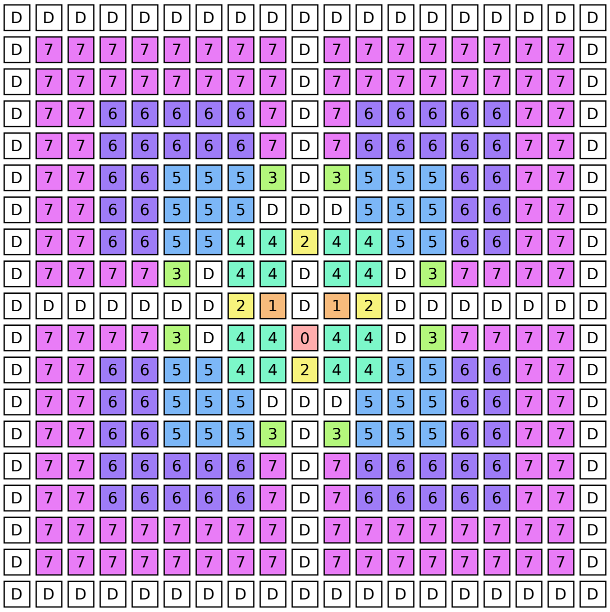
The unit capacitor placement for minimizing not only mismatch due to process variation, but also parasitics, discussed in [49], was considered here. The paper includes a diagram showing the placement of each unit capacitor, which by eye looks essentially random. This placement was tried here, and was manually routed. The extracted simulation results were not good; there was significant gain error and non-linearity. With this type of placement an automatic routing algorithm that takes the parasitics into account is required.

Therefore, a more conventional common centroid placement, shown in Fig. B.5, which is much easier to route, was used. Given the regular routing patterns it is much easier to see how the routing affects the parasitics, and to come up with tweaks to optimize it. Therefore, the extracted simulation results are much better.

The simulated INL (integral non-linearity) and DNL (differential non-linearity) graphs in Figs. B.7 and B.8 show the improvement in INL and DNL due to layout tweaks between the first, second last (v4) and last (v5) design iterations.

Fig. B.6 shows the final layout of the DAC capacitor array. The layout is an array of 19 x 19 unit capacitors, including a ring of dummy capacitors around the outside and other dummy devices in central locations, for matching. A SKILL script was used to read in a text file describing the placement and to place the unit capacitors, as well as some wiring, in the layout.

Each unit capacitor is $4\mu m$ x $4\mu m$. The array has a large N-well underneath it to shield it from the substrate. The N-well is contacted around the perimeter by an N+ guard ring (to VDD). A P+ guard ring (to GND) surrounds the N-well. The dimensions of the entire layout are $230.9\mu m$ x $230.9\mu m$ (not including the wires extending down from the bottom).



D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
D	7	7	7	7	7	7	7	7	D	7	7	7	7	7	7	7	7	D
D	7	7	7	7	7	7	7	7	D	7	7	7	7	7	7	7	7	D
D	7	7	6	6	6	6	6	7	D	7	6	6	6	6	6	7	7	D
D	7	7	6	6	6	6	6	7	D	7	6	6	6	6	6	7	7	D
D	7	7	6	6	5	5	5	3	D	3	5	5	5	6	6	7	7	D
D	7	7	6	6	5	5	5	D	D	D	5	5	5	6	6	7	7	D
D	7	7	6	6	5	5	4	4	2	4	4	5	5	6	6	7	7	D
D	7	7	7	7	3	D	4	4	D	4	4	D	3	7	7	7	7	D
D	D	D	D	D	D	D	2	1	D	1	2	D	D	D	D	D	D	D
D	7	7	7	7	3	D	4	4	0	4	4	D	3	7	7	7	7	D
D	7	7	6	6	5	5	4	4	2	4	4	5	5	6	6	7	7	D
D	7	7	6	6	5	5	5	D	D	D	5	5	5	6	6	7	7	D
D	7	7	6	6	5	5	5	3	D	3	5	5	5	6	6	7	7	D
D	7	7	6	6	6	6	6	7	D	7	6	6	6	6	6	7	7	D
D	7	7	6	6	6	6	6	7	D	7	6	6	6	6	6	7	7	D
D	7	7	7	7	7	7	7	7	D	7	7	7	7	7	7	7	7	D
D	7	7	7	7	7	7	7	7	D	7	7	7	7	7	7	7	7	D
D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D

Figure B.5: DAC capacitor array common centroid placement.

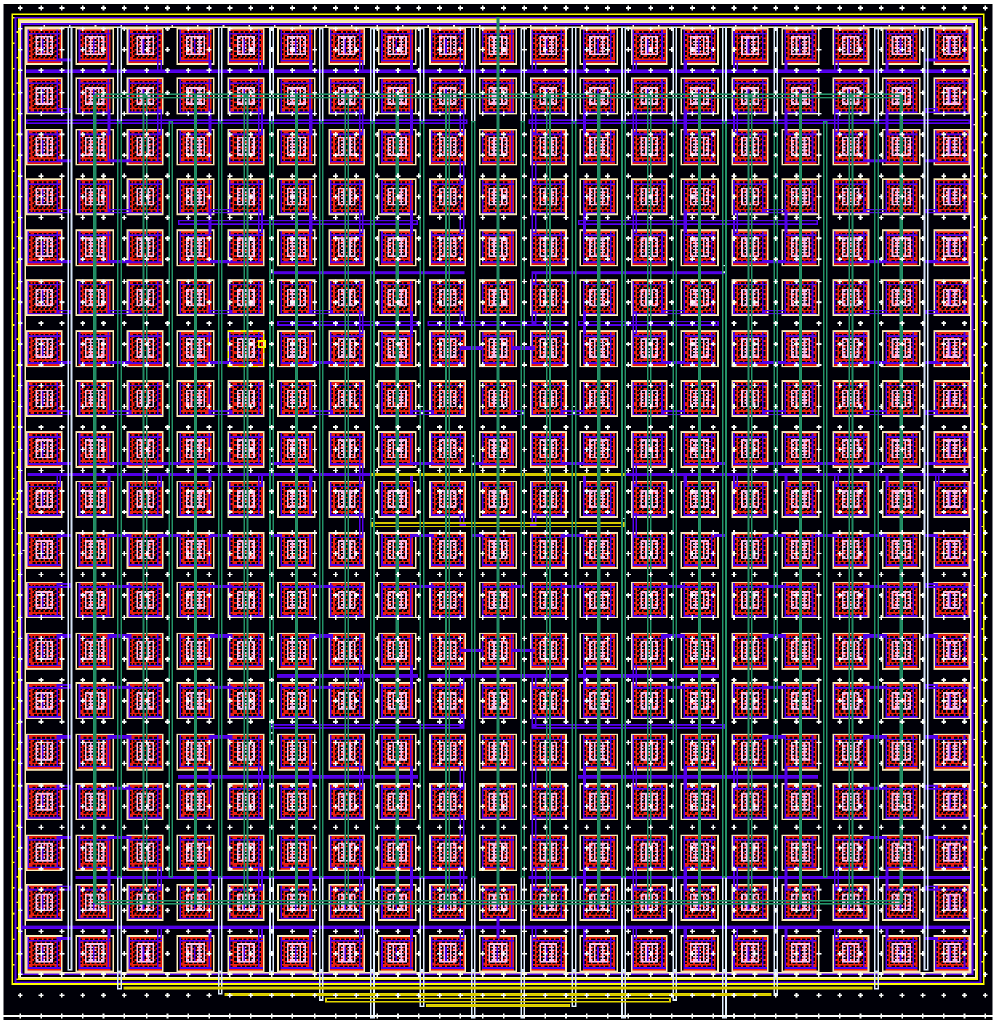


Figure B.6: DAC capacitor array layout.

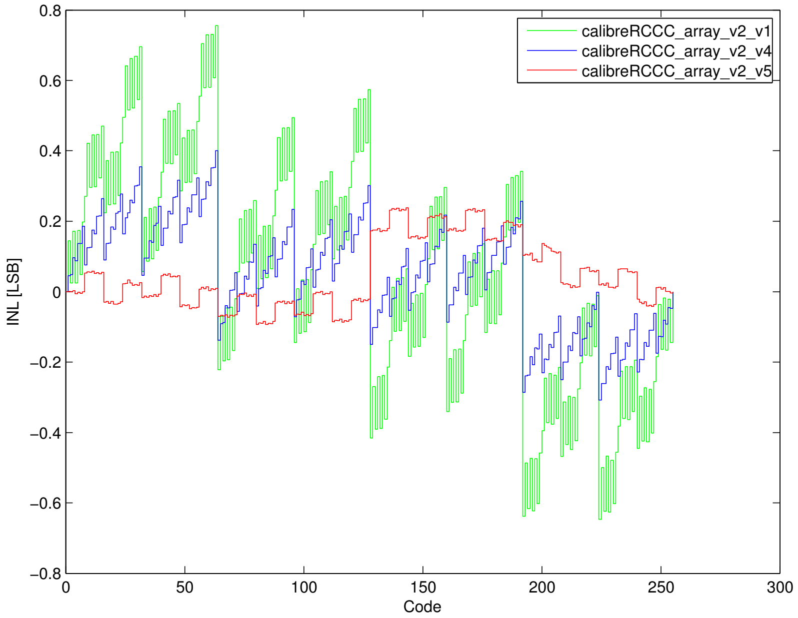


Figure B.7: DAC INL plot.

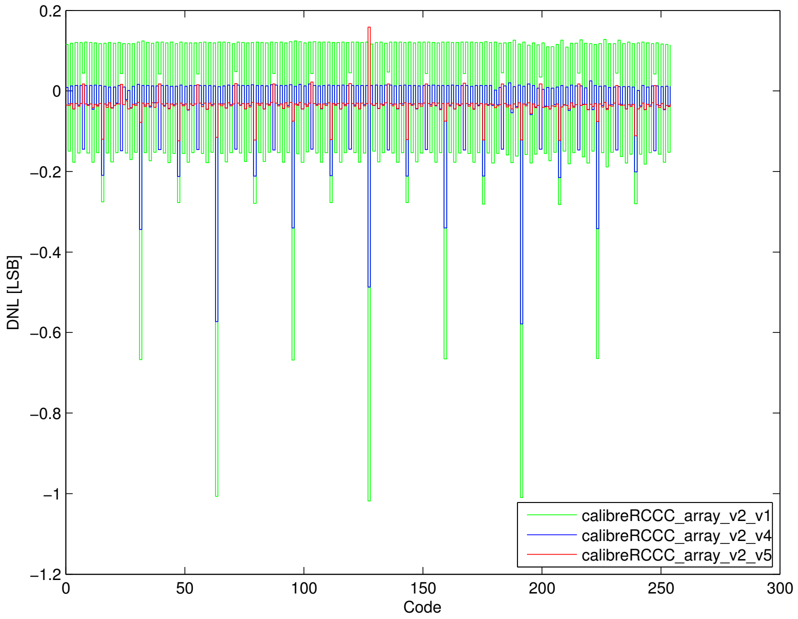


Figure B.8: DAC DNL plot.

B.2 Comparator

B.2.1 Specifications

The 8-bit ADC operating over 3.3V has $V_{LSB} = 12.89$mV. The comparator needs to meet that resolution. The goal is to have the comparator input offset voltage V_{IOS}, input-referred noise $V_{N,RMS}$, and kickback noise all less than 0.5 LSBs, or less than about 6mV. The other main specification is the speed. The comparator has half a clock cycle to settle and produce the correct output. Given the target 10MHz clock frequency, the comparator has 100/2 = 50ns to arrive at an answer.

B.2.2 Design

Fig. B.9 shows the rail-to-rail clocked comparator schematic, based on the design from [15]. When CLK is high it is in the reset mode (VON and VOP are pulled low) and when CLK is low it is in the regeneration mode (output is produced). Thus the output changes after the falling clock edge. When VP > VN, Q = VDD. When VP < VN, Q = GND. This design has good kickback noise performance due to the current mirror pre-amp approach.

VBIAS1 and VBIAS4 are generated by simple current mirrors (not shown). The currents of first stage complementary differential pairs are summed and mirrored to the second stage. Nx represents N extra inverters to implement necessary delay required for robust ADC operation. In this case, there is one inverter on each of VON and VOP, followed by 8 extra inverters each.

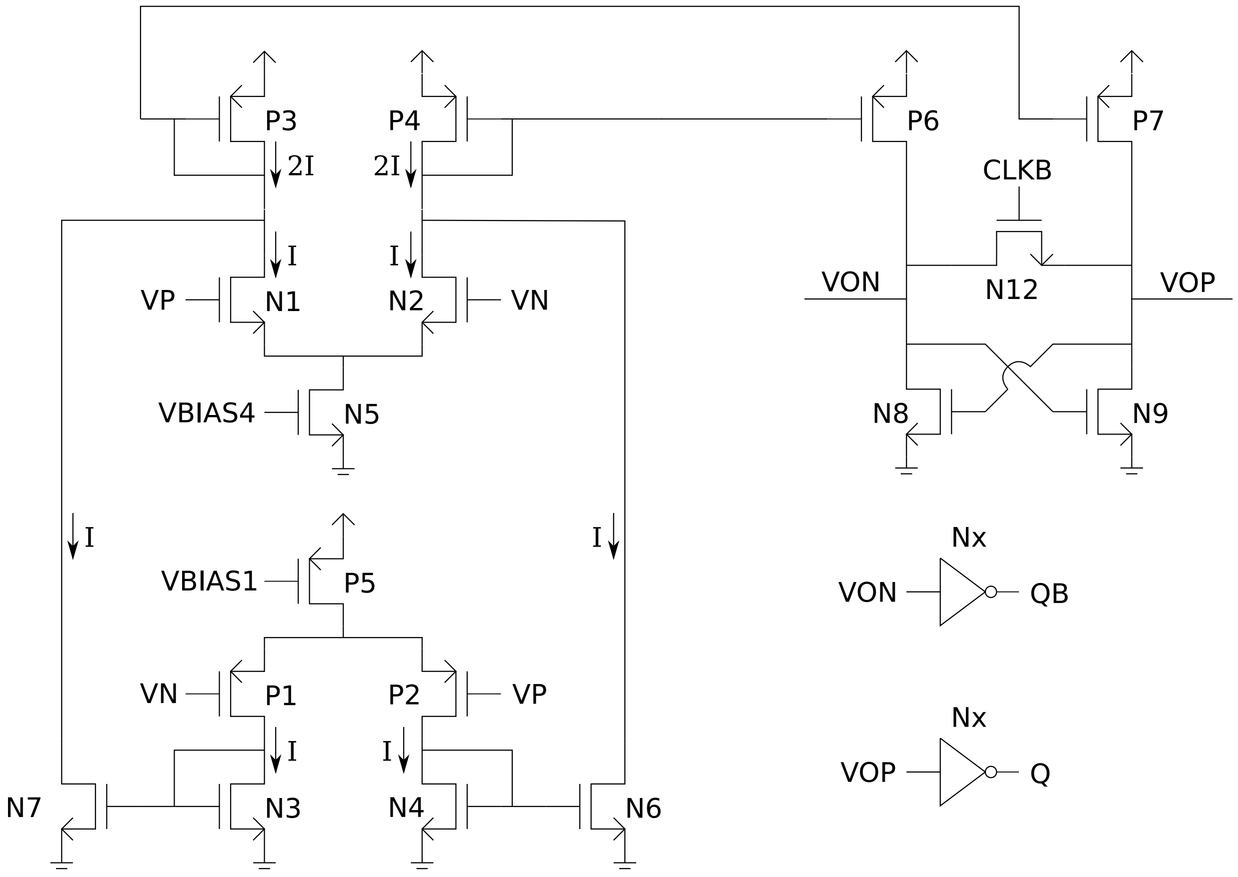
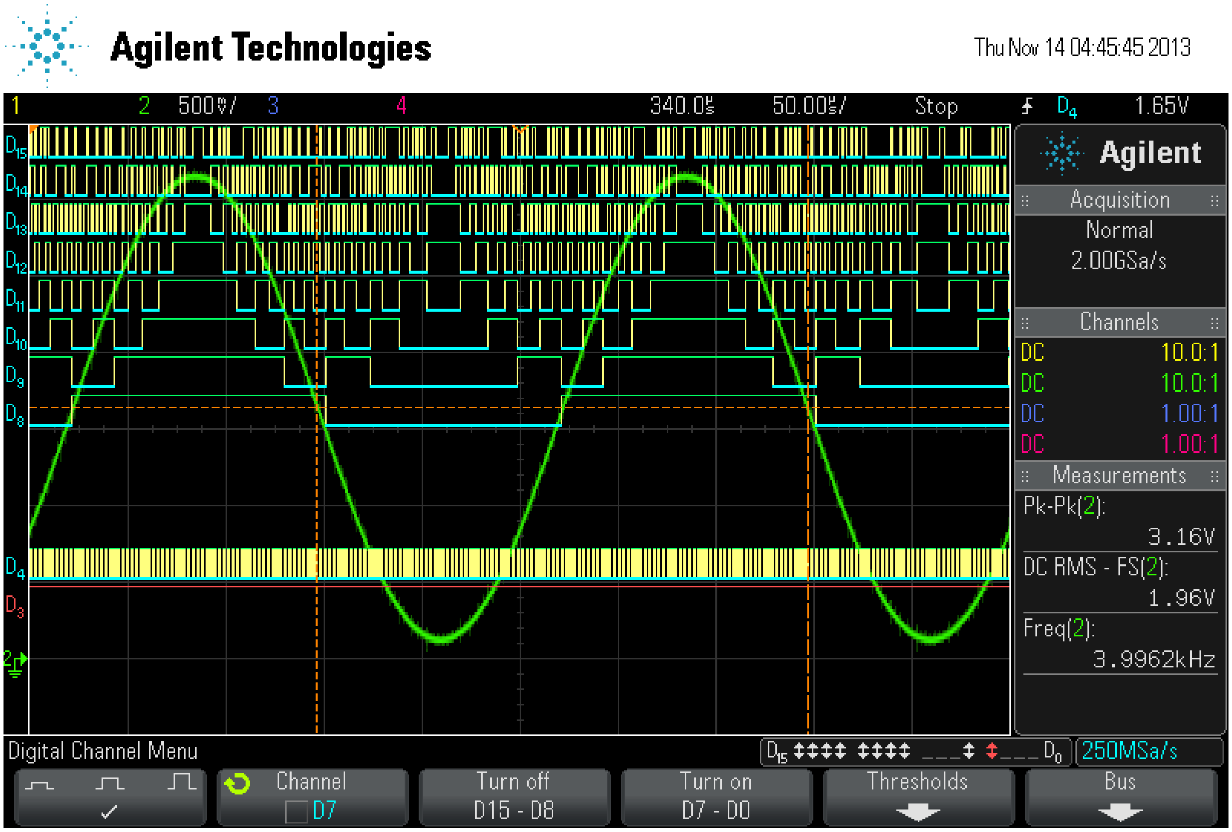


Figure B.9: Comparator schematic.

B.3 ADC Performance

Fig. B.10 shows a scope capture of the ADC output. This is sampling a 4kHz sine wave at 800kSps. Fig. B.11 shows the ADC performance from Cadence simulation of the schematic. Fig. B.12 shows the measured performance at the default imager sampling rate of 250kSps. Fig. B.13 shows the measured performance near the maximum sampling rate of 800kSps (maximum is 830kSps).



Acquisition
Normal
2.00GSa/s

Channels
DC 10.0:1
DC 10.0:1
DC 1.00:1
DC 1.00:1

Measurements
Pk-Pk(2): 3.16V
DC RMS - FS(2): 1.96V
Freq(2): 3.9962kHz

Figure B.10: ADC output scope capture. 4kHz sine wave sampled at 800kSps.

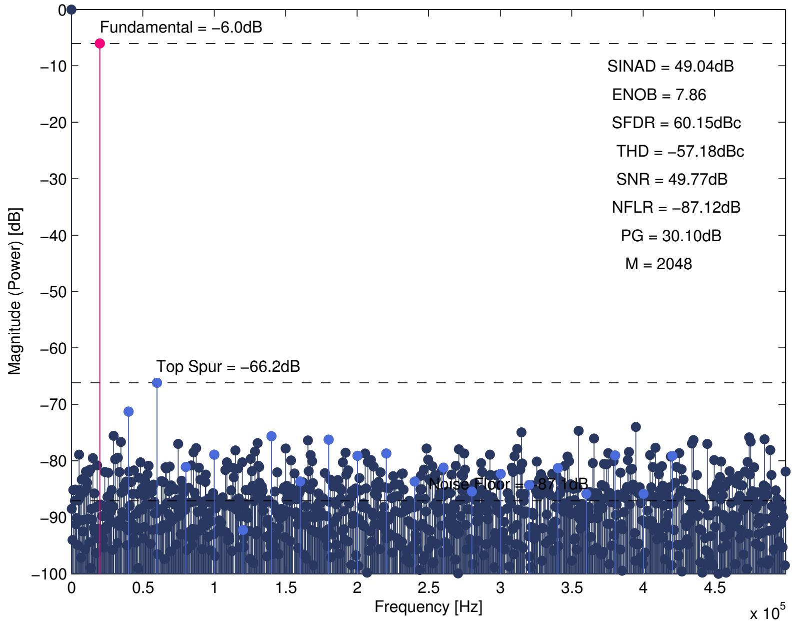


Figure B.11: ADC performance simulated at 1.0 MSps (coherent sampling, 2048 point FFT).

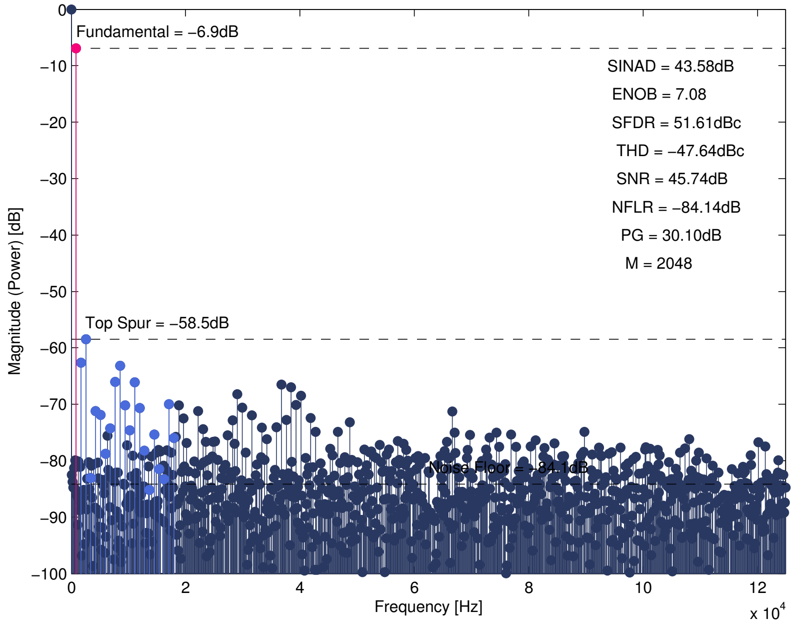


Figure B.12: ADC performance measured at 250kSps (coherent sampling, 2048 point FFT).

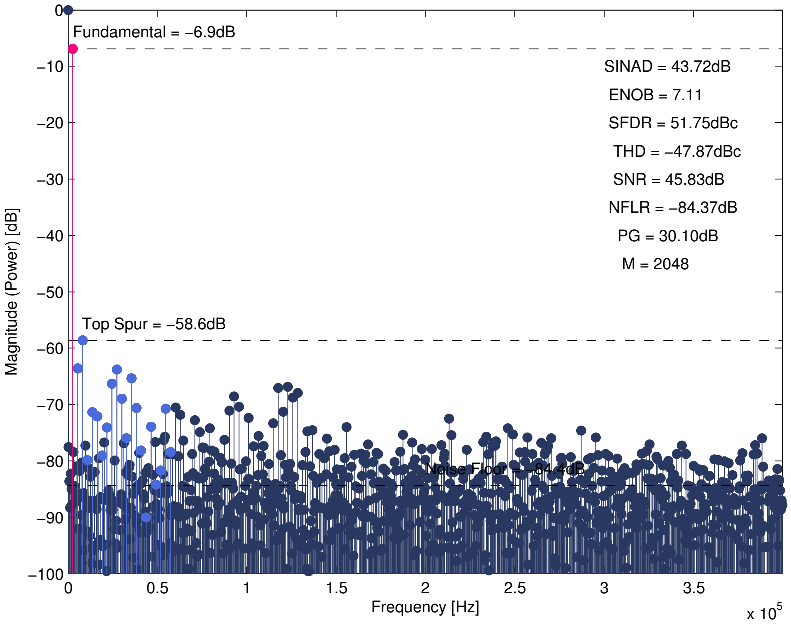


Figure B.13: ADC performance measured at 800kSps (coherent sampling, 2048 point FFT).

B.4 Potential Improvements

With an N-bit ADC it is impossible to achieve an ENOB of N in the real world. The measured ADC ENOB is over 7 bits, however. Thus there may not be a need to improve the design for our application. However, I see two main ways to improve the design. One is to improve the comparator, whose offset is limiting the performance. The layout could have better matching. The other is to relax the design requirement of rail-to-rail operation, which isn't really necessary, as the main signals that will be sampled are not rail-to-rail signals. Relaxing this requirement would make the circuit design simpler (especially the comparator).

Appendix C

Radiometer Measurement Procedure

The purpose of the irradiance measurements is twofold: 1) to find the true optical power of the LED, and its relation to current draw, and 2) to find the true optical power hitting the pixels so that the LOD of the imager can be found.

Fig. C.1 shows the parameters used in the calculations. The left side of the figure shows the LED (on LEDPCB) to imager distance, r_p, which is fixed. The right side of the figure shows the LED (on LEDPCB) to radiometer detector distance, r_d, which is variable. Note that in these measurements the geometry must be such that the LED can be treated as a point source, so the inverse-square law can be used in calculations (i.e. the diameter of the LED is < 1/10th the distance between LED and detector).

The basic idea is to measure the on-axis power emitted by the LED by placing a photodetector with a given active area at some distance away from the LED, and recording the average power hitting the photodetector. Note that the light distribution across the detector may vary if

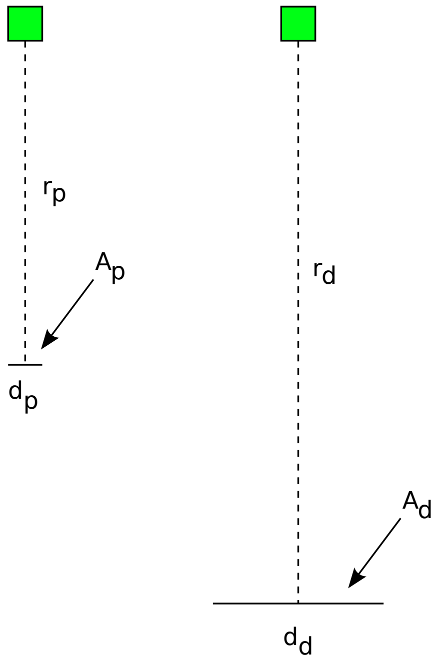


Figure C.1: Diagram of irradiance measurement geometry.

the illumination is not uniform, and we are measuring the *average* power. However, when the diameter of the photodetector is significantly smaller than the distance between the LED and photodetector (about 5x), then the irradiance will be independent of the photodetector size, and the reading can be considered accurate [28]. Then the inverse-square law can be used to find the irradiance at a different distance from the LED, as follows:

$$E_p = \frac{E_d r_d^2}{r_p^2} \tag{C.1}$$

The geometrical parameters for this work are shown in Table C.1 (distances in mm). All parameters are such that the inverse-square law is satisfied.

To perform the LED irradiance measurements, the ILT1700 Research Radiometer from In-

Parameter	LEDPCB
r_p	24.6
r_d	63.5
d_p	0.1
d_d	10.0

Table C.1: Irradiance measurement system parameters (all values in mm).

ternational Light Technologies is used, along with their SED033 detector coupled with the F flat band filter [92]. The ILT1700 can provide measurement readings in irradiance units of W/m^2.

References

[1] J. Aceros, Ming Yin, D.A. Borton, W.R. Patterson, and A.V. Nurmikko. A 32-channel fully implantable wireless neurosensor for simultaneous recording from two cortical regions. In *Proc. of the Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2300–2306, Sep. 2011.

[2] Walther Akemann, Hiroki Mutoh, Amelie Perron, Jean Rossier, and Thomas Knopfel. Imaging brain electric signals with genetically targeted voltage-sensitive fluorescent proteins. *Nature Methods*, 7(8):643–649, Aug. 2010.

[3] Jasper Akerboom, Tsai-Wen Chen, Trevor J. Wardill, Lin Tian, Jonathan S. Marvin, Sevinc Mutlu, Nicole Carreras Calderon, Federico Esposti, Bart G. Borghuis, Xiaonan Richard Sun, Andrew Gordus, Michael B. Orger, Ruben Portugues, Florian Engert, John J. Macklin, Alessandro Filosa, Aman Aggarwal, Rex A. Kerr, Ryousuke Takagi, Sebastian Kracun, Eiji Shigetomi, Baljit S. Khakh, Herwig Baier, Leon Lagnado, Samuel S.-H. Wang, Cornelia I. Bargmann, Bruce E. Kimmel, Vivek Jayaraman, Karel Svoboda, Douglas S. Kim, Eric R. Schreiter, and Loren L. Looger. Optimization of a GCaMP calcium indicator for neural activity imaging. *The Journal of Neuroscience*, 32(40):13819–13840, 2012.

[4] Altium. altium.com/en/products/altium-designer.

[5] B. Aouizerate, Emmanuel Cuny, Corinne Martin-Guehl, Dominique Guehl, Helene Amieva, Abdelhamid Benazzouz, Colette Fabrigoule, Michele Allard, Alain Rougier, Bernard Bioulac, Jean Tignol, and Pierre Burbaud. Deep brain stimulation of the ventral caudate nucleus in the treatment of obsessive-compulsive disorder and major depression. *Journal of Neurosurgery*, 101(4):682–686, Oct. 2004.

[6] B. J. Baker, H. Mutoh, Dimitrov, W. Akemann, A. Perron, Y. Iwamoto, L. Jin, L. B. Cohen, E. Y. Isacoff, V. A. Pieribone, T. Hughes, and T. Knopfel. Genetically encoded fluorescent sensors of membrane potential. *Brain Cell Biology*, 36:53–67, 2008.

[7] Mark W Barnett and Philip M Larkman. The action potential. *Practical Neurology*, 7(3):192–197, 2007.

[8] Rizwan Bashirullah, John G. Harris, Justin C. Sanchez, Toshikazu Nishida, and Jose C. Principe. Florida wireless implantable recording electrodes (FWIRE) for brain machine interfaces. *Proc. of the IEEE Int. Symp. on Circuits and Systems (ISCAS)*, 2007.

[9] M. Beiderman, T. Tam, A. Fish, G.A. Jullien, and O. Yadid-Pecht. A low-light CMOS contact imager with an emission filter for biosensing applications. *IEEE Transactions on Biomedical Circuits and Systems*, 2(3):193–203, 2008.

[10] A. Bonfanti, M. Ceravolo, G. Zambra, R. Gusmeroli, A.S. Spinelli, A.L. Lacaita, G.N. Angotzi, G. Baranauskas, and L. Fadiga. A multi-channel low-power system-on-chip for single-unit recording and narrowband wireless transmission of neural signal. In *Proc. of the Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1555–1560, Sep. 2010.

[11] David A Borton, Ming Yin, Juan Aceros, and Arto Nurmikko. An implantable wireless neural interface for recording cortical circuit dynamics in moving primates. *Journal of Neural Engineering*, 10(2):026010, 2013.

[12] Moo Sung Chae, Wentai Liu, and M. Sivaprakasam. Design optimization for integrated neural recording systems. *IEEE Journal of Solid-State Circuits*, 43(9):1931–1939, Sep. 2008.

[13] J. K. Chapin, K. A. Moxon, R. S. Markowitz, and M. A. L. Nicolelis. Real-time control of a robot using simultaneously recorded neurons in the motor cortex. *Nature Neuroscience*, 2:664–670, 1999.

[14] Cameron T. Charles. Wireless data links for biomedical implants: Current research and future directions. 2007.

[15] Sung-Min Chin, Chih-Cheng Hsieh, Chin-Fong Chiu, and Hann-Huei Tsai. A new rail-to-rail comparator with adaptive power control for low power SAR ADCs in biomedical application. In *Proc. of the IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1575–1578, 2010.

[16] Electrical Cleanroom and Brigham Young University Computer Engineering Department. Optical absorption coefficient calculator. www.cleanroom.byu.edu/OpticalCalc.phtml.

[17] Stuart F. Cogan. Electrode and electrode-tissue interface properties relevant to the design of implanted microelectronic devices. In *ISSCC Digest of Technical Papers, F8: Integrated Neural Interfaces*, Feb. 2009.

[18] Chroma Technology Corp. Chroma spectra viewer. www.chroma.com/spectra-viewer?fluorochromes=633.

[19] Lambda Research Corporation. Tracepro software. http://www.lambdares.com/.

[20] A. Coskun. *Lensfree Fluorescent Computational Microscopy on a Chip*. PhD thesis, University of California Los Angeles, 2013.

[21] Cospheric. Fluorescent orange polyethylene microspheres. http://www.cospheric.com/UVPMSBO_fluorescent_orange_spheres_density100.htm.

[22] Cree. www.cree.com.

[23] Marc Dandin, Pamela Abshire, and Elisabeth Smela. Optical filtering technologies for integrated fluorescence sensors. *Lab Chip*, 7(8):955–977, 2007.

[24] Elena Dreosti and Leon Lagnado. Optical reporters of synaptic activity in neural circuits. *Experimental Physiology*, 96(1):4–12, 2011.

[25] H. Eltoukhy, K. Salama, and A. El Gamal. A 0.18-μm CMOS bioluminescence detection lab-on-chip. *IEEE Journal of Solid-State Circuits*, 41(3):651–662, 2006.

[26] Photon Engineering. Fred software. http://www.photonengr.com/.

[27] E. V. Evarts. Relation of pyramidal tract activity to force exerted during voluntary movement. *Journal of Neurophysiology*, 31:14–27.

[28] Excelitas. Application note #5, ired axial power out measurement. http://www.excelitas.com/downloads/app_ired.pdf.

[29] S. Farshchi, D. Markovic, S. Pamarti, B. Razavi, and J.W. Judy. Towards neuromote: A single-chip, 100-channel, neural-signal acquisition, processing, and telemetry device. *29th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBS)*, pages 437–440, Aug. 2007.

[30] Lee Filters. Colour filters, 135 deep golden amber. http://www.leefilters.com/lighting/colour-details.html#135&filter=cf.

[31] E. Fluhler, V. G. Burnham, and L. M. Loew. Spectra, membrane binding, and potentiometric responses of new charge shift probes. *Biochemistry*, 24(21):5749–5755, 1985.

[32] S. C. Gebhart, W. C. Lin, and A. Mahadevan-Jansen. In vitro determination of normal and neoplastic human brain tissue optical properties using inverse adding-doubling. *Physics in Medicine and Biology*, 51(8):2011, 2006.

[33] Kunal K Ghosh, Laurie D Burns, Eric D Cocker, Axel Nimmerjahn, Yaniv Ziv, Abbas El Gamal, and Mark J Schnitzer. Miniaturized integration of a fluorescence microscope. *Nature Methods*, 8(10):871–878, October 2011.

[34] Z. Gorocs and A. Ozcan. On-chip biomedical imaging. *IEEE Reviews in Biomedical Engineering*, 6:29–46, 2013.

[35] Benoit Gosselin. Recent advances in neural recording microsystems. *Sensors*, 11(5):4572–4597, 2011.

[36] Nir Grossman, Vincent Poher, Matthew S Grubb, Gordon T Kennedy, Konstantin Nikolic, Brian McGovern, Rolando Berlinguer Palmini, Zheng Gong, Emmanuel M Drakakis, Mark A A Neil, Martin Dawson, Juan Burrone, and Patrick Degenaar. Multi-site optical excitation using ChR2 and micro-LED array. *Journal of Neural Engineering*, 7(1):016004, 2010.

[37] David Halliday, Robert Resnick, and Jearl Walker. *Fundamentals of Physics*. John Wiley & Sons, Inc., 6th edition, 2001.

[38] R. Harrison. Wireless neural recording systems: Design trade-offs at the circuit and system levels. In *Proc. of ISSCC 2009, F8: Integrated Neural Interfaces*, Feb. 2009.

[39] R. R. Harrison, P. T. Watkins, R. J. Kier, R. O. Lovejoy, D. J. Black, B. Greger, and F. Solzbacher. A low-power integrated circuit for a wireless 100-electrode neural recording system. *IEEE Journal of Solid-State Circuits*, 42(1):123–133, Jan. 2007.

[40] R.R. Harrison. The design of integrated circuits to observe brain activity. *Proceedings of the IEEE*, 96(7):1203–1216, July 2008.

[41] R.R. Harrison, R.J. Kier, C.A. Chestek, V. Gilja, P. Nuyujukian, S. Ryu, B. Greger, F. Solzbacher, and K.V. Shenoy. Wireless neural recording with single low-power integrated circuit. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 17(4):322–329, Aug. 2009.

[42] L. F. Heffer and J. B. Fallon. A novel stimulus artifact removal technique for high-rate electrical stimulation. *Journal of Neuroscience Methods*, 170(2):277–84, May 2008.

[43] Shaul Hestrin and William E. Armstrong. Morphology and physiology of cortical neurons in layer I. *The Journal of Neuroscience*, 16(17):5290–5300, 1996.

[44] Elizabeth M. C. Hillman. Optical brain imaging in vivo: techniques and applications from animal to man. *Journal of Biomedical Optics*, 12(5), Nov. 2007.

[45] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. van der Smagt, and J. P. Donoghue. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398):372–375, May 2012.

[46] L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, 442:164–171, 2006.

[47] Veronica S. Hollis. *Non-Invasive Monitoring of Brain Tissue Temperature by Near-Infrared Spectroscopy*. PhD thesis, University College London, 2002.

[48] M. L. Homer, A. V. Nurmikko, J. P. Donoghue, and L. R. Hochberg. Sensors and decoding for intracortical brain computer interfaces. *Annual Reviews in Biomedical Engineering*, 15:383–405, 2013.

[49] Wei-Hao Hsiao, Yi-Ting He, M.P.-H. Lin, Rong-Guey Chang, and Shuenn-Yuh Lee. Automatic common-centroid layout generation for binary-weighted capacitors in charge-scaling DAC. In *Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), 2012 International Conference on*, pages 173–176, 2012.

[50] Hugo et. al. *Bioinstrumentation*. John Wiley & Sons, Inc., 2004.

[51] Walter G. Jung. *Op Amp Applications*. Analog Devices Inc., 2002.

[52] Opal Kelly. www.opalkelly.com.

[53] Tae-il Kim, Jordan G. McCall, Yei Hwan Jung, Xian Huang, Edward R. Siuda, Yuhang Li, Jizhou Song, Young Min Song, Hsuan An Pao, Rak-Hwan Kim, Chaofeng Lu, Sung Dan Lee, Il-Sun Song, GunChul Shin, Ream Al-Hasani, Stanley Kim, Meng Peun Tan, Yonggang Huang, Fiorenzo G. Omenetto, John A. Rogers, and Michael R. Bruchas. Injectable, cellular-scale optoelectronics with applications for wireless optogenetics. *Science*, 340(6129):211–216, 2013.

[54] Joel M. Kralj, Adam D. Douglass, Daniel R. Hochbaum, Dougal Maclaurin, and Adam E. Cohen. Optical recording of action potentials in mammalian neurons using a microbial rhodopsin. *Nature Methods*, 9(1):90–95, 2012.

[55] M. A. Lebedev and M. A. L. Nicolelis. Brain-machine interfaces: past, present, future. *Trends Neurosci.*, 29(9):536–546, 2006.

[56] J. J. Mancuso, J. Kim, S. Lee, S. Tsuda, N. B. Chow, and G. J. Augustine. Optogenetic probing of functional brain circuitry. *Exp. Physiol*, 96(1):26–33, Jan. 2011.

[57] CMC Microsystems. www.cmc.ca.

[58] Elizabeth Munro. Implantable biosensors for neural imaging: a study of optical modeling and light sources. Master's thesis, University of Toronto, 2009.

[59] K. Murari, R. Etienne-Cummings, Nitish Thakor, and G. Cauwenberghs. Which photodiode to use: a comparison of CMOS-compatible structures. *Sensors Journal, IEEE*, 9(7):752–760, 2009.

[60] K. Murari, R. Etienne-Cummings, N.V. Thakor, and G. Cauwenberghs. A CMOS in-pixel CTIA high-sensitivity fluorescence imager. *IEEE Trans. on Biomedical Circuits and Systems*, 5(5):449–458, Oct. 2011.

[61] Hiroki Mutoh, Amelie Perron, Walther Akemann, Yuka Iwamoto, and Thomas Knopfel. Optogenetic monitoring of membrane potentials. *Experimental Physiology*, 96(1):13–18, 2011.

[62] D. C. Ng, T. Nakagawa, T. Mizuno, T. Tokuda, M. Nunoshita, H. Tamura, Y. Ishikawa, S. Shiosaka, and J. Ohta. Integrated in vivo neural imaging and interface CMOS devices: design, packaging, and implementation. *IEEE Sensors Journal*, 8(1):121–130, Jan. 2008.

[63] David C. Ng, Takashi Tokuda, Akio Yamamoto, Masamichi Matsuo, Masahiro Nunoshita, Hideki Tamura, Yasuyuki Ishikawa, Sadao Shiosaka, and Jun Ohta. On-chip biofluorescence imaging inside a brain tissue phantom using a CMOS image sensor for in vivo brain imaging verification. *Sensors and Actuators B: Chemical*, 119(1):262 – 274, 2006.

[64] K. D. Niswender, S. M. Blackman, L. Rohde, M. A. Magnuson, and D. W. Piston. Quantitative imaging of green fluorescent protein in cultured cells: Comparison of microscopic techniques, use in fusion proteins and detection limits. *Journal of Microscopy*, 180(2):109–116, Nov. 1995.

[65] Craig T. Nordhausen, Edwin M. Maynard, and Richard A. Normann. Single unit recording capabilities of a 100 microelectrode array. *Brain Research*, 726:129–140, 1996.

[66] Nusil. Med 6215 (pdms). http://nusil.com/products/findproducts.aspx?s=MED-6215.

[67] J. Ohta, A. Higuchi, A. Tagawa, K. Sasagawa, T. Tokuda, Y. Hatanaka, H. Tamura, and S. Shiosaka. An implantable CMOS image sensor for monitoring deep brain activities of a freely moving mouse. In *Proc. of the IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 269–272, 2008.

[68] J. Ohta, C. Kitsumoto, T. Noda, K. Sasagawa, T. Tokuda, M. Motoyama, Y. Ohta, T. Kobayashi, Y. Ishikawa, and S. Shiosaka. A micro imaging device for measuring neural actvities in the mouse deep brain with minimal invasiveness. In *Biomedical Circuits and Systems Conference (BioCAS), 2012 IEEE*, pages 244–247, 2012.

[69] J. Ohta, T. Noda, K. Sasagawa, and T. Tokuda. Implantable micro CMOS imaging devices for biomedical applications. In *International Conference on Optical MEMS and Nanophotonics (OMN)*, pages 5–6, 2013.

[70] Optotech. www.opto.com.tw.

[71] M. Ortmanns. Safety issues and circuit implementation for retinal stimulators. In *Proc. of ISSCC 2009, F8: Integrated Neural Interfaces*, Feb. 2009.

[72] A. Osman, Joon Hyuk Park, D. Dickensheets, J. Platisa, E. Culurciello, and V.A. Pieribone. Design constraints for mobile, high-speed fluorescence brain imaging in awake animals. *IEEE Transactions on Biomedical Circuits and Systems*, 6(5):446–453, 2012.

[73] Quik Pak. www.icproto.com.

[74] Kunal Paralikar, Peng Cong, Ofer Yizhar, Lief Ericsson Fenno, Wesley Santa, Chris Nielson, David Dinsmoor, Bob Hocken, Gordon O. Munns, Jon Giftakis, Karl Deisseroth, and Timothy Denison. An Implantable Optical Stimulation Delivery System for Actuating an Excitable Biosubstrate. *IEEE Journal of Solid-State Circuits (JSSC)*, 46(1), Jan. 2011.

[75] Joon Hyuk Park, E. Culurciello, Dongsoo Kim, J. V. Verhagen, S. H. Gautam, and V. Pieribone. Voltage sensitive dye imaging system for awake and freely moving animals. In *Proc. of the IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 89–92, 2008.

[76] Joon Hyuk Park, V. Pieribone, Dongsoo Kim, J. V. Verhagen, C. von Hehn, and E. Culurciello. High-speed fluorescence imaging system for freely moving animals. In *Proc. of*

the IEEE International Symposium on Circuits and Systems (ISCAS), pages 2429–2432, 2009.

[77] Joon Hyuk Park, Jelena Platisa, Vincent Pieribone, and E. Culurciello. A second-generation imaging system for freely moving animals. In *IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 105–108, 2012.

[78] Joon Hyuk Park, Jelena Platisa, Justus V. Verhagen, Shree H. Gautam, Ahmad Osman, Dongsoo Kim, Vincent A. Pieribone, and Eugenio Culurciello. Head-mountable high speed camera for optical neural recording. *Journal of Neuroscience Methods*, 201(2):290 – 295, 2011.

[79] Sunmee Park, David A. Borton, Mingyu Kang, Arto V. Nurmikko, and Yoon-Kyu Song. An implantable neural sensing microsystem with fiber-optic data transmission and power delivery. *Sensors*, 13(5):6014–6031, 2013.

[80] Parag G. Patil and Dennis A. Turner. The development of brain-machine interface neuroprosthetic devices. *Neurotherapeutics*, 5(1):137–146, Jan. 2008.

[81] S. Reckziegel, D. Kreye, T. Puegner, U. Vogel, M. Scholles, C. Grillberger, and K. Fehse. Flow sensor based on monolithic integration of organic light-emitting diodes (OLEDs) and CMOS circuits. volume 7219, page 72190O. SPIE, 2009.

[82] B. Richter, U. Vogel, P. Wartenberg, K. Fehse, and R. Herold. OLED-on-CMOS based bidirectional microdisplay for near-to-eye and sensor applications. In *Semiconductor Conference Dresden (SCD), 2011*, pages 1–3, Sep. 2011.

[83] A G Rouse, S R Stanslaski, P Cong, R M Jensen, P Afshar, D Ullestad, R Gupta, G F Molnar, D W Moran, and T J Denison. A chronic generalized bi-directional brain-machine interface. *Journal of Neural Engineering*, 8(3):036018, 2011.

[84] P. Ruther, T. Holzhammer, S. Herwik, P.D. Rich, J.W. Dalley, O. Paul, and T. Holtzman. Compact wireless neural recording system for small animals using silicon-based probe arrays. In *Proc. of the Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2284–2287, Sep. 2011.

[85] Massimo Scanziani and Michael Hausser. Electrophysiology in the age of light. *Nature*, 461:930–939, Oct. 2009.

[86] T. Seese, H. Harasaki, G. Saidel, and C. Davies. Characterization of tissue morphology, angiogenesis, and temperature in the adaptive response of muscle tissue in chronic heating. *Lab. Invest.*, 78:1553–1562, 1998.

[87] J. Senarathna, K. Murari, R. Etienne-Cummings, and N.V. Thakor. Design of a novel head-mountable microscope system for laser speckle imaging. In *Proc. of the IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 117 –120, Nov. 2011.

[88] R.R. Singh, D. Ho, A. Nilchi, P.G. Gulak, P. Yau, and R. Genov. A CMOS/thin-film fluorescence contact imaging microsystem for DNA analysis. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 57(5):1029–1038, 2010.

[89] R.R. Singh, Lian Leng, A. Guenther, and R. Genov. A CMOS-microfluidic chemiluminescence contact imaging microsystem. *IEEE Journal of Solid-State Circuits*, 47(11):2822–2833, 2012.

[90] Lucas Sjulson and Gero Miesenboeck. Optical recording of action potentials and other discrete physiological events: A perspective from signal detection theory. *Physiology*, 22(1):47–55, 2007.

[91] Alma E. F. Taylor. *Illumination Fundamentals*. Rensselaer Polytechnic Institute, 2000. Lighting Research Center.

[92] International Light Technologies. www.intl-lighttech.com.

[93] Life Technologies. https://www.lifetechnologies.com/ca/en/home/references/molecular-probes-the-handbook/probes-for-membrane-potential/fast-response-probes.html#head1.

[94] P. R. Thompson and Thomas C. Larason. Method of measuring shunt resistance of photodiodes. In *Measurement Science Conference*, Jan. 2001.

[95] L. Tian, S. A. Hires, T. Mao, D. Huber, M. E. Chiappe, S. H. Chalasani, L. Petreanu, J. Akerboom, S. A. McKinney, E. R. Schreiter, C. I. Bargmann, V. Jayaraman, K. Svoboda, and L. L. Looger. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods*, 6(12):875–881, Dec. 2009.

[96] Jing Wang, Fabien Wagner, David A Borton, Jiayi Zhang, Ilker Ozden, Rebecca D Burwell, Arto V Nurmikko, Rick van Wagenen, Ilka Diester, and Karl Deisseroth. Integrated device for combined optical neuromodulation and electrical recording for chronic in vivo applications. *Journal of Neural Engineering*, 9(1):016001, 2012.

[97] Christian T. Wentz, Jacob G. Bernstein, Patrick Monahan, Alexander Guerra, Alex Rodriguez, and Edward S. Boyden. A wirelessly powered and controlled device for optical neural control of freely-behaving animals. *Journal of Neural Engineering*, 8:10, 2011.

[98] J. Wessberg, C. R. Stambaugh, J. D. Kralik, P. D. Beck, M. Laubach, J. K. Chapin, J. Kim, S. J. Biggs, M. A. Srinivasan, and M. A. L. Nicolelis. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, 408:361–365, 2000.

[99] SF Witelson, II Glezer, and DL Kigar. Women have greater density of neurons in posterior temporal cortex. *The Journal of Neuroscience*, 15(5):3418–3428, 1995.

[100] Alexey Yakushenko, Zheng Gong, Vanessa Maybeck, Boris Hofmann, Erdan Gu, Martin Dawson, Andreas Offenhauml, and Bernhard Wolfrum. On-chip optical stimulation and electrical recording from cells. *Journal of Biomedical Optics*, 18(11):111402, 2013.

[101] O. Yizhar, L. E. Fenno, T. J. Davidson, M. Mogri, and K. Deisseroth. Optogenetics in neural systems. *Neuron*, 71(1):9–34, Jul. 2011.

[102] Hatim A. Zariwala, Bart G. Borghuis, Tycho M. Hoogland, Linda Madisen, Lin Tian, Chris I. De Zeeuw, Hongkui Zeng, Loren L. Looger, Karel Svoboda, and Tsai-Wen Chen. A cre-dependent GCaMP3 reporter mouse for neuronal imaging in vivo. *The Journal of Neuroscience*, 32(9):3131–3141, 2012.

[103] Michal Zochowski, Matt Wachowiak, Chun X. Falk, Lawrence B. Cohen, Ying-Wan Lam, Srdjan Antic, and Dejan Zecevic. Concepts in imaging and microscopy imaging membrane potential with voltage-sensitive dyes. *Biological Bulletin*, 198(1):1–21, Feb. 2000.