Optical Voltage Sensing: Micro- to Kilo-Scale

Jordan Edmunds
Michel Maharbiz, Ed.

Electrical Engineering and Computer Sciences
University of California, Berkeley

Technical Report No. UCB/EECS-2023-55
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2023/EECS-2023-55.html

May 1, 2023

Optical Voltage Sensing: Micro- to Kilo-Scale

by

Jordan Edmunds

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michel M. Maharbiz, Chair
Professor Dorian Liepmann
Professor Clark T.-C. Nguyen

Summer 2022

The dissertation of Jordan Edmunds, titled Optical Voltage Sensing: Micro- to Kilo-Scale, is approved:

Chair _____ Date _____

_____ Date _____

_____ Date _____

University of California, Berkeley

Optical Voltage Sensing: Micro- to Kilo-Scale

To my dad, ever understanding, supportive, and kind.

Contents

List of Figures

List of Tables

Acknowledgments

This work would not have been possible without many key people and institutions. First, my undergraduate mentors: An Do, Zoran Nenadic, and Henry Lee. I would never have become an Electrical Engineer had it not been for James Bullock, my introductory undergraduate electromagnetics professor, who showed me the incredible beauty of the subject. I would not have been able to carry out this work if not for my undergraduate institution, the University of California, Irvine, for giving their students freedom to explore and the resources to succeed. This would also not have been possible without funding from the Hertz Foundation, who funded the majority of my Ph.D. I will be forever grateful for the opportunity afforded to me by Hertz and by Lee Swanger - the donor who endowed the fellowship i received. I would not have retained my sanity or finished as fast as I did without the ever-positive and wonderfully competent staff of the Marvell Nanofabrication Facility. I would also like to thank my advisor, Michel, who gave me the freedom to explore, fail, fail, fail, fail, fail, and graduate. Michel - you were the advisor I always needed (if not always the one I wanted).

I would also like to thank my second advisor, Soner Sonmezoglu, for his friendship, honesty, hard work, and mentorship, all of which was essential to my growth and graduation in my final year.

Finally, I would not have been able to survive these last four years had it not been for my supportive friends, family, and partner. Your faith and support has always been unwavering, and was instrumental every time I lost my own.

Abstract

Optical Voltage Sensing: Micro- to Kilo-Scale

by

Jordan Edmunds

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Michel M. Maharbiz, Chair

Detection of electrical neural activity has been possible for almost a century. Despite the progress in the last century, neural recording technologies remain fundamentally constrained in their scale by the power they are able to dissipate. As channel counts have increased, and electrodes have grown flexible and microscale, connectorization to neural recording circuits has become ever more challenging. All-optical neural recording technologies hold the promise to eliminate these barriers. However, until this point, it has been impossible to quantitatively compare existing electrical recording technology with emerging optical solutions. I introduce a novel figure of merit E_Q, which overcomes this challenge, and allows direct and quantitative comparison between arbitrary technologies. I propose a novel neural recording architecture which can achieve a lower E_Q than existing electronic methods, consisting of free-floating passive optical voltage sensors. These sensors have no active circuitry and require no stored power, and can be interrogated wirelessly at will. I describe progress towards the realization of these sensors, including their design, fabrication, and characterization, as well as their use for digital communication through highly scattering tissue. Additionally, I show that the same sensors and the same framework can be applied to solve some outstanding problems in these sensors use in microgrid applications for kV-range sensing.

Chapter 1

Introduction

The work in this dissertation started out as an attempt to design a wireless, implantable neural voltage sensor. Along the way (as so often happens), I discovered another potential application for our design in power electronics - voltage sensing for microgrids. This introduction parallels that journey, starting with the idea for a neural recording sensor, and taking a detour to discuss its applications in grid voltage sensing. We then discuss properties of the ideal sensor which are common to both, and properties of the ideal electronic and optical sensor.

1.1 The Neural Sensing Problem

Every animal on planet earth uses the same underlying technology to perceive, move, and act in the world - a nervous system. This consists of a network of individual "neurons", each connected to many other neurons. The operation of this nervous system is electrical - voltages and currents are how neurons communicate with each other. The electrical nature of this system was first discovered, somewhat accidentally, in 1771 by Luigi Galvani experimenting on frog legs, and research has proceeded to this day trying to understand exactly how it operates.

Traditionally, this has been accomplished with mixed-signal electronics: a voltage is measured, either from the inside of the cell to the outside, or comparing the voltage nearby and far away from a cell of interest. This signal, typically in the μV to mV range, is then amplified and digitized. While this approach has worked well for small-scale neural recording up to hundreds to a few thousand channels, above roughly 1,000 channels it becomes incredibly difficult to connectorize [124]. New approaches have improved channel count by an order of magnitude [106], but at the cost of size and power.

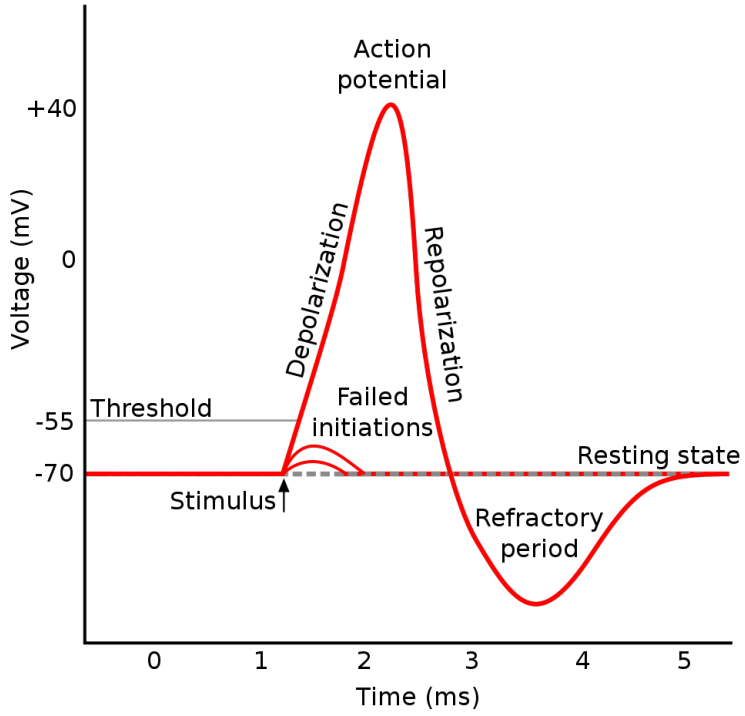

Figure 1.1: Neuron action potential, measured with the positive lead inside the cell and the negative lead on the outside. Credit: Wikimedia Commons

Neural Signal Basics

The fundamental event of interest in neural recording is thought to be the action potential [8] . When a neuron is at rest, it continually pumps sodium ions out of the cell, making the interior of the cell negatively charged, at a potential of roughly $-70mV$. When the "inputs" to a given neuron (from other neurons) are sufficiently strong, a neuron fires an action potential. A large number of ion channels on the neuron's surface open, and the inrush of sodium ions causes the potential inside the cell to rise. This action potential propagates as a wave across the surface of a neuron, and can propagate vast distances down neural projections (up to 1 meter!). Currently, the action potential is believed to be the fundamental computational action of the brain (although research is ongoing on the effects and importance of other events [11, 73, 63, 99]), and so recording these action potentials as they occur is of both fundamental scientific and clinical interest.

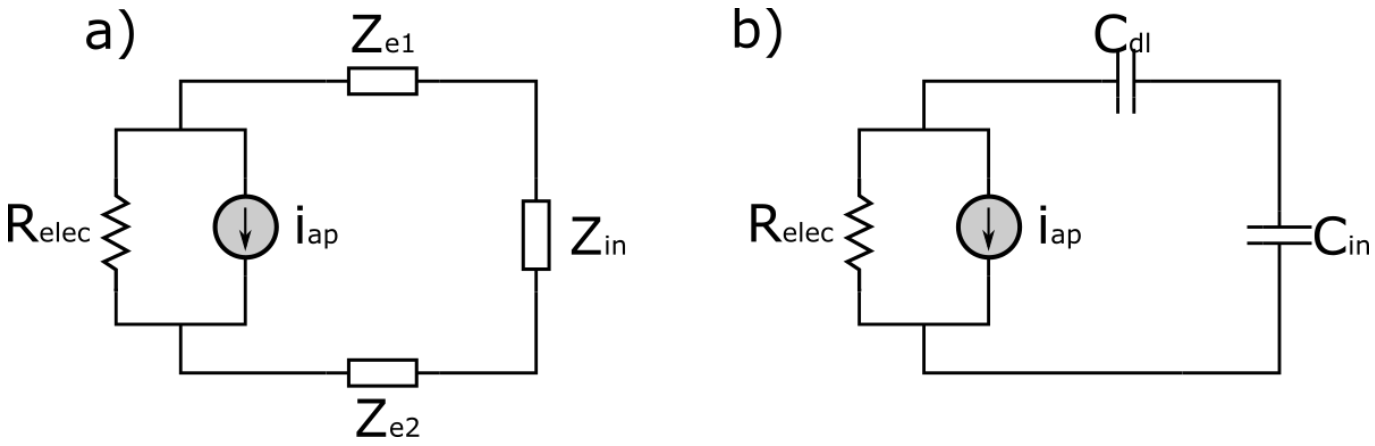

Figure 1.2: (a) Full neural recording circuit model (b) Simplified typical neural recording circuit model.

Neural Recording Circuit

Any device which wants to directly measure electric potential caused by a neuron needs to first probe that potential in space. This is typically done with metallic conductors - electrodes - placed in contact with the extracellular fluid, or in contact with the interior of the cell. The full equivalent circuit, along with its typical simplification, is shown in Fig. 1.2. The current from the action potential i_{ap} flows through the electrolyte resistance R_{elec}. Typically, R_{elec} is on the order of a few ohms. Some of that current is shunted through the electrode impedance Z_{e1} and Z_{e2}, and flows into the input impedance of the voltage sensor, typically an electronic amplifier. Typically, Z_{e2} is made very small by making the "reference electrode" very large, so it is often neglected. Z_{e1} is typically capacitive due to the double-layer capacitance of the electrode C_{dl}, although a more complete model is given in [10]. Finally, the input impedance is typically also capacitive, with capacitance C_{in}.

1.2 Techniques for Neural Recording

Many techniques exist to perform neural recording - and each has its advantages and drawbacks. Extracellular neural recording is currently the most common method, but still presents substantial challenges in terms of invasiveness, power consumption, and scalability. This dissertation started with this question: can we develop a neural recording technology that combines the benefits of optical techniques (scalability, wirelessness, favorable cellular response) with the advantages of traditional extracellular neural recording arrays (ability to record at depth, long-term use)? What would such a technology look like, and what would be its limitations? This dissertation aims to answer those questions.

Patch Clamp - The Highest-SNR Recording Technique

After initial work with extracellular electrodes, scientists also developed electrodes capable of recording from the interior of the cell [142]. By isolating a small part of the cell surface and opening a hole in the cell membrane, the potential inside the cell can be directly measured. This allows the full difference in potential between the inside and outside of the cell to be measured, on the order of $100mV$ as discussed previously. This yields much higher SNR compared to extracellular neural recording. While of fundamental scientific importance, this technique is practically challenging to scale beyond a few cells at a time - physically making a seal with the cell wall sufficient to electrically seal it is by itself difficult, and making holes in the cell membrane is a process that can result in the death of the cell.

Extracellular Neural Recording

The first neural recordings were made using extracellular electrodes - with one electrode placed nearby a firing neuron and one far away [142]. The current that flows into a cell during an action potential causes a resistive drop in the fluid surrounding the cell, and hence a change in the potential of the area surrounding the cell [47], which falls off with distance. If one electrode is extremely close to the cell, potentials as high as hundreds of microvolts can be measured relative to an electrode far away. This method of neural recording is much easier to perform, but requires that the neural signal be amplified substantially (gains of 40 to 80dB are typical) [150] prior to being digitized and recorded. It has the advantage of being reasonably scalable, as arrays of hundreds or thousands of electrodes can be monolithically fabricated [92], or simply routed out of the brain as individual wires, as is the case with microdrives [70] and bundled electrode arrays [106]. Modern microfabricated electrode arrays have managed thousands to tens of thousands of separate recording channels [133, 106], and development in these is ongoing. The major disadvantage of these large-scale arrays is that they typically require wires to be routed outside the skull, as the large volume and power dissipation of the associated electronics cannot be accomodated by the sensitive area inside the skull. Any power dissipated by electronics inside the body of a human or animal causes tissue to heat up, and a temperature increase as small as 1-2°C. can cause tissue death [155]. Those systems that do incorporate on-board wireless transmission can consume huge amounts of power transmitting the data [102], roughly 100x the power that was used to aquire it in the first place.

Microfabricated Neural Recording Arrays

Of particular note for the purpose of extracellular recording are microfabricated neural recording arrays. In an attempt to scale up neural recording to record from an ever-larger number of neurons, many groups have used microfabrication techniques pioneered by the

semiconductor industry to fabricate arrays of electrodes, an excellent historical review of which can be found here [51]. More recently, greater attention has been focused on developing flexible and microscopic [91] electrode arrays, which cause far less neural damage on implantation.

Often, but not always, these arrays have on-board integrated microelectronics which allow for the amplification, digitization, and transmission of recorded data. However, each of these functions requires power. For fully implantable systems, this power must be dissipated inside the brain itself, which causes the local temperature to rise, greatly constraining the power available to perform these functions. If the system is not fully implantable, as in [106], much more power is available to perform these functions. While useful for fundamental science, these have challenges when being translated into long-term clinically useful devices due to the ongoing infection risk and other well-documented failure modes [26].

If on-board microelectronics are not included (and often, even if they are), connectorization of the electrode arrays to the recording circuitry has become a major challenge. Physically connecting each individual electrode to its associated amplifier is an extreme assembly challenge and outstanding problem which has not been fully solved.

These two major bottlenecks have prevented further scaling of neural recording arrays, limiting the information that can be extracted from the brain.

Optical Methods of Neural Sensing

More recently, optical methods have been developed which allow neural recording to be done without any contact needing to be made to the underlying cell. Optical methods have advantages in spatial resolution, which is fundamentally limited by diffraction, and information transfer, as the underlying carrier wave has a frequency in the hundreds of THz. The most significant problem with optical techniques in general is absorption and scattering: tissue is not kind to optical radiation, and scatters it, typically with characteristics lengths of $< 1mm$ [90].

Direct Measurement of Optical Properties

It is possible to directly measure action-potential-induced changes in the physical size of neurons using optical coherence tomography [75, 2]. We will show, in the following chapter, how our metric of E_Q can be extended to use with OCT, and argues against its use as a method for detection of neural activity.

Voltage Dyes

There are also a multitude of small-molecule voltage-sensitive dyes that can be used for neural recording [6]. These are fundamentally similar to genetically encoded optical indicators, except that they are not produced by the cells themselves, but must be injected from an external source. Typically, they only last for 5-120 minutes before being removed, destroyed

by the cells, or photobleached [6]. For this reason, along with frequently being toxic, they are not typically used for long-term recording. This technique also requires relatively high optical intensities (often $> 100W/cm^2$), most of which is dissipated.

Genetically-Encoded Optical Indicators

A family of methods involves forcing neurons to produce fluorescent proteins that change their fluorescent properties in response to neural activity, usually by modifying the neuron's DNA.

Genetically-Encoded Calcium Indicators

One such optical indicator is the Genetically Encoded Calcium Indicator (GECI), which fluoresces in response to the inrush of calcium during an action potential [140]. Typically, a shorter wavelength is incident on these proteins, which absorb the incident photons and then, after a period of time, fluoresce and emit a longer-wavelength photon.

Genetically-Encoded Voltage Indicators

Even more recently, proteins which reside in the cell membrane which change their optical properties in response to the changing electric field across the cell membrane during an action potential have been developed [153, 7]. Referred to as Genetically Encoded Voltage Indicators (GEVIs) these have advantages of much faster response times. For both these indicators, the typical figures of merit used are $\Delta F/F$ (typically 5-20%), or the relative change in fluorescence due to an action potential, as well as the signal-to-noise ratio (SNR) (typically 5-20dB), and the response time (typically 5-200ms).

Another disadvantage of genetically encoded indicators is the requirement that they be genetically encoded: either the organism in question must be genetically engineered from birth, or the target neurons must have their DNA modified directly, typically using a virus to insert the desired genetic material into the target neuron's genomes [104]. As with voltage dyes, these indicators lose their fluorescence through photobleaching, and new copies must be synthesized by the cell, which takes time. This photobleaching timescale is currently in the range of minutes to hours, but has been improving. Also as with voltage dyes, the optical intensities required are typically very high.

Hybrid Optical Intensity Modulators

The approach to neural recording proposed in this dissertation is a hybrid electrical-optical method, using electrodes to capture the extracellular voltage caused by action potentials, which is transformed into a change in reflectance of optical modulators. This has the advantages of extremely fast response times - in practice, limited by the electrode dynamics, as well as the wireless nature of optical recording.

Low-Resolution Neural Sensing Technologies

On the other extreme end, many technologies exist to measure large-scale changes in the brain, with much lower spatiotemporal resolution. These include, to name a few, EEG [136], MEG [52], fMRI [5], and fNIRS [114]. These technologies and are regularly used in fun-

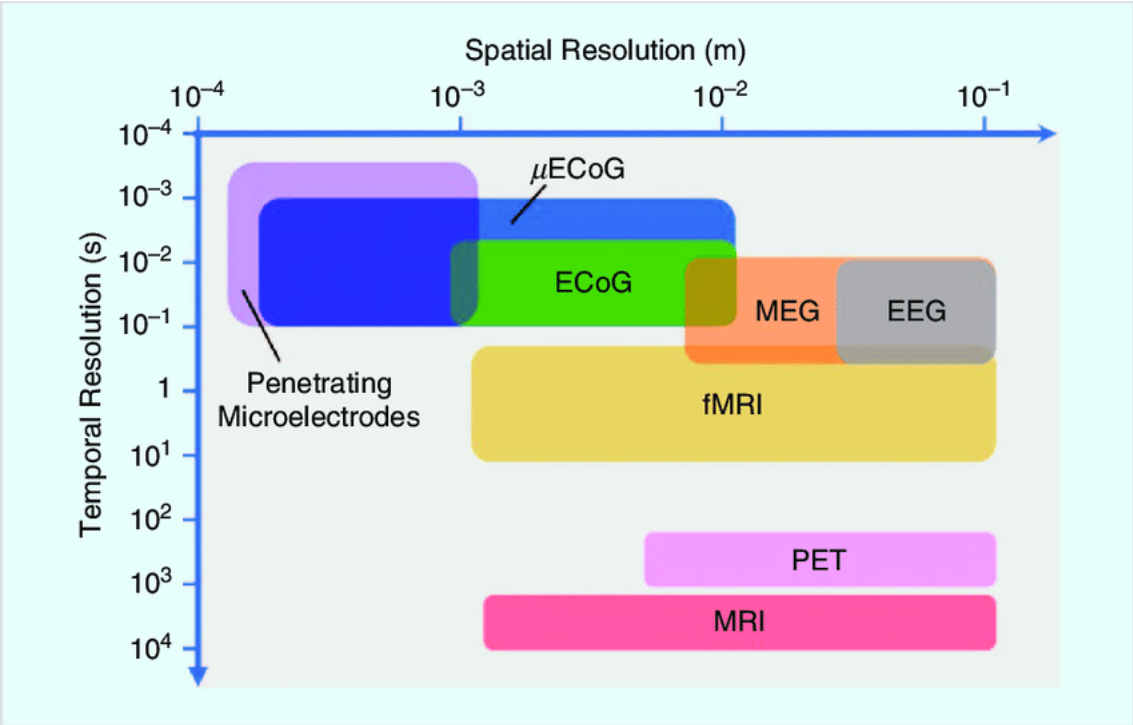

Figure 1.3: Spatiotemporal resolution of different neural recording technologies. Figure reproduced from [139].

damental science and as diagnostic tools in the clinic, and are able to measure large-scale patterns across the brain. EEG is regularly used to diagnose and localize seizures [9] and fMRI is a common tool to correlate brain regions with particular actions or behaviors. Collectively, their main limitation is in resolution: depending on the specifics of the technology and the setup, they typically measure with bandwidths $< 200Hz$ and spatial resolution of $> 1mm^3$. This captures the coordinated activity of millions of neurons, rather than the network dynamics at the individual-neuron level [64]. As such, they will not be a future focus of this work, but the reader should be aware of their existence.

1.3 Immune Response of Implantation

Implantation of devices inside the brain causes an immune response. During this immune response, glial cells migrate to the site and envelop the device, forming a "glial sheath" around the device [21]. This degrades recording performance, both by the migration or death of nearby neurons, as well as an increase in electrode impedance [125]. However, it

has been found that when the implanted device is sufficiently small (˜10μm or less) [109] or flexible [88], no detectable immune response is found. This has sparked a huge amount of activity in the development of micro-scale and flexible neural probes, which hold the promise of no long-term degradation in recording capability due to the immune response. Any sensor which can be made with a small enough cross-section may be able to take advantage of this "immune invisibility".

1.4 Optical Voltage Sensing for Microgrids

One application for optical voltage sensing which has been explored over the last few decades is voltage sensing in power grids. Power grids often require continuous monitoring of voltages in the range of tens or hundreds of kV. This is typically done via instrument transformers, which reduce the grid-level voltage to one that can be measured by conventional electronics, typically less than 100V. These transformers are simple, reasonably accurate (available in various accuracy classes in the range of 0.1 to 0.5%), robust to environmental conditions, and provide electrical isolation for the sensing equipment from the grid voltage. However, these transformers are bulky and expensive which limits their use when minimizing cost is important, such as in microgrids and distributed sensing. One potential solution is the use of optical voltage sensors, which retain the advantage of electrical isolation and can be manufactured at much lower cost.

Various authors have worked on building optical voltage sensors since the early 1980's [157]. Initial sensors were large lengths of optical fiber (several meters) wrapped around a piezoelectric material. When a voltage was applied to the piezoelectric material, it grew or shrank, increasing the fiber optical path length, which could be measured using interferometry. However, these initial sensors were very sensitive to temperature (˜300 rad/C for a 1m length of fiber) [157].

Other field-based measurement techniques were later developed [117] and commercialized [115] for high-voltage (500kV) substation sensing. Commercial solutions now exist in the ˜100's of kV range (Nxtphase), and more recently between 4-25kV (Eaton GridAdvisor) in the 0.5% accuracy class. However, these sensors remain expensive, and are not as accurate as their instrument transformer counterparts.

More recently, authors have used Bragg grating fibers, which reflect a very narrow slice of the incident spectrum. By physically coupling these filters to a piezoelectric material, a voltage can be transduced into a change in resonant wavelength by physically stretching the Bragg grating to shift the resonant frequency. However, these sensors are also very temperature sensitive (˜resonant shift of 10pm/C, or ˜1/5 of the linewidth per C), and require temperature compensation.

Different approaches have been attempted for temperature compensation , which achieved an accuracy of 30%, or closed-loop feedback [85], which has achieved much better performance (within +/- 0.1% error), and external Bragg gratings [48]. Closed-loop feedback via electrical temperature sensors poses an electrical isolation hazard, increases system com-

plexity, and decreases reliability. Bragg filter temperature compensation is an attractive approach as it does not have the same drawbacks as electrical temperature sensing. However, existing implementations using multiple sensors [154] suffers from poor accuracy (30% over the tested range), while another form of optical compensation has been shown to be accurate at DC near the bias point (0.2% accuracy over a 25C temperature range), but reported very high harmonic distortion (˜10% total for the first few harmonics), making it unsuitable for measuring AC signals.

Fundamentally, this temperature dependence is made worse in two ways: by increasing the quality factor of the interferometer or cavity used, or increasing the length of the cavity / interferometer. Temperature errors are proportional to the quality factor and inversely proportional to the length due to the linearity of expansion of materials. An ideal device would thus have a low quality factor (broad linewidth) and short cavity length. However, this is at odds with improving device sensitivity - a higher quality factor gives a higher device sensitivity. The same argument can be made for harmonic distortion - since a higher modulation depth uses more of the spectral shape of whatever reflective or transmissive elements are used, the shape of these deviates from a straight line, and nonlinearities become important.

These trade-offs can be encapsulated by the figure of merit we introduce in the following chapter - the Energy per Quanta, E_Q. A higher E_Q comes at the cost of increased harmonic distortion and temperature dependence. Here, we propose making the opposite such trade-off: deliberately increasing E_Q to achieve low temperature dependence and harmonic distortion.

1.5 Properties of the Ideal Sensor

Sensors, regardless of their application, share a common set of ideal properties, which we will lay out below. We will discuss *voltage* sensors, but the criteria below are common for a wide variety of sensors.

Linearity

First, the sensor is linear - the output of the sensor can be expressed as $a_{out} = H * V_{in}$, where a_{out} is the sensor output - whether it be a voltage, a current, a certain number of particles entering a detector, a charge, or anything else. V_{in} is the input voltage to be measured, and H is the transfer function of the sensor. Linearity allows for faithful reconstruction of the underlying signal - attempting to infer V_{in} from the measured value a_{out} is trivial - just divide by H. Sensors which are nonlinear make inferring the measured value more difficult, and, worse, make frequency-domain techniques less valuable by distorting the input signal. In practice, the typical sensor violates linearity, as there is an offset: $a_{out} = H * V_{in} + a_0$. Usually, however, the offset can be canceled or compensated for.

High Bandwidth

Second, the ideal sensor has high bandwidth. This means that the sensor can respond to fast changes - a simplification is that a bandwidth of $1kHz$ allows for a system to respond to changes on the order of $1ms$. The higher the bandwidth, the greater the capacity of the sensor for measuring faster and faster signals. Ideally, the bandwidth is infinite, but it need only be large enough to capture input voltages in the range of interest.

Low Noise

Third, the ideal sensor has low noise. Ideally, the noise is zero, but this is nonphysical and unrealistic. More realisticaly, the sensor noise should be close to that allowed by the underlying physical process. If the sensor uses photons or electrons for transduction, the underlying minimum noise is often shot noise. In other cases, it is thermal noise, or noise inherent to the process being measured itself.

Low Power Dissipation

Fourth, the sensor itself dissipates as little power as possible. This is especially critical in areas like the brain where any power dissipation leads to a local increase in temperature, and if this increase gets too high, subsequent cell death [151]. It is not necessary that the entire *sensing system* dissipates no power, but only that the elements of the system which are power-constrained dissipate no power.

Small Footprint

More generally, the ideal sensor has a small footprint on the system being sensed. In addition to dissipating a small amount of power, this might mean that the sensor is physically small, or requires minimal modification to the system being sensed.

Linear System-Level Tradeoffs

Ideal sensors also allow trade-offs between their different parameters. All else being equal, if the bandwidth of a system is increased, the noise will typically also increase. If the noise is white, this tradeoff is linear. Many systems (electronic and optical ones among them) also allow trading off power for bandwidth - consuming additional power allows electronic devices to more quickly charge system capacitances, and allows optical systems to increase their sampling rate with the same number of photons per sample. In the ideal sensor, each of these parameters may be traded off linearly for any other. These trade-offs encapsulate the need for a single quantity that captures their interaction - and I propose one which I describe in chapter 2.

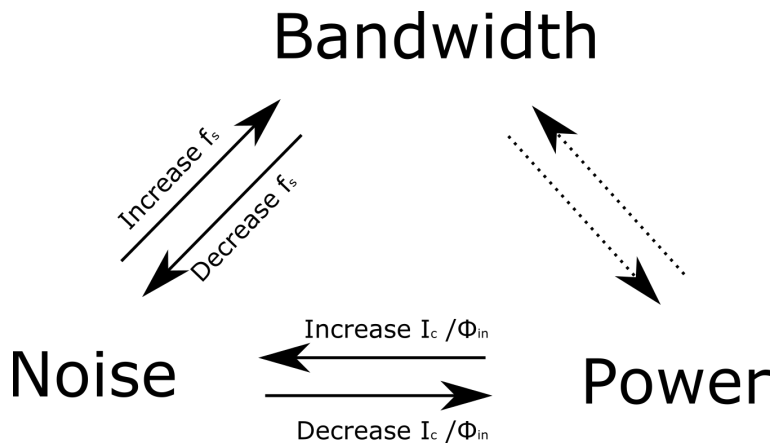

Figure 1.4: Power-SNR-bandwidth trade-off. Power/SNR and SNR/Bandwidth can be traded off directly by adjusting a single knob.

1.6 The Ideal Electronic Voltage Sensor

The lowest noise efficiency electronic voltage sensor (given by g_m/I_C) is the shot-noise-limited bipolar junction transistor. When operated in common-emitter mode, this "sensor" input is a voltage v_{in}, and it's output is a current i_{out}. To operate in its highest gain region, the amplifier needs to dissipate a certain amount of static power in the form of its collector current I_c. The transfer function H between the input and the output is simply the BJT's transconductance, g_m, which is approximately equal to I_c/ϕ_T [105], where $\phi_T = kT/q$ is the thermal voltage (k is Boltzmann's constant, T is the absolute temperature, and q is the fundamental charge unit. The lower limit for the noise of a BJT is shot noise, if thermal noise and other sources of noise can be made negligible, as they often can with proper design. These properties will be discussed further in the next chapter.

1.7 The ideal passive optical voltage sensor

There are many physical mechanisms for converting an electronic signal in the form of a voltage into an optical one, and there is not a single known 'ideal' sensor. However, regardless of the mechanism, the ideal noise properties of all optical systems are essentially similar - as with the BJT, the noise limit is given by (optical) shot noise. We will consider the case where the sensor itself is not responsible for photon generation, but there is an external source. In this case, we assume there is a total photon flux Φ_{in} incident on our sensor at an optical frequency $\hbar\omega$. Of this, some fraction A_0 is dissipated by the sensor itself. Some other fraction (at most $1 - A_0$) returns from the sensor and hits the detector. This fraction may be less than $1 - A_0$, if, for example, light is disspated elsewhere interferometrically or via

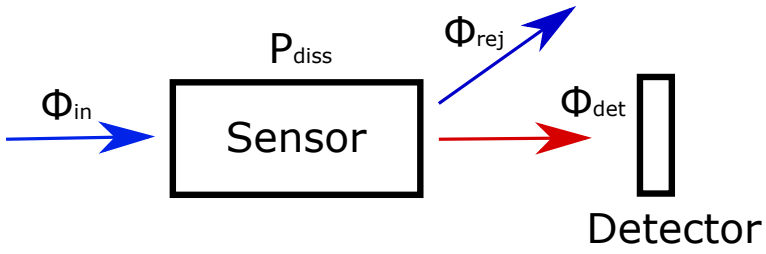

Figure 1.5: Ideal optical circuit with an input optical flux Φ_{in}, a rejected optical flux Φ_{rej}, an optical flux reaching the detector Φ_{det}, and sensor-dissipated power P_{diss}.

scattering away from the detector. Since there is no such thing as a negative photon, there will be a DC fraction of the incident photons R_0 on the detector, and another (time-varying) fraction which contains information about the signal to measure. The noise at the detector, if we assume all other sources are negligible, is just the shot noise of the photons reaching the detector.

1.8 Dissertation Contribution

There are three main contributions of this dissertation: (1) Development of a novel figure of merit - the Energy per Quanta or E_Q which allows quantitative performance comparison of devices within a fixed application (such as neural sensing or grid voltage measurement) which may involve very different fundamental physical processes, and a design framework surrounding this figure of merit. (2) Work towards the design and fabrication of a fully-passive free-floating micron-scale neural voltage sensor. Specifically, a fabrication process capable of making in-plane high-Q factor optical resonators.

1.9 Dissertation Organization

This dissertation is organized into several chapters. Chapter 1 is an overview of the problems tackled herein: neural recording and grid sensing, as well as the technologies that currently exist to tackle these problems. Chapter 2 discusses the first attempt at a piezoelectric optical voltage sensor, and provides a brief overview of the relevant physics, with extensive discussion of fabrication and characterization. Chapter 3 discusses the characterization setup built for the purposes of measuring this and future sensor. Chapter 4 discusses the addition of Bragg Mirrors to the sensor described in chapter 2, including the fabrication process optimization. Chapter 5 includes work done on developing the piezoelectric material

Chapter 2

Cross-Modality Comparison for Neural Sensing Using E_Q

In this chapter, I will present a novel, quantitative metric for multi-modality comparison of neural recording technologies - the energy per quanta, or E_Q. This metric is a direct measure of the scalability of a neural recording technology - lower values of E_Q correspond to a lower energy cost to extract the same signal, and allow for lower power dissipation per channel at a specified bandwidth. I describe comparisons for several electronic measurement systems, with a focus on electronic and optical methods of neural recording, and calculate E_Q values for the optical modulators discussed in this dissertation, several electronic systems, a representative GEVI, and optical coherence tomography.

2.1 The Energy per Quanta: Optimal devices Pareto front

Given the linear trade-offs between power, noise, and bandwidth, there are an infinite set of "optimal" devices - some designers might prefer a low noise, others might prefer low power, yet others might prefer a wide bandwidth. Trading off these parameters against each other forms a Pareto front [69]. To meaningfully compare different devices for a fixed application two of these three quantities must be fixed - and then the third can be used as a figure of merit. However, rarely is this feasible, as designs will in practice differ substantially on all three quantities. Alternatively, a composite metric can be developed, one which does not depend on where the device is on the Pareto front. This is analagous to a "radius" for the Pareto front. I propose such a metric, which I call the Energy Per Quanta, E_Q, and it is defined as follows:

$$E_Q \equiv \frac{P_{diss}}{SNR_{max} * BW} \tag{2.1}$$

Where P_{diss} is the dissipated device power (per channel, if there are multiple channels), SNR_{max} is the maximum achievable signal-to-noise ratio, and BW is the device bandwidth.

Why SNR and not Pure Noise?

One might reasonably object to this definition, as it includes the expected signal amplitude in its definition - and this is not a property of the device or technology being used. Whether to include the signal level in a figure of merit depends on whether you want to fix the *application* or the *architecture*. In the case of neural recording, the *application* is fixed: our goal is to measure neural activity. However, the input noise, by itself, is meaningless. If we want to compare very different technologies, such as analog circuits and GEVI's and optical voltage sensors and patch clamps (as this work aims to do), then what we care about is not the absolute noise, but the achievable SNR. The underlying signal level can vary dramatically depending on which technology it is aquired with - with a patch clamp the expected signal value is ˜$100mV$, with extracellular electrodes, ˜$100\mu V$. Technologies that are able to measure or utilize a larger voltage (such as GEVIs) have a higher signal power than those that measure a much weaker signal i.e. extracellular recording devices. This difference in *measured* signal level is not accounted for if only the noise is taken into the figure of merit. The disadvantage of using SNR, rather than pure noise, is that it only enables meaningful sensor comparison for a particular application. A sensor capable of multi-application use may have different possible maximum SNR.

Shot Noise - the Ultimate Limit

Both electronic methods and optical sensing metods suffer from the same fundamental source of noise - shot noise. While other sources of noise exist, and may be dominant in a particular case, shot noise represents the ultimate lower limit in achievable noise, and will feature heavily in the following discussions.

The absorption of photons and the emission of electrons over a barrier is a fundamentally random process: it has some chance of happening per unit time. In the case of electrons, the origin of this randomness is thermal motion: only some very high-energy electrons will be able to overcome an energy barrier, and the probability that any given electron will have that energy is constant over time. In the case of photons, spontaneous emission and stimulated absorption are processes which, averaged over many optical cycles, for a constant-strength electric field, have a constant probability of happening per unit time.

Constant probability per unit time does not imply that in a given unit of time there will always be the *same* number of events: one second 10 photons might be emitted, the next 9, the next 11. This type of random process is called a "Poisson Process" [74], and the result is quite simple: the average noise power is simply equal to the average number of quanta (photons or electrons) received in a given unit of time.

The Origin of Shot Noise

Optical and electronic sensing works by measurement of discrete quanta: electrons, in the case of electronics, and photons in the case of optical devices. The process of absorbing or emitting photons is fundamentally a quantum process, and the process of absorption happens in discrete events. Similarly, the process of passing electrons through a barrier also creates discrete events which cause shot noise.

Properties of E_Q

E_Q has a few nice properties which make it especially useful and suitable as a figure of merit when the application is fixed.

Can be Interpreted as a Literal Energy Cost for each Signal Quanta that Arrives at the Detector

The units of E_Q are, as the name would suggest, energy. Usingy Nyquist's theorem, the P_{diss}/BW term can be interpreted (within a factor of 2) as the energy expended to obtain a single sample of data. For any detector which uses "quanta" (that is, electrons or photons), the SNR in a shot-noise-limited system is equal to the number of quanta which arrive at the detector. Hence, E_Q is the literal energy cost to obtain a single quanta of information at the detector. Higher values of E_Q mean a higher energy cost, which is generally undesirable.

Defined in Terms of Measurable System Variables

SNR, bandwidth, and dissipated power are all measurable quantities, E_Q can be meaningfully defined for any system, regardless of the specifics of its implementation. Optical, electrical, chemical, etc. E_Q may be defined for any such system.

Allows for Meaningful Cross-Technology Comparison

Precisely because E_Q is defined in terms of measurable system-level quantities, it allows quantitative comparison to be made between an optical recording method and an electronic one, regardless of the details of implementation of those methods. Or between one optical method (i.e. optical modulators) and other (i.e. fluorescent indicators). This is a *good* metric of comparison because energy is limited when performing neural recording - any energy dissipated in neural tissue will cause heating of that tissue if not removed. If infinite energy were available to perform neural recording, this would not be a suitable metric.

Can be Bounded From Below Theoretically

Relations between power dissipation, SNR, and bandwidth can usually also be predicted theoretically, and so theoretical bounds can be set on the achievable E_Q for a given technology. We will see, for example, that optical technologies cannot have an E_Q less than the

photon energy, and electronic technologies cannot have an E_Q less than the electron energy when the signal voltage and thermal voltage ϕ_T are equal.

Does not Require Noise to be White

Rather than defining E_Q in terms of a noise power spectral density, defining it in terms of total noise power over some bandwidth allows for a definition even when noise is not white. However, in this case, E_Q will be a function of bandwidth, rather than invariant.

Invariant for Changes in Bias Current and Bandwidth in BJTs

The shot-noise-limited bipolar junction transistor is as close as electronics has come to the ideal amplifier, and is superior to the MOSFET. The primary unavoidable source of noise in this transistor is the shot noise from crossing the collector-emitter junction [25]. In the common-emitter configuration, the dissipated power, output-referred current noise power, and output-referred signal power are as follows:

$$P_{diss} \approx I_c * V_{cc} \tag{2.2}$$

$$I_n^2 \geq 2 * q * I_c * BW \tag{2.3}$$

$$I_{sig}^2 = (g_m * v_{in})^2 \approx \left(\frac{I_c}{\phi_T} v_{in}\right)^2 \tag{2.4}$$

Where $g_m \approx \frac{I_c}{\phi_T}$ is the transistor's transconductance, I_c is the DC collector current, V_{cc} is the supply voltage, q is the elementary charge, v_{in} is the input voltage, and $\phi_T = \frac{kT}{q}$ is the thermal voltage, where k is Boltzmann's constant and T is the absolute temperature. By applying the definition of E_Q in equation 2.1, we find that the collector current and bandwidth cancel out, and energy per quanta of a BJT is:

$$E_Q = 2 * q * V_{cc} * \left(\frac{\phi_T}{v_{in}}\right)^2 \tag{2.5}$$

This quantity depends only on the energy that each *quanta* carries - in this case, the electron, as well as the signal power and the temperature. It is independent of the collector current. You may increase the collector current to improve the SNR or bandwidth, but at the cost of increased power consumption. E_Q is thus a composite metric that does not depend on the specifics of how the circuit is implemented. For MOSFETs, considering only channel noise, it can be shown that E_Q takes a very similar form:

$$E_Q = 2 * q * V_{DD} * \frac{\phi_T V_{ov}}{v_{in}^2} \tag{2.6}$$

Where the overdrive voltage replaces one of the thermal voltages in the numerator. This is inferior to the BJT for overdrive voltages greater than the thermal voltage (i.e. when a MOSFET behaves like a MOSFET).

Invariant for Changes in Incident Optical Intensity and Sampling Frequency in Shot-Noise-Limited Optical Modulators

Any optical modulator must contend with shot noise, as the carriers of information are quanta - in this case, photons.

$$P_{diss} = A_0 = P_{in} \tag{2.7}$$

$$P_{noise}^2 = \eta 2 * \hbar\omega R_0 P_{in} * BW \tag{2.8}$$

$$P_{sig}^2 = (\eta\beta * v_{in} * P_{in})^2 \tag{2.9}$$

Where P_{in} is the incident optical power, A_0 is the fraction of optical power absorbed and dissipated by the modulator, R_0 is the fraction of DC photons returned to the detector, BW is the bandwidth, β is the "device gain", $\hbar\omega$ is the photon energy used, and η is the quantum efficiency of the detector used.

These equations are very similar to those of the BJT - because they represent an analagous system. P_{in} is analagous to I_c, and $R_0\hbar\omega$ is equivalent to $q * V_{cc}$. Computing the E_Q from its definition in equation 2.1, the energy per quanta is equal to:

$$E_Q = \frac{A_0 * R_0}{\eta} * 2\hbar\omega * \left(\frac{1}{\beta v_{in}}\right)^2 \tag{2.10}$$

Bandwidth may be increased by increasing the photon sampling rate, but at the expense of SNR. SNR may be improved by increasing the incident power, but at the expense of increased power dissipation. E_Q captures these trade-offs elegantly in a single quantity.

Invariant for changes in incident optical intensity and sampling frequency for single- and multi-photon fluorescent voltage indicators

A similar analysis can be performed for single- or multi-photon fluorescent systems. In these cases, the main difference between these and optical modulators is that the output wavelength is shifted relative to the input wavelength. This allows for filtering the light reaching a detector. In principle, all the light which does not interact with fluorescent indicators may be rejected, and we will use this assumption here. When narrowband sources such as lasers are used, in practice this is easy to do. In this case, it is easier to work with the signal power and noise power in terms of the number of photons, rather than the actual optical power.

$$P_{diss.} = A_0 * \hbar\omega_{in} * N_{in} \tag{2.11}$$

$$N^2_{noise.} = 2 * S_0 * \eta \frac{N_{in}}{m} \tag{2.12}$$

$$N^2_{sig.} = \left(S_0 * \eta \frac{N_{in}}{m} \beta v_{in}\right)^2 \tag{2.13}$$

Here $\hbar\omega_{in}$ is the *incident* photon energy, A_0 is the fraction of optical power which is absorbed, due to the shift in wavelength - which bounds its value from below - as well as optical power which is uselessly absorbed by nearby tissue, fluids, and non-fluorescent molecules. N_{in} is the incident number of photons, S_0 is the fraction of photons that are absorbed by a fluorescent molecule *and* return to the detector, β is the hypothetical "gain" of the fluorescent system, m is the number of photons which must be absorbed to generate a single output photon (1 for single-photon fluorescence, 2 for two-photon, 3 for 3-photon, and so on), and v_{in} is the voltage dropped across the molecule (if the protein is embedded in the cell membrane, then this is O(100mV)). Plugging these into the E_Q definition equation 2.1, we find:

$$E_Q = \frac{A_0}{\eta S_0} * 2\hbar\omega_{in} * \left(\frac{1}{\beta v_{in}}\right)^2 \tag{2.14}$$

Typically, however, β and v_{in} are not used in these applications, as they are challenging to measure. Instead, what is usually measured is $\Delta F/F_0$, or as I have written it, $\beta v_{in}/S_0$. Rewriting in terms of these quantities:

$$E_Q = \frac{A_0}{\eta S_0^3} * 2m * \hbar\omega_{in} * \left(\frac{1}{\Delta F/F_0}\right)^2 \tag{2.15}$$

As with all other E_Q, this does not depend on the incident optical intensity or sampling rate. The incident optical intensity may be increased to improve SNR, at the expense of dissipated power. The sampling rate may be increased to improve the bandwidth, but this will degrade SNR. Sampling rate and incident optical power may be simultaneously increased to improve bandwidth without degrading SNR, but at the expense of power. The piper must be paid.

Invariant for changes in incident optical intensity and sampling frequency for OCT- and other interference-based methods

Changes in optical path length can also be measured interferometrically. In the case of OCT [2], a small change in the diameter of a neuron results in a change in path length from the wavefront reflected off the surface of the neuron.

In OCT, light is reflected at various depths, yielding a total DC reflectance R_0. If a particular neuron is of interest, we will call the amount of light reflected from that neuron R_1, and is approximately equal to:

$$R_1 \approx \left(\frac{\Delta n}{n_0}\right) \tag{2.16}$$

Where Δn is the difference in refractive index between the interior and exterior of the cell, and n_0 is the average refractive index of the two. Ultimately, the change in optical path length $\Delta\phi$ from an action potential will be interfered with a reference beam. In the best case, to minimize noise, the reference beam has the same amplitude as the returning sample beam and a phase that is different by $\pi/2$.

In this case, the electric field of the two beams interfered together is proportional to $\Delta\phi$, and the optical intensity is proportional to $\Delta\phi^2$. If each beam (reference beam and sample beam) has a total optical power of $R_1 * P_{in}$, the proportionality constant is $R_1 * P_{in}$. OCT is also inherently multi-channel, so the power dissipation per channel is lower by a factor of $N_{channels}$ than the total dissipated power.

$$P_{diss} = A_0 * P_{in}/N_{channels} \tag{2.17}$$

$$P^2_{noise} = 2\hbar\omega R_0 P_{in} * BW \tag{2.18}$$

$$P^2_{signal} = \left(R_1 * P_{in}\Delta\phi^2\right)^2 \tag{2.19}$$

Plugging in these equation to the definition of E_Q:

$$E_Q = \frac{A_0 \cdot R_0}{2\eta R_1^2}\frac{1}{N_{channels}} * \hbar\omega * \left(\frac{1}{\Delta\phi}\right)^4 \tag{2.20}$$

For a typical change of 1nm [2] in cell diameter, 1300nm incident wavelength, an average refractive index of 1.36 [86], and a difference in refractive index of 0.02 [86], the change in phase, R_1 is expected to be $2 * 10^{-4}$ and $\Delta\phi$ is expected to be $6.5 * 10^{-3}$. We will assume an effective number of channels of 400, equal to the typical imaging depth [28] of 2mm, divided by a typical resolution of $5\mu m$ [28].

2.2 Cross-Modality Comparison: Power, Noise, Bandwidth, Size

Using the E_Q metric, we can quantitatively compare these dramatically different technologies, which I have attempted to do in the table below, taking a few of the more recent advances for each individual technology.

Technology	Author	P_{diss}/ch	$N_{channels}$	SNR	BW	E_Q
Optical Modulator	[41]	$55\mu W$	1	35dB	5kHz	20mJ
Optical Modulator	This work	$400\mu W$	1	5dB	4MHz	$200\mu J$* / $1.3\mu J$**
Ultrasound[2]	Seo [128]	$120\mu W$	1	-5dB	30kHz	$13\mu J$
GEVI (Archon 2)	[111]	$320\mu W$	1	28dB	1.6kHz	320pJ.
fOCT[1]	$P_{in}/N_{channels}$	400	-	-	-	540 μJ
Electronic	Musk [102]	$360\mu W$	1536	25dB	9.6kHz	$130pJ$
Electronic	Lee [81]	$1\mu W$	1	16 dB	15kHz	4.5pJ
Electronic	Ghanbari [46]	38 μW	1	26 dB	5kHz	22pJ

1 Theoretical best-case values, assuming $R_0 = 0.5$ and $A_0 = 0.5$ with an incident wavelength of $1300nm$.

2 Power consumption taken as total TX power.

Table 2.1: *As-measured in measurement setup, **With improved optical coupling (not attempted in this work. All extracellular E_Q values rescaled to use a $100\mu V$ input signal.)

We can see from table 2.1, that current fully-integrated electronic amplifier-based systems substantially outperform the best demonstrated AlN-based optical modulator for this purpose. We also see that current GEVI's, such as Archon, come close to the performance of electronic systems, and further improvements may make them more competitive. We also see that the theoretical value of E_Q for fOCT is the worst evaluated. In practice, other noise sources are likely to make the E_Q worse. For this reason, I do not believe OCT is an attractive method for measurement of neural activity.

Chapter 3

Low-Q Optical Voltage Sensor for Grid Voltage Sensing

This chapter was modified with permission from [41].

3.1 Abstract

Continuous monitoring of voltages ranging from tens to hundreds of kV over environmental conditions, such as temperature, is of great interest in power grid applications. This is typically done via instrument transformers. These transformers, although accurate and robust to environmental conditions, are bulky and expensive, limiting their use in microgrids and distributed sensing applications. Here, we present a millimeter-sized optical voltage sensor based on piezoelectric aluminum nitride (AlN) thin film for continuous measurements of AC voltages $< 350kV_{rms}$ (via capacitive division) that avoids the drawbacks of existing voltage-sensing transformers. This sensor operated with $110\mu W$ incident optical power from a low-cost LED achieved a resolution of $170mV_{rms}$ in a 5kHz bandwidth, 0.04% second harmonic distortion, and a gain deviation of +/-0.2% over the temperature range of ˜20-60°C. The sensor has a breakdown voltage of 100V, and its lifetime can meet or exceed that of instrument transformers when operated at voltages $< 70kV_{rms}$ with capacitive division. We believe that our sensor has the potential to reduce the cost of grid monitoring, providing a path towards more distributed sensing and control of the grid.

3.2 Introduction

Safe, accurate, and economical measurement of time-varying voltages in electric power systems is a significant challenge. The standard solution to this challenge is the instrument transformer, which steps down a high voltage (˜1kV+) to an appropriate level (typically $< 100V$) and isolates the stepped-down voltage, allowing safe measurement using conventional electronics. However, these transformers are bulky and expensive and sometimes

explode (˜3% of all installed instrument transformers [113]). Optical methods for direct measurement of high voltages have gained attention since the early 1980s [157], mainly due to the high available bandwidth (˜GHz), intrinsic electrical isolation, and the potential for low cost and remote monitoring. Initial optical voltage sensors consisted of long (several meters) optical fiber wrapped around a piezoelectric material [157]. In these sensors, when the piezo material was excited by a voltage, it grew or shrank proportional to the applied voltage, changing the fiber optical path length; the resulting change in the optical path length was measured using interferometry to infer the voltage amplitude. Fiber-based current and voltage sensors are currently in commercial service, such as those developed by NxtPhase [14, 116, 126]. However, these sensors are fairly complex and expensive, requiring, among other things, optical delay lines and birefringe modulators. As of this writing, they have failed to gain widespread adoption. Optical sensors with reduced cost, simplified architecture, and inherent environmental robustness may change this. More recent optical methods attempting to do just this include coupling piezoelectric material to the resonant frequency of Bragg fiber gratings [127, 154, 48]. However, the sensors based on the above optical methods are inaccurate due to their nonlinearities [48] or temperature sensitivity [154, 48, 24]. Closed-loop compensation methods were used to improve the accuracy of such sensors [61, 85], but these methods reduce system reliability and increase system complexity, electrical hazards, and cost.

Fundamentally, output nonlinearity and temperature dependence increase with an increasing quality factor (Q) of the interferometric- or optical cavity-based sensor, where Q is defined as the ratio of the energy stored in the cavity to the energy dissipated per oscillation period. Therefore, a low-Q sensor is desirable to minimize nonlinearity and temperature dependence. However, low-Q sensors suffer from low sensitivity, resulting in a low signal-to-noise ratio (SNR) and poor performance. In practice, the SNR, and hence performance, can be improved independently of the sensor sensitivity by increasing the incident optical power (P_{in}) or reducing the operating bandwidth (BW). This makes performance comparison of different optical voltage sensors difficult; they must be operated with the same P_{in} and BW for a fair comparison.

Given this, a P_{in} and BW independent figure of merit (FoM) is required to quantify the noise performance of optical voltage sensors. We propose such a metric, the *energy per quanta* (E_Q), which depends only on the sensor properties and dynamic input range. This metric extends the power efficiency factor used in analog circuit design [101] and the energy per conversion-level of analog-to-digital-converters [147, 146]. Lower Q-factors yield less sensitive sensors and a higher E_Q ($E_Q \propto 1/Q^2$).

In this work, we propose trading off E_Q for temperature insensitivity and reduced harmonic distortion, and explore the limits of this approach. To this end, we demonstrate a low-Q resonant optical cavity-based voltage sensor based on a piezoelectric AlN thin film that transduces a voltage applied across the piezo terminals into a change in the resonant frequency of the cavity. This sensor can be batch fabricated with high yield and low cost (<\$1), which makes it uniquely well-suited to reduce the cost of grid voltage measurement.

3.3 Optical Voltage Sensor Design and Sensor Fabrication

Operating Principle and Fabrication Process of the Sensor

Fig. 1 shows the operating principle of the proposed optical voltage sensor (OVS) based on changes in the measured reflectance of a resonant cavity, whose thickness varies with applied voltage. The resonant cavity is formed by an AlN thin film sandwiched between the top indium tin oxide (ITO) electrode and the bottom silicon (Si) substrate. During operation, the sensor is illuminated by a light source with an incident optical power (P_{in}) at a fixed wavelength (λ_{in}) near the resonance wavelength of the cavity (λ_r). Some fraction of P_{in} is reflected from the cavity, with the remainder dissipated in or transmitted through the cavity, as seen in Fig. 3.1. Here, the intensity of reflected light (P_r=P_{in}×R, where R is the cavity reflectance) is measured by a photodetector to detect the amplitude of an input voltage (V_{in}) applied across the cavity. P_r depends on V_{in} through R as V_{in} generates an electric field in the cavity that changes the AlN film thickness [89] and refractive index [49] and hence results in a shift in resonant wavelength $\Delta\lambda = \lambda_r - \lambda_{r0}$, where λ_{r0} is the resonant wavelength at $V_{in} = 0$. The resulting λ_r shift leads to a change in the reflectance (ΔR =R-R_0, where R_0 is the reflectance at V_{in}=0). The ΔR value at a known V_{in} can be calculated using the following expression (see supplementary material for the derivation of (1)):

$$\Delta R = \beta V_{in} \tag{3.1}$$

$$\beta = \frac{3\sqrt{3}}{4}\frac{R_{max}Q}{t}\left(d_{33} + \frac{1}{2}n_0^2 r_{33}\right) \tag{3.2}$$

where R_{max} is the amplitude of the resonant dip (always < 1), Q and t are the quality factor and thickness of the cavity, d_{33} is the thickness mode piezoelectric strain coefficient, n_0 is the (unperturbed) refractive index of the AlN thin film, and r_{33} is the Pockels coefficient (which relates the refractive index to the applied electric field). This equation is valid near $\lambda_{in} \approx \pm\lambda_r + 1/\sqrt{3}\cdot FWHM$; this corresponds to the steepest point on the reflectance curve, where $FWHM$ is the full-width-half max of the resonant dip.

Fig. 3.2(a) shows the OVS fabrication process. All lithography was done in a DUV stepper (ASML), on a 150mm-diameter silicon (Si) wafer. We first sputtered a ˜300nm-thick layer of backside aluminum (Al), and then annealed the wafer at 300°C for 15min in atmosphere to create backside ohmic contacts that serves as the bottom electrode. Next, we deposited a 300nm-thick aluminum nitride (AlN) film (endeavor AT) and a 20nm-thick indium tin oxide (ITO) film to serve as the transparent top electrode. Finally, we patterned the top ITO contacts to form 2mm diameter devices and evaporated 20nm titanium (Ti)/300nm Al to form bond pads. The fabricated $10\times10mm^2$ sensor die was attached with conductive silver epoxy and wire-bonded to a printed circuit board (PCB). An optical micrograph of the 2mm diameter sensor on its PCB are shown in Fig. 3.2(b) The die size was designed

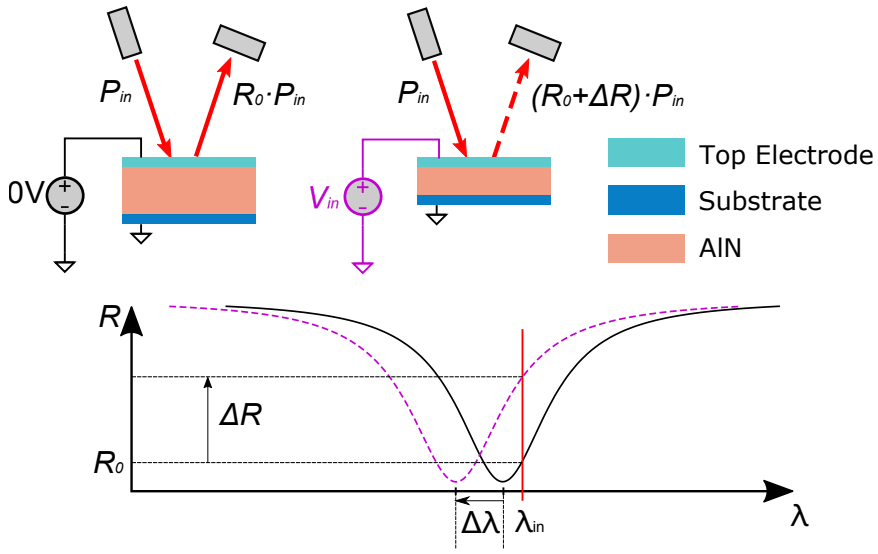

Figure 3.1: Sensor principle of operation. During operation, the sensor is illuminated by an incident optical power P_{in} at a wavelength λ_{in} (that is, the steepest point of the reflectance curve) offset from the resonant wavelength λ_r. An input voltage V_{in} applied across the cavity results in a shift in the resonant spectrum from the solid (unperturbed) to the dashed (perturbed) spectrum, causing a change in the reflectance ΔR at λ_{in}. The resulting ΔR change alters the amount of reflected light power P_r=$(R_0$+$\Delta R) \cdot P_{in})$ that is measured to determine the V_{in}.

to be larger than the actual device size to facilitate easy handling. See the supplementary information for specific process details.

Optical Shot Noise Sets a Hard Limit on OVS Performance. The performance of optical detection systems is bounded by shot noise received at the photodetector. For a shot-noise limited OVS system, the system SNR is proportional to the input optical power, which is given by $I_{pd}^2/I_{noise}^2 = \Delta I_{pd}^2/(2q(I_{pd} + \Delta I_{pd}) \cdot BW)$, where q is the electron charge, I_{pd} is the light-induced photocurrent on the photodetector, ΔI_{pd} is the change in the rms photocurrent induced by an applied input voltage (V_{in}), and BW is the system bandwidth. The SNR can also be expressed in terms of an an incident light power (P_{in}), the average device reflectance ($\langle R \rangle$), and an rms modulation depth (ΔR, that is, the change in R due to an applied V_{in}) in the following form, using $I_{pd} = P_{in} \cdot \Re \cdot R$:

$$SNR \approx \frac{P_{in} \cdot \Re \cdot \Delta R^2}{2q \cdot \langle R \rangle \cdot BW} \tag{3.3}$$

where $\Re$ is the responsivity of the photodetector. Eq. 3.3 reveals the system SNR depends not only on V_{in} through R but also P_{in} and BW, consistent with the expectation that larger

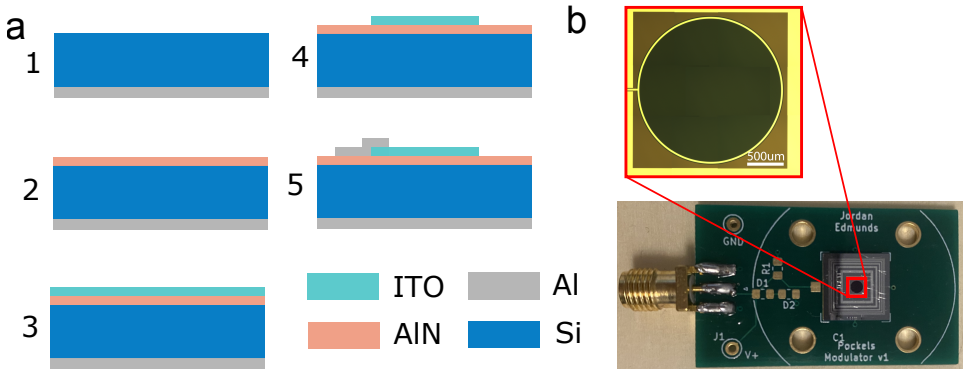

Figure 3.2: (a) Sensor fabrication process. 1. Backside Al (˜300nm) sputter and anneal at 300°C for 15 minutes 2. Frontside AlN (300nm) reactive sputter 3. Topside ITO (20nm) sputter 4. Lithographic patterning, ion milling of device mesas 5. Bond pad patterning Ti/Al (20nm/300nm) evaporation and lift-off. (b) Photograph of the fabricated sensor die on a PCB, and (inset) optical top-view micrograph of the sensor.

P_{in} and smaller BW provide a better SNR in the optical voltage sensing system. However, optical sources can only supply a limited amount of power, and system-level requirements could potentially limit the maximum P_{in} and the minimum BW in the system.

The Input-Referred Energy per Quanta (E_Q) Allows Quantitative Configuration-Independent Sensor Comparison. Since the best choices for P_{in} and BW will vary by application, it is useful to introduce a metric that normalizes SNR to P_{in} and BW. This would allow a rigorous noise performance comparison between optical voltage sensors, independent of the sensor's particular operating P_{in} or BW. In digital systems (optical and otherwise), the energy per bit has become a ubiquitous metric of device performance [141]. Here, we propose an alternative metric for analog systems, the energy per quanta (E_Q), defined as:

$$E_Q \equiv \frac{P_{in}}{\mathrm{SNR} \cdot BW} \tag{3.4}$$

This metric demonstrates how efficiently P_{in} is used, and can be interpreted as a cost paid in energy to achieve a desirable SNR. For a shot-noise limited system, E_Q can be derived, independent of P_{in} and BW, by inserting Eq. 3.3 into Eq. 3.4:

$$E_{Q,min} \approx \frac{q \cdot \langle R \rangle}{\Re \cdot \Delta R^2} \tag{3.5}$$

The form of this equation makes it clear that reducing the average cavity reflectance $\langle R \rangle$ at the expense of modulation depth ΔR at the operating wavelength can improve the noise

performance of the system, as observed in [77]. E_Q is bounded from below by the incident photon energy captured by the photodetector.

Optical Voltage Sensor (OVS) Design

We designed our OVS to measure grid-level AC voltages in the range of tens to hundreds of kVs via capacitive division. The system bandwidth (BW) was set to 5kHz, satisfying the BW requirement of most grid applications, including inverter-based solar [42]. To minimize sensor nonlinearity and temperature sensitivity, we chose to deliberately design the sensor to have a high E_Q, as $d\hat{\beta}/dT \propto 1/E_Q$ and $d\hat{\beta}/dV_{in} \propto 1/E_Q$ (where β is the sensor gain, shown in Eq. 2, and $\hat{\beta} = \beta(V_{in}, T)/\beta(0, 0)$ is the normalized gain). To achieve a high E_Q, we minimized the sensor Q-factor by designing the cavity as thin as possible ($L = 300nm$) and excluding mirrors (other than the material interfaces).

Furthermore, most previous sensors use lead zirconate titanate (PZT) as a piezoelectric material to form the resonant cavity as it provides a large piezoelectric strain coefficient ($d_{33} \approx 500pm/V$ [29], $d_{31} \approx 200pm/V$ [83]). However, the piezoelectric coefficients of PZT are extremely temperature sensitive, and changes between ˜5-25% over a $60°C$ range, depending on the orientation and method of measurement [83, 152]. In our sensor system, the system gain almost linearly changes as a function of the piezoelectric coefficient. Therefore, for example, a 1% change in the d_{33} coefficient due to temperature variation will cause an ˜1% error in the estimated input voltage. Therefore, here we elected to use aluminum nitride (AlN) for our sensor to further minimize the sensor temperature dependence as its d_{33} is independent of temperature [65, 121].

Fabrication Process

A detailed fabrication process is given below with settings used, using the same numbering as Fig. 2 in the main text.

Prior to this process, we deposited 0.3um of UV210 resist, and patterned alignment marks for layer-to-layer alignment in the ASML 5500/300 DUV stepper. We then etched 6 alignment marks into the Si substrate in a Poly-Si etcher (TCP 9400SE Lam Research) with 300W plasma power, 150W substrate power, 50 standard cubic centimeters per minute (sccm) chlorine gas (Cl_2), 150sccm hydrogen bromide (HBr), 4sccm helium (He), at a process pressure of 12mTorr for 40 seconds, with a target etch depth of 120nm.

1. Al Sputtering and Anneal- We first sputter etched the backside of the wafer at 200W RF Power / 10sccm argon (Ar) flow for 120 seconds. Without breaking vacuum, we then sputtered ˜300nm of aluminum (Al) at 4.5kW DC power / 15sccm Ar flow for 60 seconds. Target was conical and 99.999% pure. To make the backside Al contact ohmic to minimize contact resistance and nonlinearity, we annealed the wafer in an AccuThermo 610 at 300°C for 15 minutes, with an initial ramp rate of ˜20°C/s at

atmospheric pressure flowing pure N_2. The specific contact resistance at 0V dropped from $600M\Omega \cdot \mu m^2$ to $5M\Omega \cdot \mu m^2$ (˜110x improvement), and became linear instead of rectifying.

2. AlN Deposition - Prior to each deposition, we ran a bare Si conditioning wafer Al sputter to remove built-up AlN from the target. We sputtered the target at 4kW / 6sccm Ar flow for 5 minutes, followed by a brief (30 second) period where we introduced N_2 gas, flowing 7 sccm Ar and 22 sccm N_2 at 4kW power. We then sputter etched the frontside of the device wafer at 200W RF Power / 10sccm Ar flow for 120 seconds. Without breaking vacuum, we then sputtered ˜300nm of aluminum nitride (AlN) at 4.5kW, flowing 7sccm Ar and 22 sccm N_2 for 420 seconds at a chamber pressure of 5mTorr. The chamber used was dedicated to only deposit AlN.

3. Topside ITO Deposition - We deposited ITO from an 8" target at 500V DC and 2mTorr, flowing 40sccm of Ar into the chamber for 60 seconds, for a target ITO thickness of 20nm.

4. Topside ITO patterning - We spun on 0.3um of UV-210 photoresist in an automated Picotrack system, exposed a full wafer of 10mm x 10mm dies in an ASML DUV stepper (Model 5500/300) at $20mJ/cm^2$ dose. We then developed the resist in MF26A developer for 60 seconds in the same automated Picotrack coating system, and UV-hard baked the resist in a Fusion Systems M200PCU. We then ion milled the ITO layer in a Pi Scientific ion mill at 500V / 250mA at an angle of 20 degrees away from normal incidence for 60 seconds, followed by an etch at 70 degrees away from normal incidence for 20 seconds. The resist was then removed by pressurized and heated NMP (20MPa / 80C / 1000rpm).

5. Al bond pad patterning - We then spun on 1um of LoR-5A liftoff resist in an automated track coater (SVG 8626) for subsequent liftoff, followed by 0.87um of UV210 resist in an automated resist coating system (Picotrack). We then patterned the bond pads and traces connecting the pads to the ITO layer in an ASML 5500/300 DUV stepper. We then evaporated 20nm titanium (Ti) / 300nm Al in an electron beam evaporator (CHA Solutions). We deposited the Ti at 10kV beam voltage / 90mA beam current and the Al at 10kV / 50mA. We monitored the thickness with a 6MHz crystal monitor.

Following the main fabrication process, individual 10mm x 10mm chips were coated with 2um of protective i-Line resist and then diced in an automated dicing saw (Disco DAD324). The resist was then removed with acetone, and individual chips removed for subsequent bonding onto a PCB. We bonded the Al-coated backside of each die to an IPA-cleaned gold (Au) 10mm x 10mm pad on a custom PCB with two-component silver epoxy (MG Chemicals 8331), cured at 50°C for 30 minutes. We then wirebonded the top 150um x 150um Al bond pad to a bond pad on the PCB with $50\mu m$ Al wire with a manual wirebonder (West Bond Model 7400B).

Optical Reflectometer System Design

One component of this system which was absolutely critical was the optical reflectometer. While it was possible to use a commercial FTIR system for measuring the raw reflectance data, it was not possible to use it to measure small changes in reflectance. In practice, we observed that even with very short integration times, drift in the optical source made it impossible to reliably measure changes in reflectance spectra at a certain wavelength or range of wavelengths. For this reason, we decided to design and build our own optical reflectometer. This reflectometer was designed not only to suit our measurement needs for this particular project, but also for future projects using the same kind of devices.

Drift Compensation

Since we were trying to measure very small changes in reflectance ($< 0.01\%$) this required either (1) minimizing source drift to below this level or (2) moving the measurement away from DC to a higher frequency. While in principle (1) is possible, it is much more difficult than (2). Optical source fluctuations are typically on the order of 0.1-1%, and minimizing them with closed-loop feedback required substantial effort. Performing the measurement at a higher frequency, in contrast, simply requires changing the frequency of the applied voltage to the device, which is trivial. However, this is not possible out of the box with available commercial reflectometers, so we decided to design and build our own.

Choice of Optical Source

The first step in designing our optical system was to choose an optical source. This turned out to be more subtle an exercise than expected.

Fundamentally, to maximize signal-to-noise ratio, we needed to maximize the number of photons reaching the detector in a given period of time. At a fixed wavelength, this is equivalent to maximizing the optical power.

However, our experiments were to be done in a fairly narrow wavelength range, so we only cared about the photons within our narrow frequency range. This would indicate we want to maximize the optical power per unit wavelength (spectral flux, or Φ_{lambda}).

We next realized, however, that because the devices we wanted to interrogate were fairly small (¡1mm), that the emitting area of most optical sources were physically larger than this, and so the quantity we *truly* cared about was the optical power per unit wavelength per unit area of the emitter, or spectral irradiance E_λ.

However, in lens-based optical systems (which we intended to use) only a finite solid angle from the source is usable. Sources which spread their optical power over a large output solid angle are less desirable than sources which are highly directional. For this reason, the final figure of merit we settled on for a source was the power per unit wavelength per emitter area A per emitter solid angle (called the spectral radiance, $L_{e,\Omega,\lambda}$). Emitter area A and solid angle Ω can be combined into a composite metric referred to as the geometric extent $G = \Omega * A$.

We then proceeded to search for available sources in the near-IR (900nm - 1500nm) where we could feasibly fabricate initial devices. Of the thermal sources we investigated, the best were found to be Quartz-Tungsten-Halogen (QTH) lamps, two of which were investigated - a high-power lamp and a low-power lamp. Other sources investigated included an LED and a tunable laser source (Santec 510).

Table 3.1: Available optical sources and their typical characteristics

Source Type	A	Ω (sr)	G $(mm^2 * sr)$	$\Delta\lambda_{FWHM}$	λ_{peak}
QTH (50W)[1]	1.6 x 3.3mm	4π	54	$1\mu m$	900nm
QTH (1kW)[2]	6mm x 16mm	4π	1200	$1\mu m$	900nm
IR LED [3]	0.126 mm^2	1.0	0.125	60nm	970nm
TLS	4.4 [4] * 10^{-5}	0.063	2.8 * 10^{-6}	100nm	1300nm

1 https://www.newport.com/p/6337
2 https://www.newport.com/p/6315
3 https://www.thorlabs.com/thorproduct.cfm?partnumber=M970F3 Area, solid angnle taken from output fiber, not emitter area.
4 Santec 510 Tunable Laser Source (TLS). Area from 7.5μm core diameter fiber with NA = 0.1.

Table 3.2: Available optical sources and their figures of merit, using reported values from the manufacturer and emitter areas and solid angles in table 3.1.

Source Type	$\Phi_{e,\lambda}$ (mW/nm) [1]	$E_{e,\lambda}$ [1] $(mW/nm/mm^2)$	$L_{e,\Omega,\lambda}$ $(mW/nm/mm^2/sr$ [1])
QTH (50W)	38mW/nm	7.0	0.56
QTH (1kW)	690mW/nm	2.1	0.17
IR LED	0.14mW/nm	1.07	1.08
TLS [2]	10	2.3 * 10^5	$3.6 * 10^6$

1 reported at its peak value.
2 Reported using a bandwidth of 1nm for comparison purposes (this was a typical bandwidth for other sources when doing measurements with our in-house monochromator, and so is meaningful for comparison purposes.)

As is obvious from the tables above, the tunable laser source has, by far, the highest figure of merit, in terms of its spectral radiance. However, this type of source is very expensive ($20-40k at the time of writing) and at the time we did not have one available. The next-best source was the infrared LED, which we used for characterizing modulation spectra in this work. However, the relatively narrow bandwidth of the LED made it impossible to use for

the purpose of broadband characterization. When we wanted to do this, we swapped out the LED for a thermal source. Initially, we mounted and fiber-coupled the source ourselves. However, this was time-consuming, and ended up being more difficulty than it was worth. In the end, we settled on purchasing a pre-coupled source from Thorlabs (SLS201L), whose performance was nearly identical, but in a package much easier to use.

Optical System Design

Anticipating that we may need to move the system at a future date, we elected to use a cage-based optical system. As Thorlabs has the largest selection of optical components, we used their 30mm cage system structure. As described in figure 3.5, in some experiments we placed a monochromator between the fiber-coupled source and the reflectometer.

Electronic System Design

When first designing this reflectometer, I decided it would be wise to also design the electronic stack as well, to better understand the limitations of such a system in the future. Such a system would need to be as close to the shot noise limit as possible, as I was trying to measure very small changes in that optical signal. I designed this initial system to have relatively low sampling rate (125kHz) as this was within the processing capabilities of available microcontrollers.

Maximum Achievable Photocurrent

To begin the design, I theoretically calculated the expected maximum photocurrent I would measure from an InGaAs and Si photodiode using the photodiode responsivity, monochromator grating efficiency, theoretical blackbody radiation spectrum, and specified beamsplitter reflection/transmission values. Predicted and actual values are shown in figure 3.4, (dashed lines).

The maximum photocurrent of the LED was predicted to be about 4x higher at 800nA.

ADC Choice

Per the suggestion of Analog Devices, I started my system design by choosing an appropriate ADC. To measure at the shot-noise limit with 200nA of current required an input-referred noise of $2.5pA/\sqrt{Hz}$. Typically, ADC noise is reported as total noise v_n and not spectral noise density e_n. If AC and DC measurement are to be done simultaneously, the system gain from photodiode to ADC cannot be higher than V_{DD}/I_0, which causes the ADC to saturate. In this case, taking the system bandwidth to be limited by the sampling rate of the ADC, we find that:

$$v_n < V_{DD} * \sqrt{\frac{qf_s}{I_0}} \tag{3.6}$$

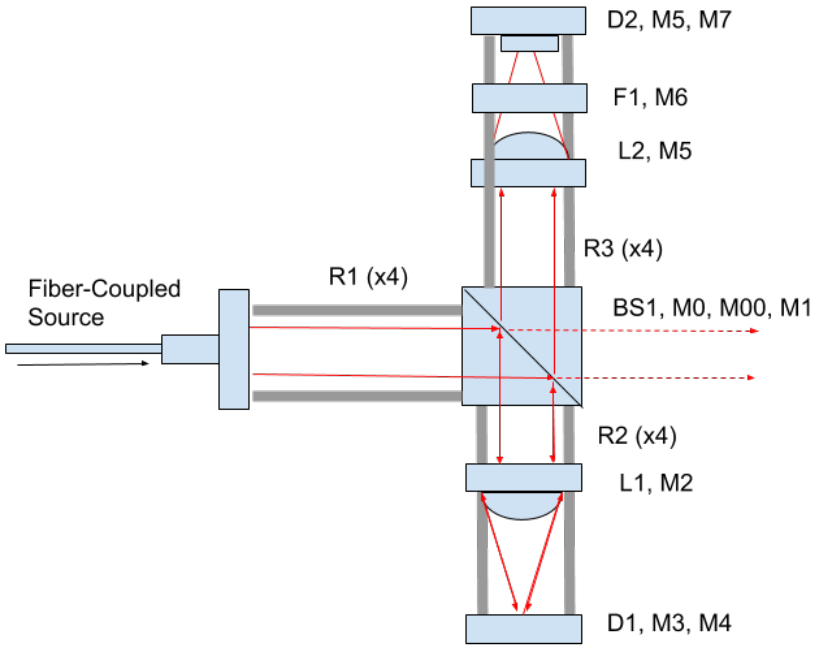

Figure 3.3: Optical cage system-based reflectometer, using fiber-coupled source described earlier. The fiber-coupled source is routed via a SMA to FC/PC fiber (M36L01) and collimated with a FiberPort adapter (PAF2A-A15C) mounted to a cage plate (CP08FP). We used 1/2" cage rods (ER05) to mount the FiberPort adapter to a beamsplitter cube (C4W) containing a plate beamsplitter (BSW22). The reflected light was then focused through a 40mm focusing lens (LA1422-C) mounted in a z-translation cage mount (SM1ZA). This light was then incident on the device, which we mounted on a 2.5mm translation mount (ST1XY-S). For additional translation along one dimension in later experiments, we added an in-line translator (CPX1) between the lens and the device. This allowed the use of a side-by-side optical reference. The reflected light then transmitted through the beamsplitter, and was either focused via a 40mm lens directly onto a photodiode (FGA10) through an optical filter, or, in later experiments, was coupled into an optical fiber for measurement with external hardware.

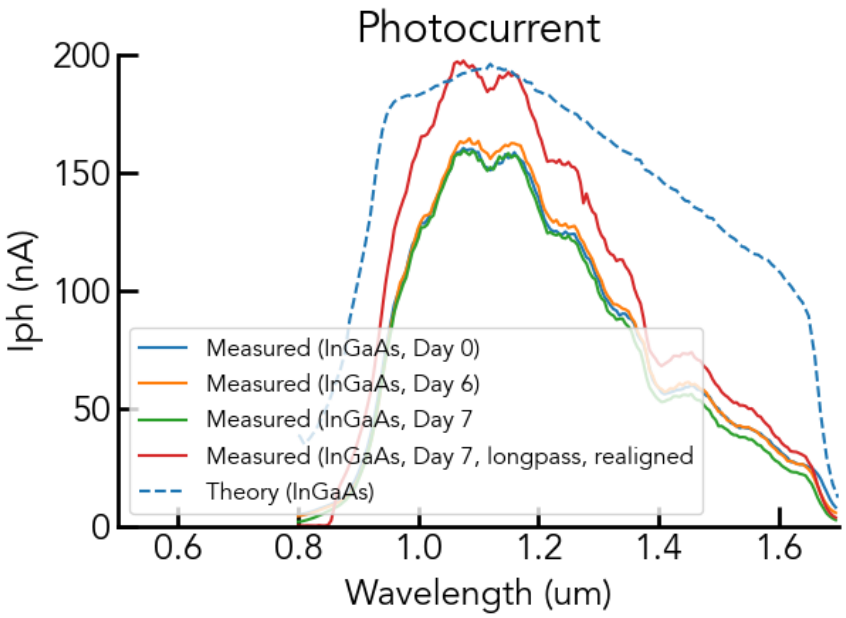

Figure 3.4: Theoretical and measured photocurrent vs. monochromator center wavelength after reflection from a thorlabs 50/50 beamsplitter, reflection from the surface of copper tape, and transmission back through that beamsplitter. Copper tape was optically rough and very scattering.

If the ADC is differential, then its maximum SNR is $(V_{DD}/v_n)^2$, and the condition for shot-noise-limited operation can be rewritten as:

$$SNR > \frac{I_0}{qf_s} \tag{3.7}$$

I chose to limit my sampling rate to 125kHz, as faster than this would require specialized hardware. At 800nA (the maximum current of the LED with the attached monochromator), it requires at least 76dB of SNR. At 1mA of photocurrent, which was higher than we were likely to achieve with the unfiltered LED, this requires 107dB of SNR. To allow this design to be used flexibly with multiple sources at the shot noise limit, I chose the highest plausible SNR I was likely to need. I chose the MCP3561 for this, as it had close to the desired SNR, and a reasonable choice of sampling rates up to a maximum of 128kHz.

Transimpedance Amplifier Design

To ensure the TIA has a noise below that of shot noise is even simpler than with the ADC, as it turns out to reduce to a single parameter - the supply voltage of the TIA. The TIA

has several regions of operation. At low input currents and high feedback resistances, the input-referred noise current is likely to be dominant. At high input current and low feedback resistances, either the input-referred noise voltage or the resistor noise current is likely to be the dominant noise source. If we assume the feedback resistor is always chosen to bring the DC current near the positive rail (which is usually possible in practice), the relevant equations simplify greatly. To ensure the input-referred noise current is always below the shot noise:

$$i_n < \sqrt{2 * q * I_{min}} \tag{3.8}$$

Where I_{min} is the minimum DC current we want to be able to measure at the shot-noise limit and i_n is the TIA's input-referred spectral noise density. To ensure the input-referred noise voltage is always below the shot noise:

$$e_n < V_{DD} * \sqrt{\frac{2 * q}{I_{max}}} \tag{3.9}$$

Where e_n is the input-referred spectral noise voltage, I_{max} is the maximum DC current we are likely to want to measure, and V_{DD} is the supply voltage. To ensure the resistor noise is always below the shot noise:

$$R > 2\frac{\phi_T}{I_0} \tag{3.10}$$

Where $\phi_T = \frac{k*T}{q}$ is the thermal voltage and I_0 is the DC current to be measured. With the initial assumptions that we made, the last of these equations always holds when $V_{DD} > 2*\phi_T$, which is always the case in practice. Assuming $I_{min} = 200nA$ and $I_{max} = 1mA$, the remaining two equations give us that $e_n < 59nV/\sqrt{Hz}$ and $i_n < 250fA/\sqrt{Hz}$, assuming a supply voltage of 3.3V. I chose the ADA4625, which has $e_n < 10nV/\sqrt{Hz}$ for all frequencies above $1Hz$, and $i_n = 4.5fA/\sqrt{Hz}$, ensuring the TIA will never be the dominant noise source in any measurement configuration.

3.4 Experimental Results

The optical voltage sensor (OVS) was characterized using the setup shown in Fig. 3.5, where we used an LED (Thorlabs, M970F3) with a peak intensity at ˜$970nm$ ($\pm 10nm$) and a mean intensity at ˜950nm. The light from the LED is fiber-coupled to collimating lens L1, beam-splitter BS1, focusing lens L2, focusing lens L3, 850nm long-pass filter F1 (Thorlabs FELH0850) and Si photodiode PD (Thorlabs, SM05PD2B). The PD (Thorlabs, SM05PD2B) was connected to a transimpedance amplifier (TIA) with a 1MΩ feedback resistance to convert a light-induced photocurrent on the PD to a voltage. The resulting voltage was digitized by an analog-to-digital converter (ADC) (NI myDAQ; National Instruments) with a 10kHz sampling rate and then sent to a computer through a serial link for data storage and further analysis.

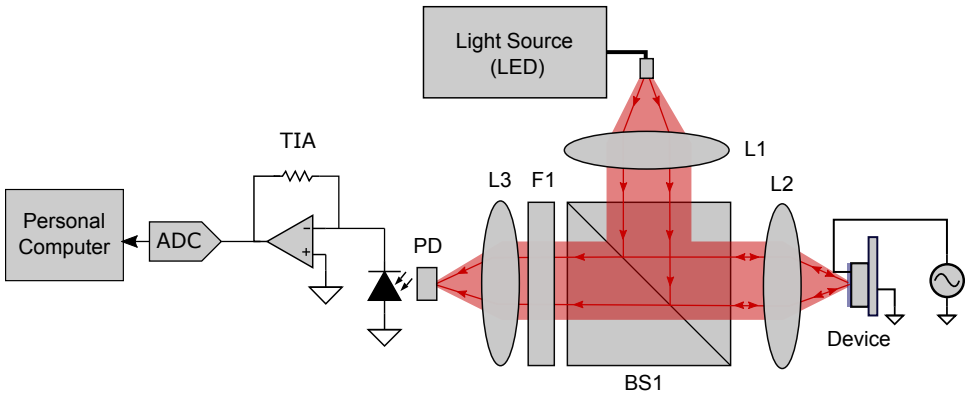

Figure 3.5: Sensor characterization setup. Light is emitted by an LED source, focused onto device, imaged onto a photodiode, and measured with a TIA. The TIA output is sent to the ADC to be recorded in the computer.

To extract the modulation depth (ΔR_{rms}) spectrum of the OVS, we applied a 105Hz sinusoidal signal (V_{in}) to the sensor using a function generator (Agilent 33210A) and used a 2.4nm bandwidth monochromator (Dynasil, DMC1-05G) with the LED. Note that the monochromator is only used in the ΔR_{rms} measurement. The collected photocurrent spectrum data was normalized to the data obtained in the same manner using a 120nm gold-coated sample. Fig. 3.6(a) depicts the measured ΔR_{rms} spectra for the sensor operated at V_{in} of $20V_{pp}$, showing good agreement with the ΔR_{rms} spectra predicted from a thin-film Fresnel equation model (supplementary materials) using the parameters provided in Supplementary Table S1. The difference in the measured and predicted spectra can be mainly attributed to the Si substrate becoming increasingly transparent at longer wavelengths, causing the transmitted light through the AlN layer (subsequently reflected by the Si/Al interface) to partially cancel the reflected light from the cavity. The R_0 spectra is given in Supplementary figure S5 (Dataset 9 [39]).

The Sensor Operated With an Incident Optical Power ($\mathbf{P_{in}}$) of 110μW Exhibits a Resolution of $170\mathbf{mV_{rms}}$ in a 5kHz Bandwidth With a Full Scale of $140V_{rms}$.

Fig. 3.6(b) depicts the output photocurrent as a function of V_{in} (105Hz sinusoidal signal) (Dataset 2, Ref. [32]). The sensor operated with P_{in} of $110\mu W$ shows a sensitivity of 2.09nA/V and a noise floor of $5.0pA/\sqrt{Hz}$, yielding a voltage resolution of $2.4mV/\sqrt{Hz}$, corresponding to $170mV_{rms}$ in a 5kHz bandwidth. Note that the measured noise floor was in excess of the shot noise limit ($2.3pA/\sqrt{Hz}$) by 6.6dB, dominated by the noise from the LED.

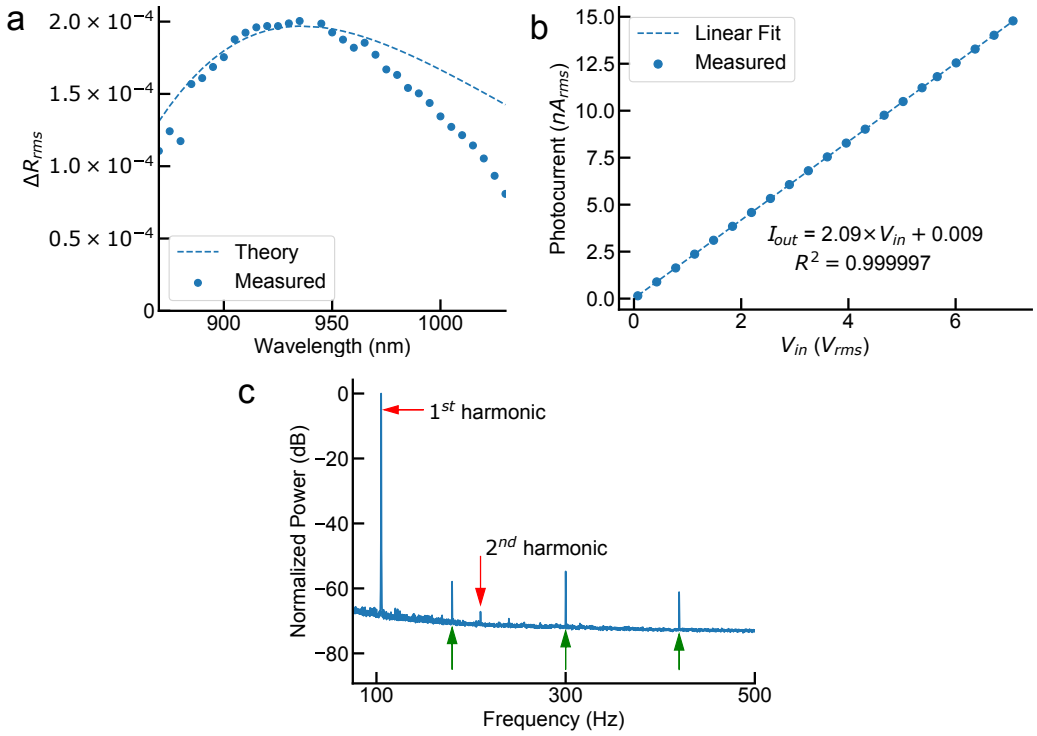

Figure 3.6: (a) Measured and predicted modulation depth spectra of the sensor operated with an input voltage (V_{in}) of $20V_{pp}$ (Dataset 1, Ref. [31]). (b) Output photocurrent versus V_{in} applied to the sensor (Dataset 2, Ref. [32]). (c) Power spectral density of the sensor output (averaged 500 times, the duration of each measurement was 10s), normalized with respect to the first harmonic amplitude (Dataset 3, Ref. [33]). Red arrows indicate applied signal, and green arrows indicate mains interference.

The Sensor Second-Harmonic Distortion Results in a First-Harmonic Measurement Inaccuracy of only 0.04%.

In the nonlinearity test, the sensor was operated with V_{in}=$20V_{pp}$ and P_{in}=˜$110\mu W$. Fig. 3.6(c) shows the normalized power spectral density of the sensor output (Dataset 3, Ref. [33]). The second harmonic power alone (-67.3dB) corresponds to a measurement inaccuracy of 0.04% in the sensor output; this result is consistent with the nonlinearity of AlN reported in [15]. The third harmonic is invisible in the spectrum as its power is below the sensor noise level. However, each subsequent harmonic is expected to fall in amplitude from the previous one by roughly an order of magnitude [15], and we expect the total harmonic distortion to be close to that of the second harmonic. The other tones seen in Fig. 3.6(c) are the 60Hz

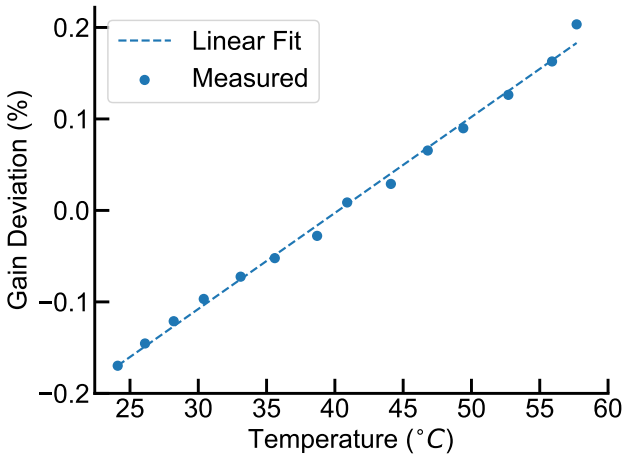

Figure 3.7: Normalized change in sensor gain ($\Delta\beta/\beta$) versus temperature. Raw data available in Dataset 4, Ref. [34].

interference tone and its numerous harmonics.

The Sensor Output Varies Only +/-0.2% Over a ˜40°C Temperature Range.

We measured the temperature dependence of the sensor gain (β, see Eq. 2). During the measurement, the sensor, operated with a $20V_{pp}$, 105Hz sinusoidal signal (V_{in}), was placed on a 250µm-thick polyimide heater (Omega) that was controlled using a PID controller (Omega, CSi32) and a K-type thermocouple attached to the sensor die. The measurement result in Fig, 3.7 showed that the sensor gain varies only +/-0.2% in the temperature range of 23-57°C, yielding a temperature sensitivity of ˜0.01%/°C. We fit the measured $(\Delta\beta/\beta)/\Delta T$ using a monochromatic Fresnel equation based optical model [56], with the incident wavelength as the fitting parameter. From this model, the incident wavelength was fit to 940nm, which is 10nm lower than the LED's mean emission wavelength, within the manufacturer's tolerance. This is about 30nm away from the optimal operation wavelength of 911nm at which point the error is predicted to be quadratic with temperature, with a max deviation of less than +/-0.02%.

Sensor Lifetime Can Meet or Exceed that of Instrument Transformers.

To determine the maximum electric field (E) that could be safely applied to the sensor along with the expected sensor lifetime, we measured the breakdown charge Q_{bd} of 8 sensors with low leakage currents of less than 0.1nA at 10V input voltage (V_{in}) [143]. During the measurement, we subjected these sensors to a linear voltage ramp to $43V$, followed by a temperature ramp ($\tilde{}5°C/min$) to $180°C$, and recorded the current over time until the point of failure. The obtained data were fit to a Weibull distribution [93], whose cumulative distribution function (CDF) is

$$1 - e^{-(Q/Q_0)^{\gamma}} \tag{3.11}$$

where γ is the Weibull slope and Q_0 is the characteristic charge. We extracted values of γ=0.67 and Q_0=379mC (Supplementary figure S2). Combining the extracted Weibull CDF (Supplementary figure S2) with the leakage current at room temperature (Supplementary figure S3), we estimated the likelihood of sensor failure over time at various operating input voltages (see Fig. 3.8). While this measurement was performed under DC conditions, this device would normally be operated under sinusoidal AC conditions. To estimate AC lifetime, we conservatively assume that the AC failure statistics are similar to the DC failure statistics with the peak AC amplitude in place of the DC voltage. This has been shown to be the case using square waves of varying frequencies in insulator systems similar to AlN [78, 57].

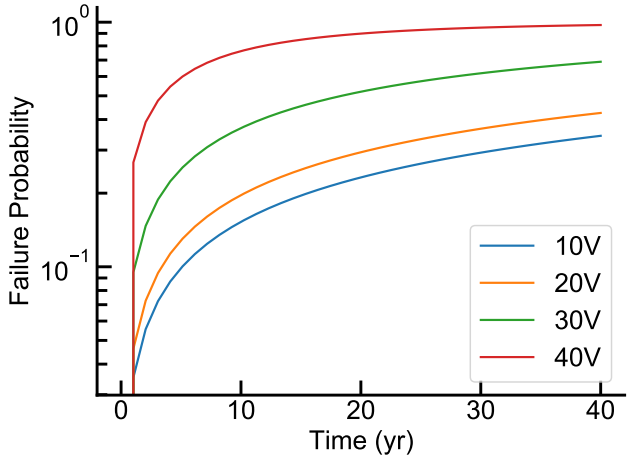

Figure 3.8: Probability of device failure versus time at various DC operating voltages. Data underlying curve available in Datasets 6-7, Ref. [36, 37]

Transient Waveform

A representative time-domain plot of the measured current as a function of time is shown in Fig. 3.9. The $40nA_{pp}$ signal is riding on top of a $20\mu A$ offset current, for a modulation depth of $\tilde{}2 * 10^{-3}$.

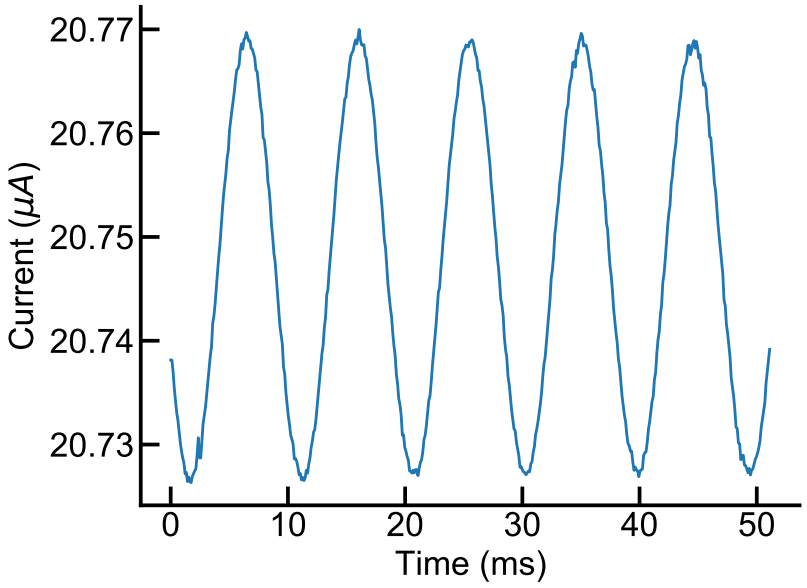

Figure 3.9: Representative time-domain transient plot with $V_{in} = 20V_{pp}$ at an input frequency of 105Hz.

Breakdown Field Indicates Device can be Operated Directly up to 90V

The overall Weibull slope γ for the breakdown field was measured to be 3.8, with a characteristic field E_0 of $3.0MV/cm$, which corresponds to an applied voltage of 90V, above which most devices start failing.

Breakdown Charge Indicates Devices can Withstand $12C/cm^2$ Before Failure

The overall Weibull slope γ for the breakdown charge was measured to be 0.67 with a characteristic field E_0 of $380mC$, which corresponds to a characteristic breakdown charge density of $12C/cm^2$.

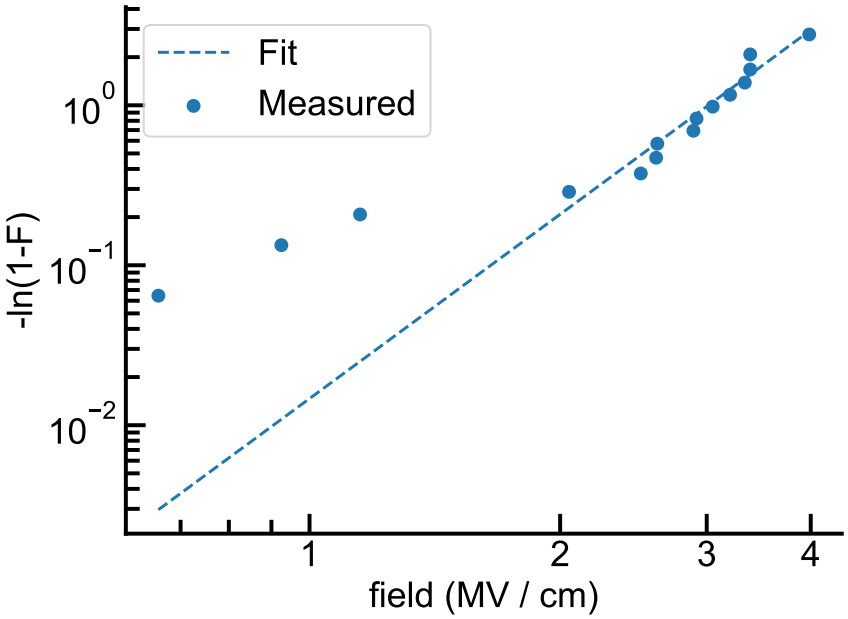

Figure 3.10: Cumulative number of failures vs. applied electric field (Weibull plot). F: CDF of the failure probability distribution. Raw data available in Dataset 5.

Leakage Current over Voltage, Temperature

The leakage current reported in the device lifetime figure was at a temperature of 20C. To ensure that these conclusions are valid at elevated temperatures, and to predict lifetime versus operating condition, we measured the IV curve of one low-leakage devices at temperatures from 20C-100C. Above approximately 20V, the leakage current increases by an order of magnitude from 20C - 100C, but only by a factor of 2 between 20C and 60C. Colder temperatures would have lower leakage currents than shown here, and we may expect the average temperature over the device lifetime to be near room temperature.

Device Impedance Yields a Device Capacitance of 1nF and Theoretical Bandwidth of 400kHz

Device impedance was measured using an Agilent 4284A LCR meter, and the measured impedance of a typical device, along with the best-fit RC model is shown in Fig. 3.13.

Device Reflectance Spectra, Model Fitting

The reflectance of a single device was measured in the range 870-1020nm, and the reflectance is shown in Fig. 3.14. This was fit to a Lorentzian function with Q and R_{max} as fitting

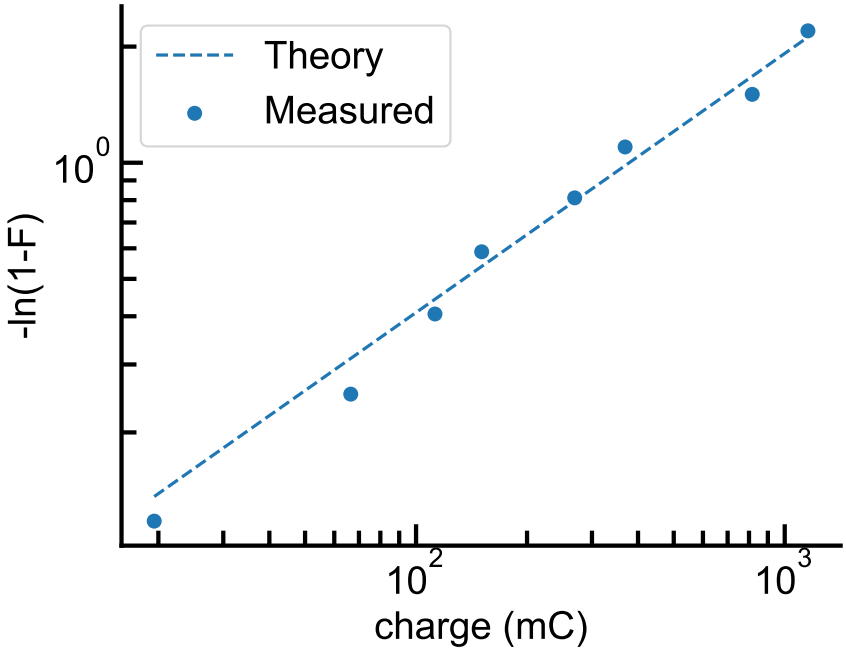

Figure 3.11: Cumulative number of failures vs. applied charge (Weibull plot). F: CDF of the failure probability distribution. Raw data available in Dataset 6.

parameters which are shown in table 3.4.

Optical Frequency Response is flat and TIA-limited

We also measured the frequency response of the optical transfer function, which is plotted, normalized, and fit to a single-pole system in Fig. 3.15, described by:

$$\beta_{norm}(f) \equiv \frac{\beta(f)}{\beta_{max}} = \frac{1}{\sqrt{1+(f/f_0)^2}} \tag{3.12}$$

Where f is the frequency of the input signal and f_0 is the pole frequency. The best-fit pole location f_0 was at 5.12kHz, which is consistent with the designed TIA pole location (5kHz) and the capacitor tolerance used ($\pm 5\%$). This shows the device frequency response is dominated by the TIA, as expected from device impedance measurements.

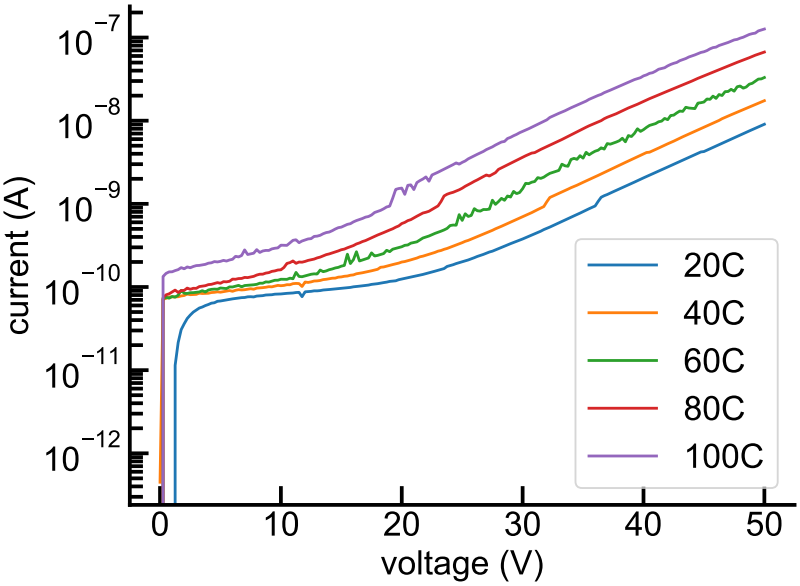

Figure 3.12: Leakage current vs. voltage at temperatures from 20C to 100C. Raw data available in Dataset 7.

3.5 Discussion

Monitoring Grid-Level Voltage up to ˜350kV_{rms} is Possible Through Capacitive Division.

Since the optical voltage sensor (OVS) is based on a 300nm AlN thin film, a large device capacitance ($> 1nF$) is possible despite the small sensor size (2mm diameter) and relatively low-permittivity dielectric (ϵ_r=9.5). We measured the sensor impedance using an HP2484A LCR meter from 10kHz-500kHz (Supplementary figure S4, Dataset 8, Ref. [38]), and extracted the device capacitance (C=0.997nF) and resistance (R=389Ω) from fitting the measured data to a series R/C model. This relatively high capacitance allows for off-the-shelf capacitors (¿˜1pF) to be used for capacitive division and hence to facilitate measuring high voltages in the order of tens to hundreds of kVs. Fabricating the device on an insulating substrate would also enable capacitive division. For example, a quartz substrate (ϵ_r=4.5 [72]) with a typical thickness of 675μm can be used to fabricate the sensor and form a capacitive divider on the same die; the quartz capacitance density ($0.06pF/mm^2$) is much lower than the AlN film capacitance density ($280pF/mm^2$) and allows approximately 5000:1 capacitive division for the same size sensor and capacitor. With a breakdown voltage of $100V$ (Dataset 5 [35]), this could enable voltage sensing up to $350kV_{rms}$, but one can choose to operate

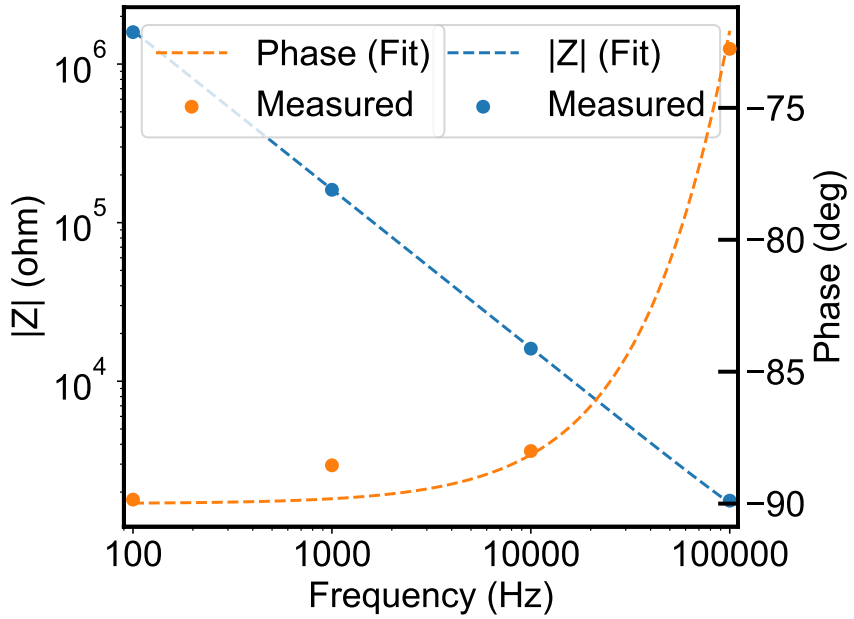

Figure 3.13: Representative device impedance vs frequency from 10Hz-500kHz. Theoretical curve was fit using an R/C series model, fitting to the logarithm of the magnitude of Z and the phase directly using the python symfit package. Fitted series resistance was 389Ohms and fitted capacitance was 997pF. Raw data in Dataset 8.

the sensor with voltages $< 350kV_{rms}$ to extend the sensor operational lifetime. In designing such a capacitive divider, the thermal variation of both the device and dividing capacitor should be carefully considered. The temperature coefficient of permittivity (TCP) for AlN is roughly $7*10^{-4}/K$ [59], so a capacitor with lower or matching thermal characteristics should be chosen to minimize errors resulting from temperature variations. Quartz, for example, which has a TCP of approximately $-4*10^{-5}/K$ [53], would cause an 0.4% error over a $60°C$ range.

The Low-Cost OVS System Can be Built Using our Device and Inexpensive Optical Components.

To our knowledge, we present the first OVS that uses an LED for operation, rather than a specialized light source such as an amplified spontaneous emission source (ASE) [48, 154] or superluminescent LED (SLED) [24] (see Table 3.3). Our low Q-factor device enables using a conventional broadband light source LED without the need of an optical filter, which allows sensor operation at higher powers near baseband with lower noise floors. This is because LEDs do not suffer from the same low-frequency excess noise that other narrowband light

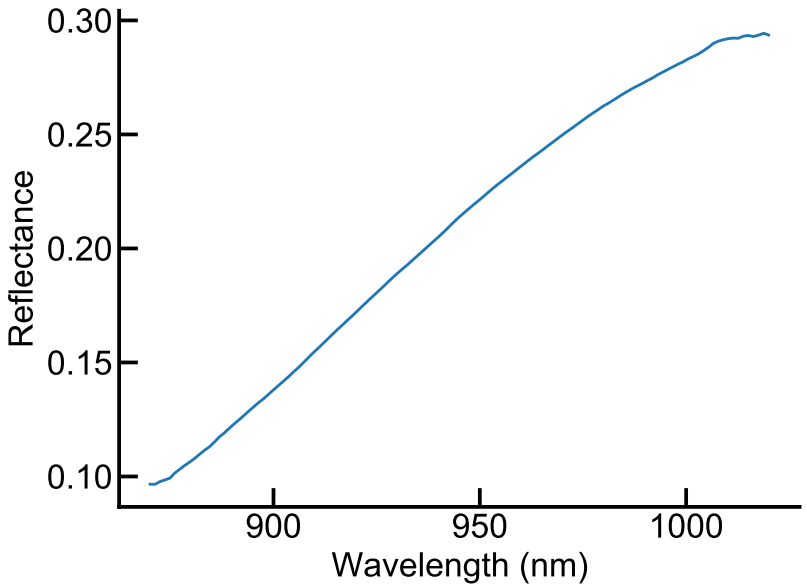

Figure 3.14: Measured reflectance spectra of optical modulator. Raw data available in Dataset 9.

sources, such as lasers, do and can get very close to the shot noise limit [123]. Additionally, the sensor can be interrogated at an angle or in a transmissive configuration; this would allow input and output fibers to be coupled directly to the sensor, without the necessity of using optical circulators or beamsplitters. This proposed architecture would eliminate all bulk optical components from the signal path; greatly reducing system cost and complexity. Future work should focus on implementation and optimization of this architecture. Since this device is a relatively large size (mm-scale) it does not require precise alignment. Taken together, these properties can enable building an optical voltage sensing system based on our sensor and inexpensive optical components (LEDs and optical fibers), which could bring the cost of production below that of a commonly used instrument transformer or prior OVS systems [48, 154, 24]. The major disadvantage of using an LED in this manner is drift of the output light power and center wavelength over time and temperature. In practice, this problem can be solved by using an additional photodiode, positioned close to the LED, in the system which enables a differential measurement of a reference light beam and hence thermal drift compensation.

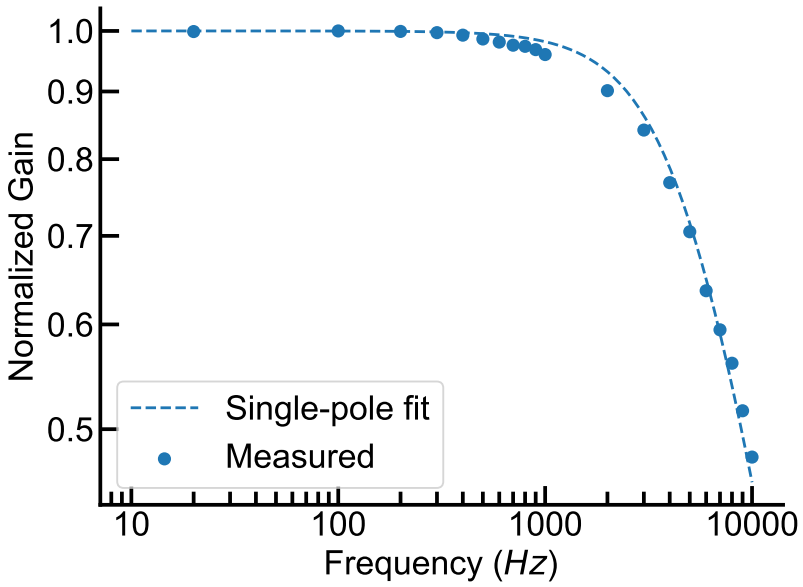

Figure 3.15: Magnitude of normalized device frequency response (optical output / electrical input), fit to single-pole model. The best-fit pole location was at 5.12kHz.

The OVS Trades off *Energy per Quanta* (E_Q) for Temperature Insensitivity and Linearity.

Our sensor represents one extreme within the spectrum of all possible OVSs by deliberately excluding mirrors to reduce the Q-factor; the low Q-factor allows for our sensor to achieve better total harmonic distortion (THD) and temperature-induced relative gain error ($\Delta\beta/\Delta T$) (without any compensation) than that of prior work, as shown in Table 3.3.

As we have shown, the noise efficiency of OVSs, represented by E_Q, can be traded off directly with nonlinearity and temperature dependence of the sensor output; E_Q can be improved at the expense of an increase in nonlinearity and temperature sensitivity by increasing the sensor's Q-factor (supplementary materials). For a sensor that is desired to be operated with a low E_Q, the sensor design can be modified to incorporate mirrors on either side of the piezoelectric AlN thin film; this will improve the sensor Q-factor and hence improve the E_Q. Alternatively, the sensor thickness can be increased to increase the Q-factor (supplementary materials) and reduce E_Q, allowing the sensor to operate at higher input voltages.

Our OVS Operates Within 6.6dB of the Shot Noise Limit with no Source Feedback.

Previous systems, as shown by their high E_Q values (see table 3.3), are operating well above the shot noise limit, wasting input photons. Our system operates within 6.6dB of the shot noise limit, with the primary excess noise due to the optical source. This efficient use of photons allows a low noise floor to be achieved despite a low modulation depth; this eliminates the need for closed-loop feedback to reduce the noise of the optical source, further reducing the cost and complexity of a potential OVS system.

Limitations of the Energy per Quanta (E_Q) Metric.

The E_Q is a useful figure of merit when trying to use optical systems as sensors rather than (digital) communication devices because it is not a direct measure of energy per information (bit) used to express the energy efficiency in digital communication systems. Here, it represents a lower limit of noise performance for a shot-noise limited optical voltage sensor (OVS) operated at a fixed bandwidth (BW) and incident light power (P_{in}) and reveals the trade-offs between sensor SNR, BW, and P_{in}. In order to design an efficient OVS, the operating BW can be traded off directly for SNR, and P_{in} can be traded for either BW or SNR.

3.6 Conclusion

This work presents an AlN thin film based optical voltage sensor for power grid applications; the sensor is fabricated using standard microfabrication techniques. We demonstrate the advantages of this sensor in terms of nonlinearity and robustness to temperature variations, and articulate a figure of merit - the *energy per quanta* (E_Q). The E_Q fully captures the trade-offs between sensor parameters, enabling the design of high-performance optical voltage sensors. Future work should focus on system-level integration and optimization.

3.7 Equation Derivations

Device Transfer Function, Gain

This device operates by a shift the resonant frequency, and the change in output reflectance ΔR can be linearly related to the input voltage V_{in} through the gain β: $(\Delta R = \beta \cdot V_{in})$. We can approximate the lineshape as a Lorentzian (this approximation is equivalent to a small-angle approximation, and gets better with increasing Q):

$$R(x) \approx R_{max}\left(1 - \frac{1}{1+x^2}\right) \tag{3.13}$$

Table 3.3: Comparison of state-of-the-art optical voltage sensors.

	[154]	[48]	[24]	**This work**
Architecture	Dual FBG	FBG	FBG	**Thin-film**
Light Source	Broadband ASE	Broadband ASE	SLED	**LED**
P_{in}	~1mW	25mW[c]	$500\mu W$[f]	**110μW**
BW	20kHz	5kHz	1kHz	**5kHz**
SNR_{max}	21dB	3dB[d]	54dB	**35dB**
E_Q	~370	$2.5\mu J$	$1.9pJ$	**13pJ**
THD	-47dB[a]	-23dB	-51dB	**-67dB**
$\frac{(\Delta\beta/\beta)}{\Delta T}$	$44\%/°C$[b]	$0.2\%/°C$[e]	$0.2\%/°C$[e]	**0.011%/°C**
Compensation	FBG	Thermal screws	Bias point tracking	**None**

[a] Estimated from V_{out}/V_{in} transfer function.
[b] Maximum differential change in gain β from $6°C$ to $8°C$.
[c] Input power not specified. Assume maximum power of typical ASE light source (500mW) with an insertion loss of 13dB, identical to our own.
[d] Estimated from PSD noise of 0.01V in 16Hz bin and signal (including spectral leakage) of 0.93V across 4 bins.
[e] Estimated from the temperature coefficient of PZT [83], which was not accounted for by the authors.
[f] Input power not specified. Assume power of typical SLED light source (10mW) with an insertion loss of 13dB, identical to our own.
FBG: Fiber Bragg Grating
$\Delta\beta/\beta$: Normalized sensor gain

Where x, the normalized frequency, is equal to $\frac{\omega_r-\omega_{in}}{FWHM/2}$, where $FWHM$ is the full-width-half-max of the spectral lineshape, ω_r is the resonant frequency at which destructive interference occurs, and ω_{in} is the input frequency. By expanding this in terms of derivatives and finite differences, we find:

$$\Delta R \approx \frac{dR}{dx} \cdot \Delta x \tag{3.14}$$

$$\frac{dR}{dx} = \frac{-2R_{max}x}{(1+x^2)^2} \tag{3.15}$$

$$\left(\frac{dR}{dx}\right)_{max} = \mp\frac{-2R_{max}3\sqrt{3}}{8} \tag{3.16}$$

$$\Delta x = \frac{\Delta\omega_r}{FWHM/2} \tag{3.17}$$

$$\Delta\omega_r \approx \omega_{r0}\left(\frac{\Delta L}{L_0} + \frac{\Delta n}{n_0}\right) \tag{3.18}$$

A change in physical length occurs due to the converse piezoelectric effect:

$$\frac{\Delta L}{L_0} = \frac{d_{33} * V_{in}}{L_0} \tag{3.19}$$

and a change in refractive index due to the Pockels effect:

$$\frac{\Delta n}{n_0} = \frac{n_0^2 \cdot r_{33} * V_{in}}{2L_0} \tag{3.20}$$

Combining these equations together at the point of maximum gain, and recognizing that $Q = \omega_{r0}/FWHM$, we find that:

$$\beta = \frac{3\sqrt{3}}{4} \frac{R_{max} Q}{L_0} \left(d_{33} + \frac{1}{2} n_0^2 r_{33} \right) \tag{3.21}$$

Q-factor Length dependence

By solving the Fresnel equations [56] for a cavity with equal mirror reflectivities, we can find an exact form of the cavity reflected intensity:

$$R = 1 - \frac{1}{1 + F * sin^2(\phi)} \tag{3.22}$$

Where F is the coefficient of finesse, which can be written in terms of the mirror reflectivity R_m:

$$F = \frac{4R_m}{(1 - R_m)^2} \tag{3.23}$$

We can solve the above equations for the full-width-half-max of the spectra, and can obtain a form of the quality factor which depends on the incident wavelength λ_{in} and the length and refractive index of the cavity:

$$Q = \frac{2\pi n L}{FWHM_\phi \lambda_{in}} \tag{3.24}$$

Where $FWHM_\phi = 2Arcsin\left(\frac{1}{\sqrt{2+F}}\right)$.

Gain Nonlinearity

To quantify the nonlinearity, we assume that the device is biased at the wavelength which corresponds to maximum linearity. This corresponds to the best-case device nonlinearity, and accounts only for the nonlinearity in the reflectance lineshape. In practice, this will be dominant as the quality factor becomes large. At very low Q-factors, other sources of nonlinearity, such as material nonlinearity, may be dominant.

Table 3.4: Parameters used in the calculation of sensor gain for analytical model and Fresnel model

Parameter	Value
r_{33}	$1pm/V$ [49]
d_{33}	$5pm/V$ [89]
L_0	$295nm^a$
n_0	2.13 [108]
R_{max}	0.39^b
Q	4.0^b

[a] Measured using interferometry.
[b] Best fit of Lorentzian to measured reflection spectra.

In terms of normalized frequency / wavelength x_0 this is $\pm\frac{1}{\sqrt{3}}$. The nonlinearity can be quantified in terms of gain compression, or normalized change in linear gain:

$$\frac{\Delta\beta}{\beta_{max}} = -\frac{9}{4}\Delta x^2 \tag{3.25}$$

where ΔH is the absolute deviation in gain and H_{max} is the gain at maximum linearity, which is also equal to the maximum gain. H_{max} is $\mp\frac{3\sqrt{3}}{8}$. We can now combine this with the energy per quanta, defined in the main text in terms of the modulation depth ΔR. Without loss of generality, we can write $\Delta R \approx G(x_0) * \Delta x$. A change in voltage will change the normalized frequency, which results in a change in reflectance due to shifting of the reflectance spectrum. Combining the terms in the energy per quanta $E_Q = \frac{2q\cdot R_0}{\Re\Delta R^2} = \frac{2q\cdot R_0}{\Re G(x_0)^2 * \Delta x^2}$ and the above expression we have:

$$E_Q * \frac{\Delta\beta}{\beta_{max}} \approx \left(\frac{2q\cdot R_0}{\Re H_{max}^2 * \Delta x^2}\right) * \left(-\frac{9}{4}\Delta x^2\right) = 4\sqrt{3}\frac{q\cdot R_0}{\Re} \tag{3.26}$$

The product of the energy per quanta and the nonlinearity is equal to a constant. In other words $E_Q = const * \frac{1}{NL}$. A better (lower) energy per quanta can only be achieved at the cost of a worse (higher) nonlinearity.

If the incident wavelength is not at the optimal wavelength, as will usually be the case, this result is worse by a factor of approximately $2\frac{x_{os}}{\Delta x}$, where x_{os} is the normalized offset frequency. Alternatively, we can say that E_Q trades off against the *square* of the nonlinearity, as quantified by the relative change in gain.

Gain Temperature Dependence

The prior section generalizes to anything which changes the resonant frequency of the device - it need not be voltage. We can thus take the results above to apply for temperature as

well - the energy per quanta trades off linearly temperature sensitivity, as quantified by the normalized change in the gain $E_Q \propto 1/\frac{\Delta G}{G_{max}}(T)$.

By expanding the relative gain change $\frac{dH}{H}$ in terms of the resonant frequency:

$$\frac{d\beta}{\beta} = -\frac{9}{4}\Delta x^2 \tag{3.27}$$

Where Δx is the shift in normalized resonant frequency away from its optimal value, dH is the absolute change in gain and H is the gain.

$$\Delta x = \frac{\Delta\omega_r}{\gamma/2} \tag{3.28}$$

Where ω_r is the resonant frequency (rad/s), and γ is the full-width-half-max linewidth of the spectra.

$$\Delta\omega_r = \omega_r \Delta T \left(\alpha_n + \alpha_L\right) \tag{3.29}$$

Where ΔT is the change in temperature, α_n is the relative change in refractive index (sometimes called the thermo-optic coefficient), and α_L is the Putting these three equations together, and recalling that $Q \equiv \omega_r/\gamma$ we find the relative change in gain to be:

$$\frac{d\beta}{\beta} = -9Q^2\Delta T^2 \left(\alpha_n + \alpha_L\right)^2 \tag{3.30}$$

For a the parameters used to model this device, we expect the change in gain over the temperature range used to be 0.05%

Chapter 4

High-Q Optical Voltage sensor for neural sensing

4.1 Abstract

Optical links for medical implants have recently been explored as an attractive option primarily because it provides a route to ultrasmall wireless implant systems. Existing devices for optical communication either are not CMOS compatible, require large bias voltages to operate, or consume substantial amounts of power. Here, we present a high-Q CMOS-compatible electro-optic modulator that enables establishing an optical data uplink to implants. The modulator acts as a pF-scale capacitor, requires no bias voltage, and operates at CMOS voltages of down to 0.5V. We believe this technology would provide a path towards the realization of millimeter (mm)- and sub-mm scale wireless implants for use in bio-sensing applications.

4.2 Introduction

It is currently axiomatic in the bioelectronics field that wireless medical implants should move to smaller dimensions (mm or even smaller sub-mm) to enable implantation with minimal tissue damage and long-term impact on surrounding tissue and to ease subsequent removal of such devices. Given this, many researchers have recently focused on developing wireless ultrasmall implant systems to revolutionize healthcare by improving their applicability in tissue for diagnostics and therapies.

Most existing implantable wireless systems use radio frequency (RF) electromagnetic (EM) waves in the GHz frequency range to transfer power to and data from the implant [13, 79]. However, coupling these EM waves through an RF coil (antenna) to sub-mm implants becomes severely inefficient [149]. This is because GHz EM waves experience high attenuation in tissue [138], and the miniaturization of the RF antennas to sub-mm scales results in a small coil impedance, making it challenging to harvest enough RF energy to

power up the implant [68]. Ultrasound waves with sub-mm wavelengths in a few MHz range have recently gained attention as they enable efficient coupling to tiny implants [112, 131]. However, the miniaturization of ultrasonic implants to sub-mm scales is still challenging as scaling down the size of the piezoelectric bulk crystal (antenna) used in ultrasonic implants degrades power transfer efficiency and data transfer reliability of the acoustic link in tissue [20, 130]. Light, on the other hand, is very attractive to use for designing wireless implants at the sub-mm or even μm scales as it has short wavelengths (sub-μm and μm scale) to enable efficient coupling and the efficient conversion of light to electrical energy through a semiconductor photovoltaic cell is possible [4, 98]. In biosensing applications, the wireless optical implants consist of a sensor and an integrated circuit (IC) incorporating interface electronics, power management, and data communication components [100, 82].

This work focuses specifically on an optical wireless data communication modality to advance the realization of low-power wireless implants with ultrasmall footprints. In practice, uplink data (data emitted by the implant) can be transmitted from the inside to the outside of the body by using SiGe optical modulators or light sources (LEDs or lasers) [18, 100, 82, 1, 23]. However, SiGe modulators require large reverse bias voltages (1-3V), restricting the circuitry that can be used to drive them. Therefore, light sources, including LEDs and lasers, have recently gained popularity for optical data transfer, mainly due to their ease of integration with CMOS IC. However, these light sources require turn-on voltages with specialized drive circuitry to operate and consume substantial amounts of power.

In this work, we present a high-Q electro-optic modulator (EOM) based on piezoelectric actuation that is fabricated using standard microfabrication techniques. In prior work, we demonstrated a low-Q optical voltage sensor based on the same operating principle as the EOM for measuring high voltages in the order of tens of volts [41]. Here, the device incorporates high-reflectance mirrors on both sides of the piezoelectric film and anti-reflection coating (ARC); this enables achieving high-Q and hence operation at small voltage amplitudes. This EOM shows great promise of establishing an optical wireless link to an implant for high-bandwidth uplink data transfer and of avoiding the drawbacks of the current optical data transfer methods mentioned above. The EOM was designed to operate at a wavelength of ˜1300nm to minimize light attenuation mainly due to scattering in tissue [30]. The device operates as a simple capacitor, eliminating the need for active power consumed by the device to initiate uplink data transfer. The device can be heterogeneously integrated with a low-power CMOS IC, required for sensing, energy harvesting, and data communication to achieve extreme implant miniaturization, Fig. 1(a), and be operated at voltages down to 0.5V, compatible with CMOS electronics, with extremely minimal drive circuitry requirements. The device size is demonstrated to be scaled down to micrometers ($< 100\mu m$).

4.3 Design and Fabrication

The Electro-Optic Modulator (EOM) Operates as a Normal-Incidence Reflective Optical Intensity Modulator Requiring a Narrowband Tunable Light Source.

Fundamentally, the EOM is a resonant cavity with a half-wavelength piezoelectric active layer sandwiched between two high-reflectance mirrors. In this work, the active layer is formed by Aluminum Nitride (AlN), and the mirrors are distributed Bragg reflectors [129]. Fig. 4.1 (b) shows the operating principle of the EOM. When light is incident on this device at its resonance wavelength (λ_r), the reflected light intensity is at a minimum. Moving away from λ_r, the intensity increases until it reaches a maximum after roughly a linewidth. By using incident light at a wavelength between these two extremes, a slight shift in the λ_r can be made to result in a change in the reflected light intensity. An applied voltage (V_{in}) across the device produces an electric field inside the active AlN layer that causes it to change its thickness, resulting in a shift in the λ_r and hence a change in the reflectance (R). For the EOM operated at a wavelength (λ_{in}) where the R slope is the steepest near the λ_r, the modulation depth ΔR can be approximated by

$$\Delta R = \beta \cdot V_{in} \tag{4.1}$$

$$\beta \approx \frac{3\sqrt{3}}{4} \frac{R_{max} \cdot Q \cdot d_{33}}{t} \tag{4.2}$$

where β is the modulation gain of the device, R_{max} is the amplitude of the resonant dip, Q and t are the device quality factor and active layer thickness, and d_{33} is the thickness mode piezoelectric coefficient. Note that Equation (2) is a simplified version of the original equation [41], which is obtained by omitting the AlN refractive index term as its contribution to the ΔR is negligible compared to that of the AlN layer thickness.

The EOM operates without the need of a bias current or voltage, and acts as a simple capacitor. Electrically, the device contains no active components, does not operate using charged carriers, and is well-modeled as a single capacitor. Any electric field applied within the inner piezoelectric layer causes it to deform, but this requires no bias current, other than leakage through the layer.

The Device was Designed to Operate at a Wavelength of 1310nm that is a Minimally-Scattering Wavelength for Transmission Through Biological Tissue. A central problem for optical communication through tissue is overcoming absorption and scattering in tissue [60]. At increasingly longer wavelengths, Mie and Rayleigh scattering becomes less problematic, but at wavelengths above ˜$2\mu m$, absorption by water begins to dominate the optical loss [62]. Along with the availability of efficient optical detectors in this region [120], this has led researchers to conclude that the $1.0\mu m$ to $1.7\mu m$ near infrared region is an optimal one for imaging in biological tissue [30]. Fortunately, the near IR is also a minimally-absorbing wavelength in glass used in optical fibers [67], and optical telecommunications components are readily available in this region. Operation at wavelengths ¿$1\mu m$ also minimizes the absorption of several materials commonly used in microfabrication (among them SiN, SiO_2, $\alpha - Si$). For these reasons, we chose a target wavelength of 1310nm.

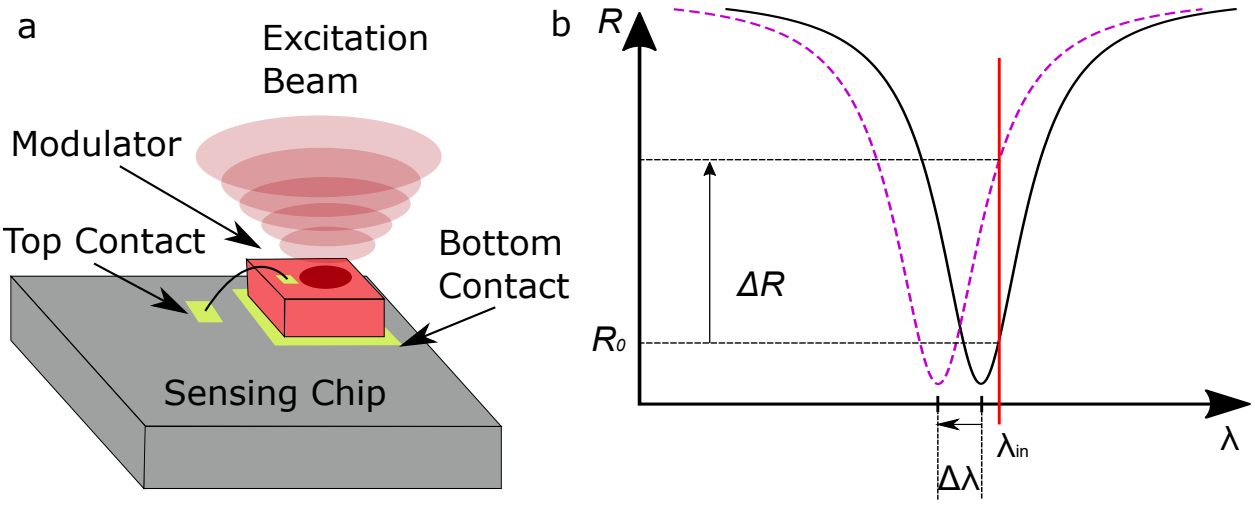

Figure 4.1: (a) Example integrated sensing system incorporating separately-fabricated optical modulator. Excitation beam is coupled into a resonant modulator. The intensity of this beam reflected from the modulator surface can be modulated by the voltage provided by the sensing chip applied between the top and bottom contact. (b) The device operates on the principle of light intensity modulation, wherein light at a wavelength λ_{in} near the resonance dip (that is, the steepest slope of the reflectance R curve) is incident on the device surface. A voltage applied across the device results in a shift in its resonant wavelength (dashed line), enabling to modulate the amount of light reflected off the device.

The Electro-Optic Modulator (EOM) was Fabricated Using CMOS-Compatible Micromachining Techniques. The EOMs were fabricated in the Marvell nanofabrication facility at the University of California, Berkeley. Fig. 4.2(a,b) shows the fabrication steps and the cross-sectional SEM of the device. Briefly, we deposited an anti-reflection (AR) coating of 50nm titanium (Ti) and 200nm amorphous Si ($\alpha - Si$) on a bare p-Si wafer. We then deposited 9 quarter-wavelength layers of alternating 240nm silicon dioxide (SiO_2) and 90nm $\alpha - Si$) using PECVD at 350°C. We then sputtered 310nm-thick aluminum nitride (AlN) layer, and deposited a second set of 9 alternating SiO_2 and $\alpha - Si$ layers for the top mirror. We then lithographically patterned circular regions with sizes between 5μm and 1280μm and etched the top mirror and the AlN layer. We confirmed full clearing of the AlN layer with an AlN etchant (dilute TMAH) drop test. Next, we sputtered 10nm titanium (Ti)/100nm Al on the backside of our wafer to compensate for film stress and provide an ohmic contact to the p-Si. After this, we sputtered a 60nm-thick layer of indium tin oxide (ITO) on top of the entire structure, lithographically patterned it, and ion milled all but a circle overlapping the original devices. Finally, we patterned a final liftoff layer and evaporated 300nm-thick Al on the surface of the wafer to form bond pads. The fabricated 5×5mm^2 die containing different size devices and test structures, shown in Fig. 4.2(c), was bonded with conductive silver

epoxy and wire-bonded to a printed circuit board (PCB) for testing. Note that multiple devices were placed on the same die to enable fair performance comparison of different size devices. See Supplementary Section 1 [MAY CHANGE] for further details of the fabrication steps.

Fabrication of High-Quality and Extremely Smooth SiO_2 and $\alpha - Si$ Layers Allows the Entire Structure to be Deposited at Once. A major challenge with deposition of Bragg mirrors is that as the number of layers increases, the mirror roughness tends to increase [22]. This degrades the reflectivity (R) of the mirror, lowering the device quality factor (Q). To avoid this, we performed a 17-run definitive screening experiment for SiO_2 films deposited on bare Si substrates and for $\alpha - Si$ layers deposited on thermally oxidized substrates. We varied the temperature, pressure, and available gas flows and measured the deposition rate with cross-sectional SEM, film roughness using AFM, and stress using wafer curvature (see Supplementary Materials). Both optimal SiO_2 and $\alpha - Si$ recipes achieved rms roughnesses of $< 1nm$ and film stresses of ˜200MPa. Fully-deposited 9-layer mirrors maintained rms roughnesses of 0.8nm, and the roughness did not increase substantially with an increasing layer number.

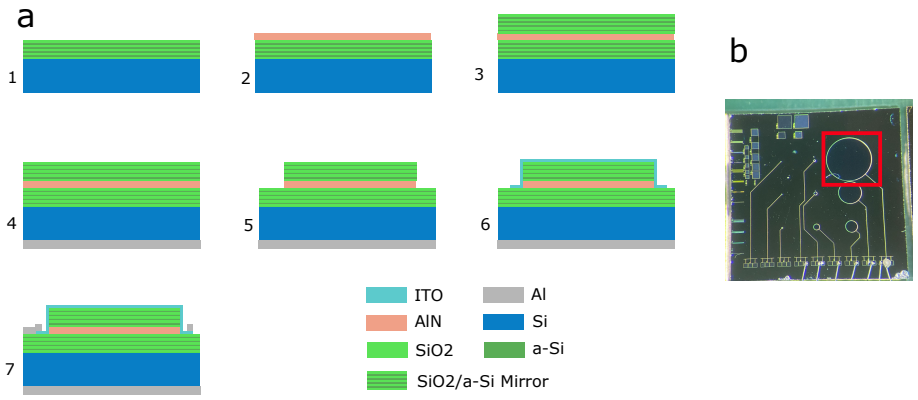

Figure 4.2: (a) Device fabrication process. 1. Deposited backside AR-coating and mirror. 2. Deposited AlN layer and frontside mirror. 3. Deposited backside W/Al metal contact. 4. Etched device mesas, stopping on back mirror. 5. Sputtered, patterned top ITO contact. 6. Evaporated, patterned Al bond pad layer using a liftoff process. (b) Cross-sectional SEM of the fabricated device. The measured thicknesses of the layers are as follows: ITO: 77nm, SiO_2: 212nm (mean, top), 207nm (mean, bottom), $\alpha - Si$: 109nm (mean, top) and 103nm (mean, bottom), AlN: 314nm., α-Si ARC layer: 229nm, Ti ARC layer: 93nm (c) Fully-fabricated devices on a 5×5mm^2 die.

4.4 Experiment Results

We Measured Reflectance (R) Spectra of the Electro-Optic Modulator (EOM) Using a Beamsplitter-based Cage System Reflectometer. Fig. 4.3(a) shows a schematic diagram of the setup used to measure the EOM reflectivity. In this setup we used a fiber-coupled thermal broadband light source (Thorlabs SDS201L). The light was collimated with a 15mm fixed-focus collimation package L1, reflected with a 50/50 plate beamsplitter BS1, and focused onto the device with a 40mm plano-convex lens L2. The reflected light from the device was then coupled through a focusing lens L3 into a $100\mu m$ core / 0.1 NA fiber, which was measured by an optical spectrum analyzer (OSA; Yokogawa AQ6370C). To obtain a R spectra of the EOM, the measured spectrum data was normalized to the data collected in the same manner using a 100 nm gold-coated reference sample; the sample has approximately unity reflectance above 800nm [12]. Fig. 4.3(b) shows the measured and predicted R spectra of the EOM with anti-reflection coating (ARC), showing that the experiment data agree well with the data predicted from a transfer matrix model (Supplementary Materials) using the device parameters provided in Table S5 [MAY CHANGE]. We believe the difference in the measured and predicted spectra mainly results from difficult-to-predict variations in the oxide absorption coefficient throughout the layer stack.

The AR-Coated Devices Achieved a Maximum Quality Factor (Q) of 2900, but Optical Coupling Reduces Modulation Gain (β) at Smaller Device Sizes. We measured the Q of each device by sweeping the tunable laser source (Santec 550) near resonance, and fitting the resultant measured spectra to a Lorentzian at a fixed Q using least-squares linear regression. A typical fit curve is shown in Fig. S6 [MAY CHANGE]. The Q measured for the five different device sizes tested are shown in Fig. 4.4(a), and range from ˜2400 to ˜2900. These are representative of devices measured elsewhere on the same wafer and across wafers, with typical Q values in the range of 1000-3000. The largest device Q (˜2900) measured with the tunable laser was almost three times that (˜1000) measured with the optical spectrum analyzer. We suspect this is due to the poor M^2 of the fiber-coupled thermal source compared to the tunable laser, and the resultant much wider range of incident angles present when using the thermal source, leading to a smeared-out spectra near resonance.

While the Q-factor for smaller devices remained high, the device β decreased substantially from $2.3 * 10^{-3}V^{-1}$ to $9 * 10^{-5}V^{-1}$ (Fig.4.4(b)). We believe this is primarily due to the poor optical coupling from the beam to the device, due to the relatively high reflectance observed when aligning to the device.

The Anti-Reflection Coating (ARC) Enables High Quality Factor (Q) by Eliminating Back Reflection of Light due to the Transparent Substrate. One challenge of operating the device at a wavelength of around 1300nm is that Si substrate is transparent in this region. This is because the incident light near resonance that passes through the device layers will be reflected off the bottom electrode surface, degrading the device's Q and therefore capability of modulating light intensity. We alleviated this problem by adding

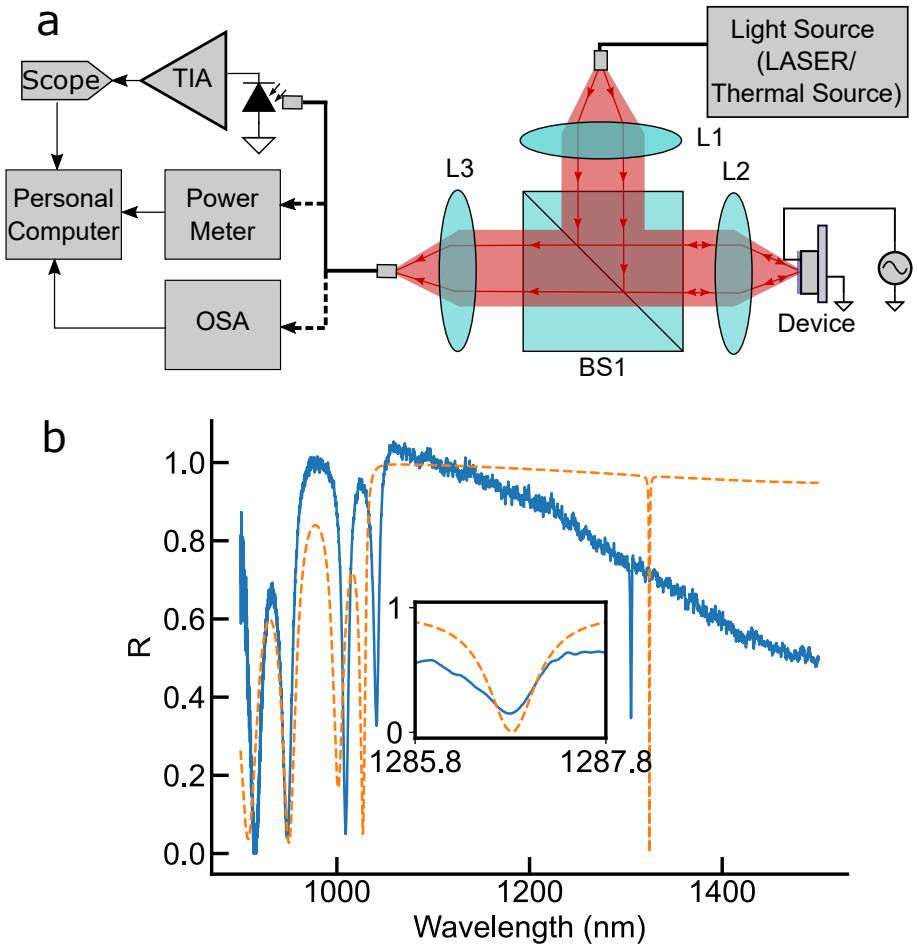

Figure 4.3: (a) Device characterization setup. Light is coupled in either from the tunable laser source or the broadband thermal source into the reflectometer and collimated by lens L1 (f=15mm). The beam then reflects off beamsplitter BS1, is then focused and de-magnified (M=0.38) by lens L2(f=40mm) onto a device. It then reflects back by lens L3 (f=15mm) and is coupled into a second fiber cable. That cable is then either routed to a TIA / oscilloscope (for narrowband/modulation experiments), a power meter (for measuring reflected power / attenuation) or an OSA (for broadband reflectance). (b) Reflectance of planar film with ARC measured using a thermal source and an OSA. Dashed line represents simulation results using TMM package (see Supplementary Materials). Inset near resonance, theoretical inset is offset to force overlap of resonances. Inset data measured using swept-source narrowband laser for more accurate Q-factor estimation. Best-fit Q is 2200 theoretical (with $\kappa_{SiO2} = 0.0002$) and 2100 measured.

.

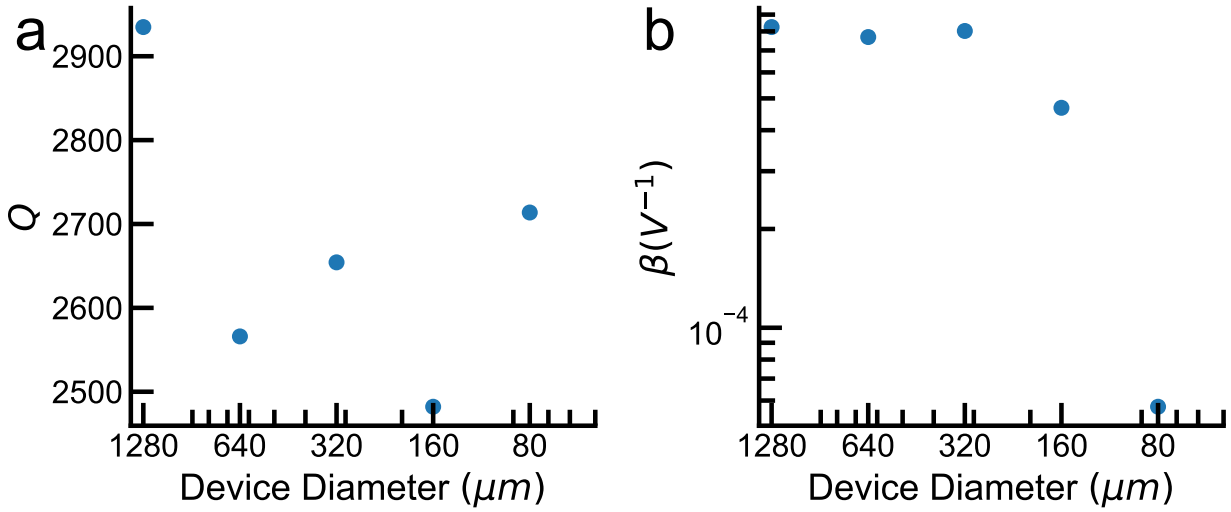

Figure 4.4: (a) Measure quality factor (Q) values for devices with differently-sized diameters, from $80\mu m$ to $1280\mu m$. Q was measured by sweeping the tunable laser source near resonance, and fitting a Lorentzian lineshape to the resulting spectra using least-squares linear regression. (b) Measured modulation gain (β) for differently-sized devices.

an ARC based on a $\alpha - Si$ ($n \approx 3.5$) layer on top of a Ti layer underneath the bottom mirror; these materials have high compatibility with standard fabrication tools. Any light near the device resonance which makes it through the bottom mirror is absorbed by the ARC, eliminating the effect of the substrate on the device performance. Fig. S8 [MAY CHANGE] shows the measured and predicted reflectance (R) spectra of the device without ARC, revealing that the device without ARC has a substantially lower Q (440) than that (2900) of the device with ARC as expected.

To Extract the Modulation Depth (ΔR) of the Electro-optic Modulator (EOM), we Utilized the same Cage System with a Fiber-Coupled Tunable Laser Source. As this device has a high-Q, the optical source linewidth must be much smaller than the resonant peak linewidth for light intensity amplitude modulation to function properly. The only light source capable of providing this at sufficient power is the laser. Therefore, we used a tunable O-band laser (Santec 550) as a light source, coupled to a $8\mu m$ core diameter fiber. We measured the change in device photocurrent using an integrated fiber-coupled photodiode and a transimpedance amplifier (TIA) (Thorlabs PDB450C) with a gain of 10^5V/A and a bandwidth of 4MHz. The TIA output was digitized using an oscilloscope (Keysight DSOX2024A) with a 5MHz sampling rate and then sent to a computer for data storage. Prior to ΔR measurements, we swept the laser wavelength around the resonant wavelength to precisely identify the operating wavelength. To maximize the ΔR during operation, the laser wavelength was set slightly off-resonance where the resonant slope was the steepest. In

the measurement, we applied a 100Hz sinusoidal signal with various peak amplitudes from 10mV to 10V to the EOM and measured the change in optical power reflected off the device. The measured power value at 100Hz was normalized to the DC level to extract the ΔR. Fig. 4.6(a) depicts the measured ΔR versus voltage amplitude applied to the EOM with a diameter of $1280\mu m$, showing that the device is highly linear ($R^2 = 0.9999$) in the typical CMOS voltage range of 0.5-5V and maintains linearity even well above the typical maximum voltage tolerated by CMOS electronics. The measured device gain β of $2.3 * 10^{-3}V^{-1}$ is in good agreement with our predicted analytical model value of $1.8 * 10^{-3}V^{-1}$ (See Table S6).

The Input-Referred Voltage Noise is $280\mu V/\sqrt{Hz}$ with 1mW of Incident Optical Power and is Nearly Shot-Noise Limited In this work, there are three main sources of noise: shot noise, excess laser noise, and TIA noise. The DC photocurrent was typically around $4\mu A$, yielding a theoretical shot noise of $1.1pA/\sqrt{Hz}$. The TIA used had a specified typical noise of $1.45pA/\sqrt{Hz}$ of $10^5\Omega$. The total measured noise above the $1/f$ corner of the laser of 10kHz was $1.46pA/\sqrt{Hz}$, about 3dB above the shot noise limit, and consistent with the two aformentioned noise sources.

The integrated noise current from 2Hz to 10kHz was $4.8nA$, corresponding to a laser relative intensity noise (RIN) of 0.1%. The input-referred noise spectrum from 2Hz - 500kHz with a $10mV$ input amplitude sinewave is shown in Fig. 4.5.

We computed the input-referred spectral noise by dividing the measured photocurrent by $i_{max} * \beta$, where i_{max} is the photocurrent due to a perfectly reflective surface ($R = 1$), and computing the power spectral density.

The Electro-Optic Modulator (EOM) can Enable Uplink Data Communication with Implants Through Digital Amplitude Modulation of the Reflected Light from the Device with a a Bit Error Rate (BER) of $< 10^{-5}$ in a Bandwidth of 4MHz. To test the device's ability to transmit digital data, we used the same setup as was used in the ΔR experiment and operated the $1280\mu m$ diameter EOM with square-wave voltages with a peak-to-peak amplitude of 0.5V and frequencies of 9kHz and 100kHz, generated by a function generator (Agilent 33210A). To reduce the effect of laser low-frequency noise, such as 1/f noise, on the output, we post-filtered the data from the device at 1kHz through a 3rd-order butterworth high-pass filter. Fig. 4.6(b) shows the time-domain transient optical response of the EOM operated with a 9kHz square-wave signal at a 1MHz sampling frequency. For the BER measurement, we collected 500,000 data points from the device operated with a 100kHz square-wave signal at a 5MHz sampling frequency and excluded the first and last five time constants (5000 points) of the data due to filtering. We measured the signal-to-noise ratio (SNR) by summing the power of the first harmonic and all subsequent odd harmonics, taking this to be the signal power, and the remaining power to be the noise power. We took each individual sample of the data to be a "bit", excluding the samples during a high-to-low or low-to-high transition, as well as the adjacent samples. To experimentally calculate the BER as a function of SNR, the laser optical power was varied between 0.3mW and 10mW during the measurement. Fig. 4.6(c) shows the measured BER versus SNR, showing perfect agreement with the BER predicted using AWGN channel model [CITATION]. It is important to know that the BER performance of the device can be

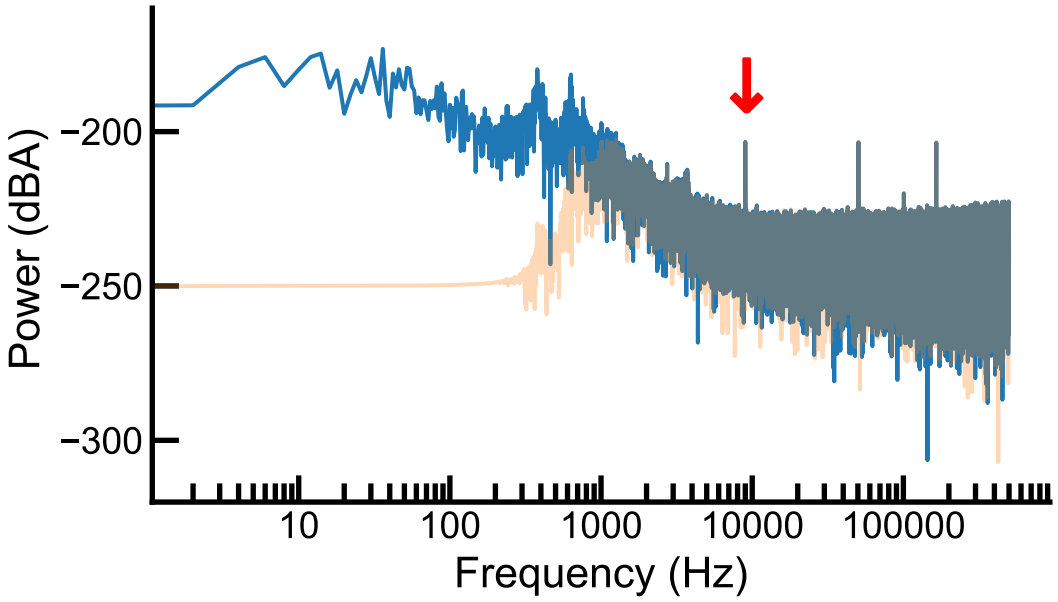

Figure 4.5: Measured input-referred power spectral density with an input 10mV amplitude sinewave to the device. High-frequency ($> 10kHz$) PSD was 153 $\mu V/\sqrt{Hz}$

improved when the device is operated with larger modulation voltages, as applying higher voltages increases the device ΔR.

As expected, SNR increased roughly linearly with output power, from -5dB to +11dB when incresing the power from 0.3mW to 10mW. Most experiments done here, when highpass filtered at 1kHz, were within 3dB of the shot noise limit. At frequencies ¡10kHz, the laser drift and 1/f noise becomes dominant (see Fig. 4.5).

The Device Operated Through 0.8mm-thick Dry Chicken Skin Demonstrates Robust Operation. As a proof-of-principle to determine whether this device could be implanted subcutaneously, we measured the ΔR of the $1280\mu m$ diameter EOM as a function of applied modulation voltage by placing it below the 0.8mm-thick dried chicken skin layer. For this experiment, we set the laser power to 1mW and used a measurement frequency of 9kHz. The measurement result is depicted in Fig. 4.6(d), showing that the device exhibited a 30 times lower ΔR than the device operated in free space. The ΔR reduction is mainly due to tissue light scattering and results in a concomitant reduction in the BER. Furthermore, the reflected light power from the device coupled into the output fiber was approximately 10 times lower than without the skin present due to the light attenuation in tissue.

The $1280\mu m$ Diameter Device can operate with a 5MHz bandwidth (BW), which increases to ˜300MHz for the smallest $80\mu m$ device. The underlying mechanism by which this device operates is a mixture of the piezoelectric and electro-optic effects. The speed of piezoelectric actuation is ultimately limited by the speed at which the material

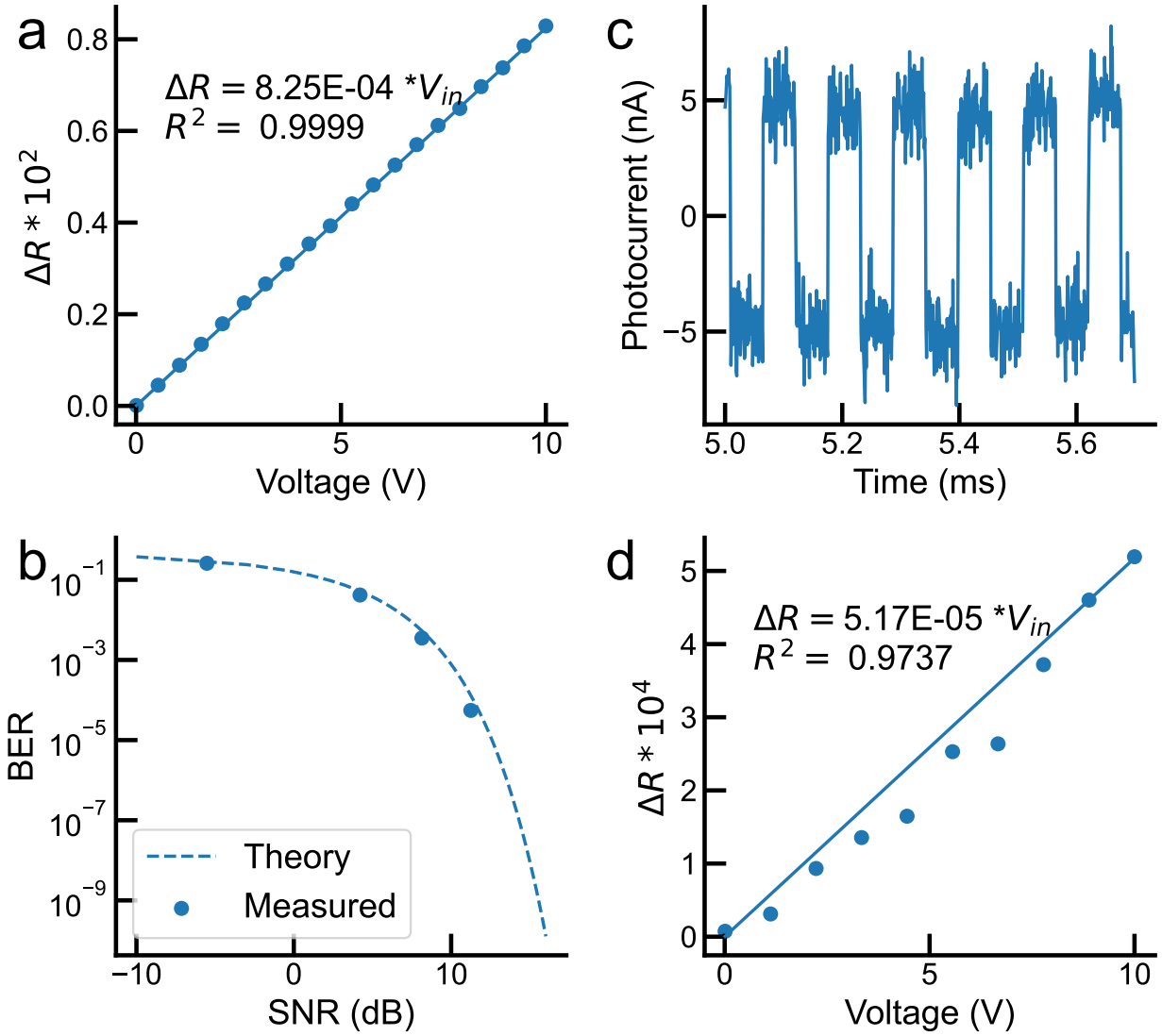

Figure 4.6: (a) Measured modulation depth (ΔR) versus input voltage in free space. (b) 1kHz highpass-filtered time-domain data from an $0.5V_{pp}$, 9kHz square wave applied to device. (c) BER versus SNR, obtained by varying incident optical power from 0.3mW to 10mW. The data was collected at 5MHz sampling rate with a 4MHz TIA bandwidth and a 50% duty-cycle, 100 kHz square-wave input modulation signal. The BER was predicted using an AWGN channel model, and calculated treating all individual sample points as bits, excluding the square wave HIGH-¿LOW and LOW-¿HIGH transition points. (d) ΔR versus input voltage measured through 0.8mm-thick chicken skin resting on the device. The device used for these measurements has a $1280\mu m$ diameter.

can deform - the speed of sound. For these devices operating in the longitudinal mode along the thickness direction, which have an active thickness of 300nm, this yields a theoretical BW of ˜30GHz using a speed of sound of $10^4 m/s$ in AlN [45]. For this reason, the device operating BW in this work is set by its electrical BW, or its RC time constant. We extracted the device series resistance R and capacitance C by fitting the lumped element RC model to the measured impedance data, shown in Fig. 4.7(a) for the $1280\mu m$ device. As we used a thin layer of relatively low-quality ITO, our device series resistance was obtained to be ˜1kΩ for all device sizes. Device capacitance values from fitting vary from ˜28pF for the largest $1280\mu m$ device to ˜0.5pF for the smallest $80\mu m$ device, after correcting for bond pad capacitance. This makes the theoretical BW range from ˜6MHz for the largest device to ˜300MHz for the smallest device. Fig. 4.7(b) shows the measured optical frequency response of the $1280\mu m$ device that matches well with the theoretical BW predicted using the R and C values from the fitting. The fitted high-frequency rolloff fitted to 8MHz was close to that expected (4MHz from the limiting TIA or 5MHz from the device), we also found an unexpected low-frequency rolloff, which fitted to a pole-zero pair at 9Hz. This may be due to leakage in the AlN layer, and would be consistent with a leakage resistance of 600MΩ. This is substantially higher (˜6x) than measured, but also higher than the values our impedance meter could reliably measure, and corresponds to a resistivity of $2.6*10^{11}\Omega\cdot$cm, in the typical AlN resistivity range of 10^{11} - $10^{13}\Omega\cdot$cm [16].

4.5 Conclusion and discussion

This work presents a high-Q electro-optic modulator (EOM), formed by a piezoelectric AlN thin film sandwiched between two high-reflectance Bragg mirrors, for wireless medical implants. The EOM is fabricated using standard fabrication processes with the highest temperature of 350°C, making it CMOS compatible. The device operates at a wavelength of ˜1300 nm in the NIR window with low tissue absorption, enabling to establish an efficient optical wireless link to implants for high-bandwidth uplink data transmission. The device operation is successfully demonstrated ex vivo in chicken skin. We also demonstrated that the device is scalable down to micrometer in size ($< 100\mu m$) and capable of operating at CMOS voltages down to 0.5V. Although there is extensive further work required to demonstrate the integration and operation of the EOM with CMOS electronics and to fully address biocompatibility and other challenges related to each targeted application, this work provides a path toward the realization of minimally invasive wireless implants for use in diagnostics and treatments.

Future versions of the device could trade off bandwidth (BW) for device modulation gain (β). In this work, we placed the device electrodes on the top of the Bragg mirror and the bottom of the Si substrate. This enables relatively easy fabrication, and reduces the total device capacitance, increasing the BW. However, the resulting capacitive division mainly between the mirrors with the capacitance of 25pF/mm^2 and the active layer with the capacitance of 240pF/mm^2 (See table S5 and S6) [MAY CHANGE]) results in a smaller voltage acting on the active layer than the voltage applied to the device, reducing

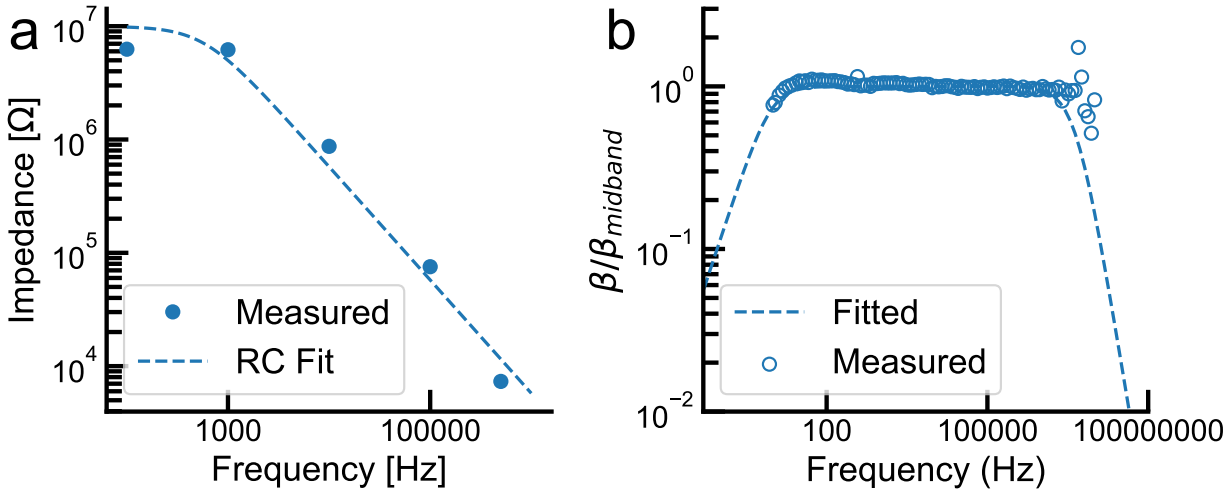

Figure 4.7: (a) Device impedance of representative 1280μm diameter device. Dashed line represents lumped element parallel RC model, with a fitted R value of 1kΩ and C value of 28pF (zero at ˜5MHz). Theoretical C value from literature permittivities and SEM thicknesses is 32pF. (b) Device normalized gain versus frequency for the 1280μm device. Fitted single-pole corner frequencies of 9Hz and 8MHz using a TIA bandwidth (BW) of 4MHz. $\beta_{midband}$ refers to gain at 10kHz. Other device sizes have virtually identical frequency responses limited by the TIA BW.

the maximum achievable β. The β can be theoretically improved by a factor of ˜9 if the electrodes are sandwiched between the mirror layers and the active layer, preventing this capacitive division (see Supplementary Materials). Since ITO is highly absorbing around the operating wavelength of 1300nm Fig. S4 [MAY CHANGE], this also requires very thin ITO layers ($< 5nm$) to form the device electrodes, which increases the device contact resistance and therefore further reduces the BW.

The choice of a different lower-absorption transparent conducting oxide to form the device top electrode could increase the device BW by two orders of magnitude. Here we used ITO - a well-characterized and readily available transparent conducting oxide - to form an electrode. However, the ITO absorption near the operating wavelength of interest (˜1300nm) is very high ($\kappa \approx 0.2$). This forced us to use a thin ITO layer, leading to a relatively low BW due to its relatively high resistivity (˜$10^{-3}\Omega * cm$). By using a different conducting oxide such as ZnO with absorption and resistivity lower than those of ITO [134], the device BW can be improved by two orders of magnitude.

Modification of the fabrication process would enable more finicky piezoelectric materials to be used. We used aluminum nitride (AlN) as the active piezoelectric material for its relatively low permittivity ($\epsilon \approx 8.5$) and ability to deposit directly on top of thermal SiO_2, as shown in this work. However, AlN has a relatively small piezoelectric constant

($d_{33} \approx 5$pm/V) [89] compared to other piezoelectric materials such as BTO ($d_{33} \approx 300$pm/V) [144]. These materials with higher d_{33} coefficients would allow for a much greater modulation depth (ΔR) at a fixed applied voltage, leading to improved SNR and hence BER, especially if the device electrodes are placed between the mirror layers and the active layer. However, they also have much higher permittivities. BTO, for example, often has $\epsilon_r > 1000$ [44], which results in a smaller device BW.

System level improvements are possible. The system's optical insertion loss from laser output to photodiode input was ˜20dB as we used fiber-coupled equipment for ease of switching between different pieces of equipment. For example, direct coupling of light reflected off the device to a collecting photodiode without using an optical fiber could reduce the insertion loss and hence increase the received optical power by ˜15dB, to as high as 17dB when the device is optically biased at a DC reflectance of 50%.

In this work, we post-processed the data from the device using a high-pass filter to reduce the effect of the laser 1/f noise on the system output. After high-pass filtering, our measured noise was typically within 1dB of the shot-noise limit. An another slight improvement can be achieved by using closed-loop noise reduction techniques [58] that could enable to minimize the laser 1/f noise and further improve the system noise performance.

Unlike traditional integrated photonic optical modulators, the electro-optic modulator (EOM) does not require a very precise alignment of the light beam. One of the most challenging aspects of existing high-Q optical modulators - which typically make use of integrated photonics - is optical coupling into and out of the device. Traditionally, this is achieved by placing an optical fiber directly above an edge or grating coupler, and requires sub-micron to single-micron accuracy, depending on the method used [80]. It is necessary to dramatically relax this alignment requirement for a modulator intended to be used to build a wireless implant because such alignment precision is extremely difficult to achieve. In contrast to the existing photonic modulators, the EOM eliminates the need for high alignment accuracy as it only requires the light beam focused to within the diameter of the device. For the largest $1280\mu m$ diameter device, the transverse misalignment tolerance is roughly ± 0.5mm with the current setup.

Device operation is sensitive to angular misalignment. The EOM operation is sensitive to incident angle (θ) between the beam direction and the device normal due to its high-Q. Any deviation in θ will cause a shift in the device resonant wavelength that can be approximately calculated using the formula: $\Delta\lambda_r = \lambda_r \cdot (1 - cos(\theta)^2)$ and a reduction in the power coupled to the device, which is also quadratic with θ [CITATION]. Therefore, angular alignment of the device to the laser beam should be carefully performed during the surgical placement of an EOM-based wireless implant in tissue.

Overall, we have shown here that substantial improvement is possible at the device level and at the system level, with the potential for two orders of magnitude increase in bandwidth, and two orders of magnitude improvement in noise performance from improved optical coupling. This, combined with the results presented thus far, present a compelling case for use of this device as a platform for wireless readout of digital information from implantable devices. While we have demonstrated microscope-free alignment of devices down to $80\mu m$

in size, further reductions in size are possible with more careful beam shaping and alignment procedures. Theoretically, devices with diameters as small as a single half-wavelength (600nm) are possible to fabricate, at the expense of increased difficulty with alignment and optical coupling.

4.6 Fabricaton Process

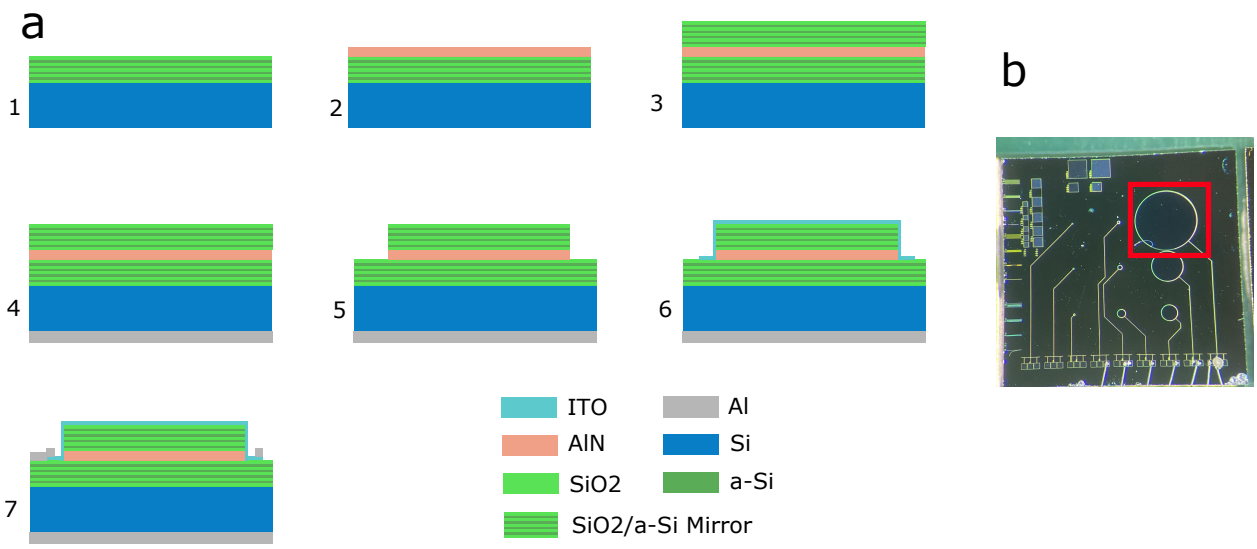

Figure 4.8: Fabrication process for high-Q optical modulators

Mirror Layer Optimization

To allow for arbitrarily thick mirrors to be deposited, I had three main concerns: first, excess stress can cause failure of films, especially $\alpha - Si$, which I have found to be particularly temperamental and prone to high stress. High stress also causes warping of the underlying substrate, which restricts the equipment that can subsequently use - particularly our lithography stepper. Second, film roughness causes degradation of mirror reflectivity, and often increases with an increasing number of layers. Third, the films should have low optical absorption near our wavelength of interest. Due to time- and resource- constraints, I did not optimize for low absorption, but primarily for low stress and roughness. If I were to redo this project, I would have optimized for all three simultaneously.

To optimize our SiO_2 film, I used the Oxford Plasmalab System100 plasma enhanced chemical vapor deposition (PECVD) tool. This particular system can provide variable silane,

Silane flow (sccm)	5 - 50
N_2O flow (sccm)	400 - 1000
Ar Flow (sccm)	0 - 200
Temperature (C)	200-350
Pressure (mtorr)	600-1500
Power (W)	15 - 50

Table 4.1: Factor ranges chosen for statistically designed experiment for SiO_2 deposition

Silane flow (sccm)	50
N_2O flow (sccm)	700
Ar Flow (sccm)	200
Temperature (C)	350
Pressure (mtorr)	600
Power (W)	15

Table 4.2: Process parameters for optimal SiO_2 deposition

N_2O, and Ar flow, and can vary the temperature, pressure, and RF power to the plasma. From previous development work, I chose parameters that I knew would sustain a plasma and could be feasibly controlled by the tool, shown in Table 4.1.

I chose to perform a 17-run definitive screening experiment, as this tool can run whole cassettes of wafers without much difficulty. This experimental design allows for quadratic effects to be estimated and provides slight confounding of two-factor interactions. I generated this experiment using JMP statistical software, ensuring run order randomization and running (2) additional center points. I deposited all films for 21 minutes on bare Si wafers, and measured the thickness using a reflectometer (NanoSpec), stress using a optical interferometer for wafer bow measurement (Tencor FLX-2320). Our films had measured thicknesses between 250nm and 2900nm, stresses between -4 and -530MPa, and roughnesses between 0.16 and 9.16nm rms. I chose a film which had both excellent roughness (0.8nm rms for a 1700nm thick film) and reasonable stress (-250MPa), which had the following process parameters (see Table 4.2):

To optimize our $\alpha - Si$ film, I used the same Oxford Plasmalab System100 plasma enhanced chemical vapor deposition (PECVD) tool used for SiO_2 deposition, and used the factors and ranges shown in Table 4.3. In this experiment, I started with thermally oxidized Si wafers, as these films were to be deposited on top of SiO_2. I also found deposition onto bare Si to yield bizarre formation of large (¿200μm) spherical mounds on the surface, which did not occur with deposition onto SiO_2.

I designed a similar 17-run experiment for $\alpha - Si$ optimization, but terminated the experiment after 9 runs as I found an acceptable film. I deposited all films for 21 minutes

Silane flow (sccm)	5 - 50
H_2 flow (sccm)	0 - 100
Ar Flow (sccm)	0 - 500
Temperature (C)	200-350
Pressure (mtorr)	600-1500
Power (W)	10 - 50

Table 4.3: Factor ranges chosen for statistically designed experiment for $\alpha - Si$ deposition

Silane flow (sccm)	5
H_2 flow (sccm)	0
Ar Flow (sccm)	500
Temperature (C)	350
Pressure (mtorr)	600
Power (W)	10

Table 4.4: Process parameters for optimal $\alpha - Si$ deposition

on bare Si wafers, and measured the thickness using a reflectometer (NanoSpec), stress using a optical interferometer for wafer bow measurement (Tencor FLX-2320). Our films had roughnesses between 0.5 and 3.4nm rms, and stress between -200 and -1900MPa, and thicknesses between 30 and 101nm. The best film of these had a stress of -200MPa and a roughness of 0.7nm rms, which had the process parameters listed in Table 4.4. Conveniently, the temperature was the same as for the SiO_2, allowing us to avoid temperature ramping between layers, a time-consuming process.

Once these mirror layers were optimized, I sent them out for optical characterization of refractive index in the range of 600nm - 1700nm (Covalent Metrology).

Step 1: Mirror Deposition

I deposited a 9-layer stack as the bottom mirror, starting and ending with SiO_2 layers. The target thickness for the SiO_2 layers was 240nm, and the target thickness for the $\alpha - Si$ was 90nm. The deposited 9-layer mirror had a total roughness of 0.8nm rms, similar to what was expected from the films themselves.

Step 2: AlN Sputtering

I used previously-developed process recipes for sputtering aluminum nitride. I used the Endeavor AT sputtering tool, which is a reactive sputter tool has a conical aluminum target and can flow N_2 gas during deposition. I used 7sccm of Ar and 22sccm of N2 with 4.5kW of

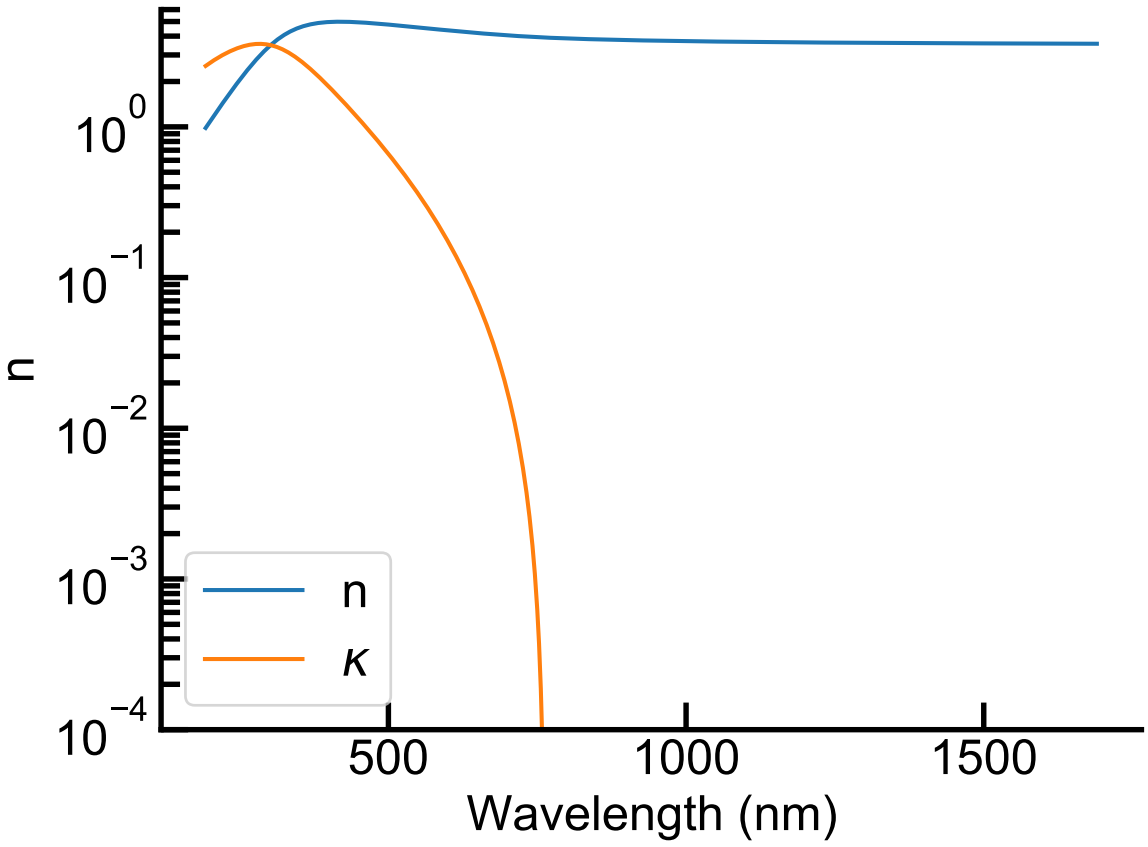

Figure 4.9: $\alpha - Si$ complex refractive index from 250nm - 1700nm, deposited on Fused Silica substrate measured by Covalent Metrology. Complex index at 1300nm is 3.61 + 0i.

DC power at a pressure of XXX, typically for 350-450s. The typical deposition voltage was 395-401V. Prior to deposition, for each wafer, I would also run a conditioning wafer which first cleaned the target for 100s with 4kW of power and 6sccm of Ar gas, followed by a brief 30s target poison using 4.5kW of power, 22sccm N2 / 7sccm Ar. As the thickness of this layer is critical to the device resonant wavelength, I ran pre- and post- thickness monitor wafers to ensure the deposition thickness was within spec, and calibrate the depositin time and potential deposition rate drift.

Step 3: Top Mirror Deposition

To deposit the ttopside mirror, I used the same process as in step 1.

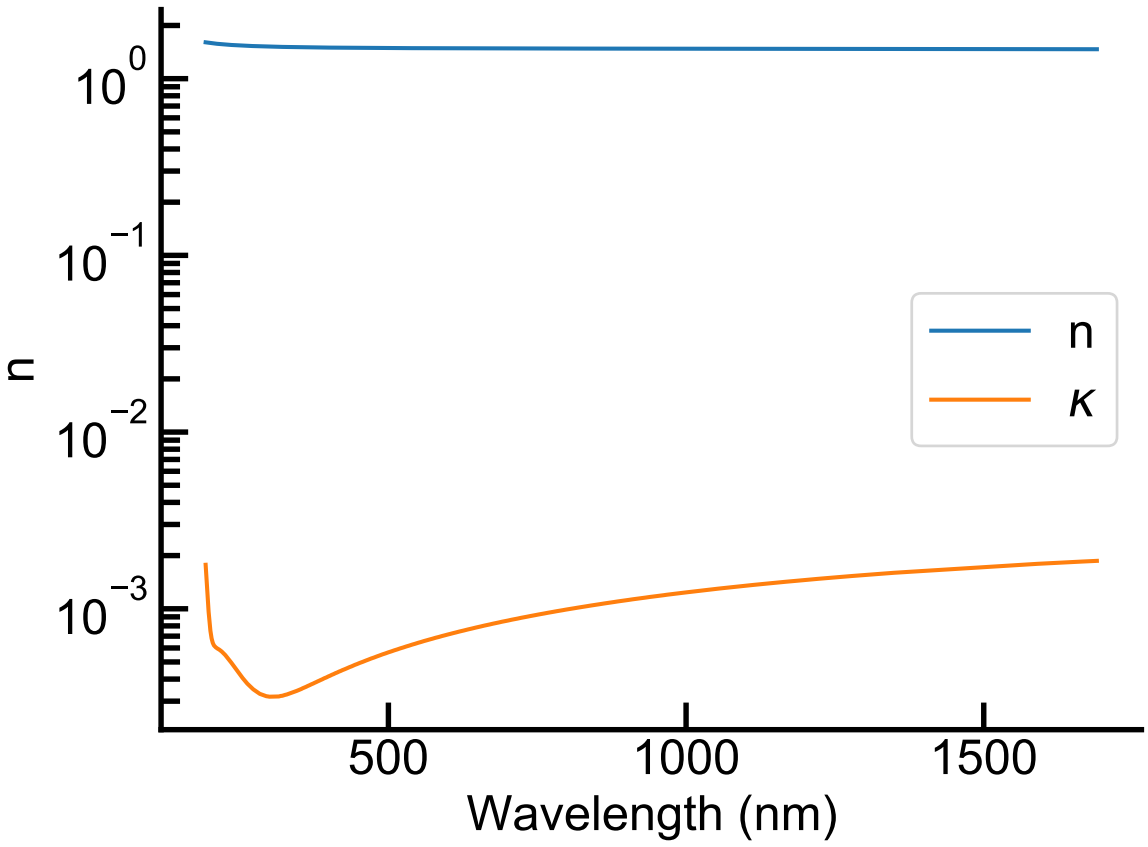

Figure 4.10: SiO2 complex refractive index from 250nm - 1700nm deposited on Fused Silica substrate, measured by Covalent Metrology. Complex index at 1300nm is 1.47 + 0.0016i.

Step 4: Al / W sputtering

To form a backside contact to p-Si, I elected to use aluminum with a post-deposition anneal of 300C in air. However, after deposition of the mirrors and AlN layers, the wafer was highly bowed, typically between 60 and 90 μm. This posed a challenge for our lithography stepper, which can only tolerate bows $< 50\mu m$. To compensate for this, after depositing 10nm of Al to form the contact, I deposited an additional 100nm of tungsten to de-bow the wafer. I chose tungsten as it can tolerate large stresses before film failure, and as-deposited was highly stressy (˜1-1.4GPa). This reduced the wafer bow to ˜$30\mu m$ to allow for subbsequent lithography.

Step 5: SiO2 / a-Si / AlN etching

To isolate individual "devices", I patterned circular mesas using 2.5 μm of UV26 negative resist, exposed at $35mJ@cm^2$ in our DUV lithography stepper (ASML DUV Stepper Model 5500/300), and developed with a 130C post-exposure bake and 45 seconds of soak time in MF-26A. I then hard baked the resist at 120C for 30 minutes.

To etch through the topside mirror, I alternated between using an inductively coupled plasma oxide etcher (STS Advanced Plasma Systems) and a transformer-coupled plasma etcher (Lam Research, TCP 9400SE). Each SiO_2 layer etch ran for 80 seconds with a strike pressure of 10mtorr, a process pressure of 4mtorr, with C_4F_8 / H_2 / He flow of 15 / 8 / 174sccm, a coil power of 1500W, and a bias power of 350W at 13.56MHz. The $\alpha - Si$ etch ran for 30 seconds with a TCP RF power of 300W, a bias power of 150W, Cl_2 / HBr flows of 50 / 150sccm at a pressure of 12mtorr. I alternated between the two different etchers until I reached the underlying AlN layer. To confirm I had breached the final oxide layer, I applied a small drop of MF-26A developer to the surface, and confirmed that it etched the AlN.

To etch through the AlN, I etched the AlN in a dedicated aluminum etch chamber (Centura Decoupled Plasma System) using a Cl_2 based etch. I used a plasma power of 800W, a bias power of 85W, a pressure of 5mtorr. I used the gases Cl_2 / H_2 / Ar / BCl_3 at flows of 50 / 50 / 70 / 5sccm for 240 seconds. I used the same MF-26A spot test to confirm that the AlN was fully cleared across the wafer.

Once the AlN was cleared, I removed the remaining photoresist using a 1kW oxygen plasma (Centura MxP+).

Step 6: ITO Sputtering, Patterning

To form the top device contact, I sputtered 60nm of ITO using a multi-target co-sputtering system. The system used a 3" ITO target with a process pressure of 3mtorr, Ar flow of 200sccm, and target power of 100W. No process optimization was attempted. I sent out a thick film of ITO for optical constant characterization (Covalent Metrology)

To pattern the ITO, I spun on 2.5μm of UV26 negative resist, exposed in our DUV lithography stepper, developed with a post-exposure bake of 130C and an MF-26A soak of 45 seconds, and hard-baked at 120C for 30 minutes. I then ion milled the ITO at 30 degrees for 5 minutes at 500V / 250mA, followed by a 70-degree polish for 1 minute. I targeted an over-etch of 50 % to ensure the ITO was fully cleared. After the ion mill, I used a probe station to confirm the layer was non-conductive, and removed the resist with a heated soak in NMP heated to 80C.

Step 7: Al evaporation, liftoff, dicing

To make contacts to the devices, I used aluminum deposited using a liftoff process. First, I deposited 1 μm of LoR-5A resist, followed by 0.87 μm of UV-210 negative resist. I patterned

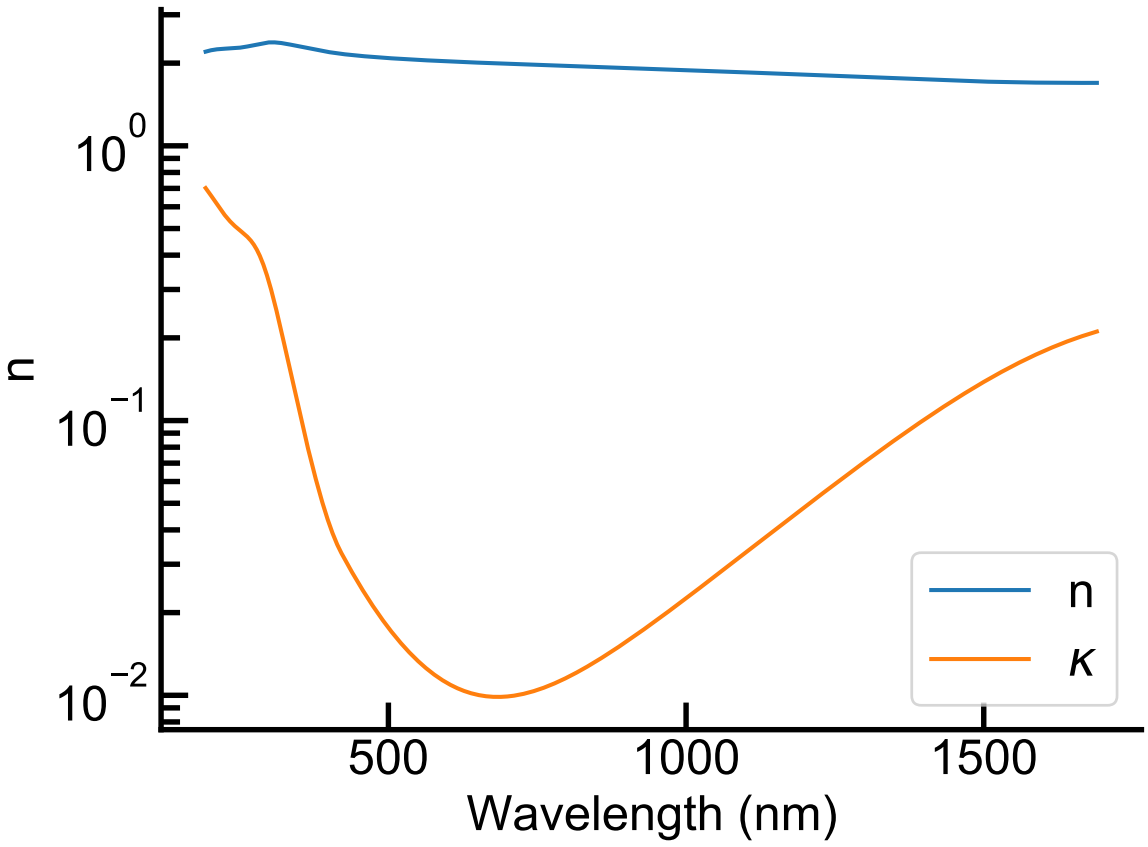

Figure 4.11: ITO complex refractive index from 250nm - 1700nm, measured by Covalent Metrology. Complex index at 1300nm is 1.78 + 0.071i.

the resist in the ASML stepper used previously, and developed with a 130C post-exposure bake using 45 seconds of MF-26A. I then evaporated 300nm of Al and performed a gentle liftoff (as the adhesion to the underlying ITO was poor) by soaking in NMP at 80C for 30 minutes. I then coated the wafers in $2\mu m$ of i-line resist to protect during dicing, diced 5mm x 5mm square chips, and rinsed them in acetone to remove the protective resist. I made electrical contact to the substrate by bonding each die to a printed circuit board with printed epoxy, and the top of each die by wirebonding to bond pads.

Future Materials - initial attempt to deposit barium titanate (BTO)

While this process relied on direct deposition of a crystalline piezoelectric material (aluminum nitride) on an amorphous Si_O2 film, this is not always possible.

Most piezoelectric materials do not easily self-organize into a crystalline structure, and must be grown on a crystalline substrate that matches the target crytsalline lattice. One material which is an attractive replacement for AlN is barium titanate (BTO). This material has a piezoelectric coefficient as high as 300pm/V [144], roughly two orders of magnitude that of AlN. However, it cannot be grown directly on amorphous material, and requires a seed layer or lattice-matched crystalline substrate. I attempted to use crystalline platinum as a seed layer, which other authors reported had been used to make high-quality BTO thin films [107].

Platinum Seed Layer Deposition

Initially, I attempted to deposit crystalline platinum in our sputtering chamber. However, my first attempt had an extremely broad rocking curve full-width-half-max (FWHM) of 20 degrees, indicating the platinum was poorly orietned. To optimize the crystallinity of the platinum, I ran a statistically designed response surface experiment, varying the power on a 3-inch target from 50W to 125W, and the pressure from 1.4mtorr (the lowest achievable) to 15mtorr (the highest achievable). The full experiment consisted of 10 runs, with a center replicate at the beginning and one at the end. After collecting the data on our X-ray diffractometer, I fit a model to the data, and obtained an excellent two-term model. With only the two-factor power-pressure interaction and a quadratic main effect of pressure ($R^2 = 0.99$), with a prediction expression of:

$$RFWHM = 11.5 + -0.0167pr * P + 0.326 * pr^2 \tag{4.3}$$

Where P is the power in Watts and pr is the pressure in mtorr. Ideally, the pressure is low (and it cannot be mode lower than 1.4mtorr), and the power is high. Unfortunately, this expression indicates a power of 500W is required, which is more than can be delivered with our existing pulsed DC supply. For this reason, we ordered a HIPIMS power supply, which can achieve instantaneous power densities of ¿2kW while maintaining low average power densities. Unfortunately, this HIPIMS unit was broken for the remainder of my Ph.D., and so I could do no further work to optimize the platinum seed layer.

Layer Transfer process

Platinum is not nearly reflective enough to serve as a mirror on its own, so it must eventually be removed and replaced with a Bragg mirror. To do this, I developed a film transfer process, which first grew the piezoelectric material on a seed layer, followed by Bragg mirror

deposition, wafer bonding, etch-back, and deposition of the second mirror. The Bragg mirro deposition followed the same process described previously.

SiO2 - Si/SiO2 wafer bonding

The most challenging part of the layer transfer process was developing and optimizing the wafer bonding process. Wafer bonding is extremely finicky, it requires first that the surface roughness of each wafer to be bonded has atom-scale roughness ˜$< 1nm$ [50], and that both surfaces to be bonded be completely free of microscopic dust particles. This requires a higher level of cleanliness than the typical academic cleanroom, and extreme caution to avoid contaminating the surfaces with particles. Even a single microscopic particle can cause a very large (˜1-5cm) void in the bonded wafer pairs, preventing film transfer at that location and potentially compromising the mechanical integrity of the bonded wafer pair.

I chose to use SiO_2-SiO_2 wafer bonding because (1) SiO_2 is compatible with virtually all equipment in the cleanroom, (2) my Bragg mirrors were already going to be made out of SiO_2 and (3) it is easy to CMP to achieve $< 1nm$ surface roughness. To perform the bond, I deposited SiO_2 by PECVD or thermal oxidation. I used the SiO_2 recipe developed previously (see Table 4.2), and performed the thermal oxidation in a wet oxidation furnace on bare p-Si wafers at 1000C.

After performing CMP to reduce the surface roughness, I cleaned the wafers in dilute (1:1000 v/v) ammonium hydroxide, forcefully wiping the cleaned surface with a sponge to remove any particles, followed by rinsing using pressurized wator, ultrasonicating in dilute ammonium hydroxide, quick-dump-rinsing the wafers, and drying them with compressed nitrogen. I confirmed the cleaning process removed all the CMP particles by AFM at several points on the wafer, and found no particles.

After the CMP cleaning, I ran both wafers to be bonded through our standard furnace pre-clean process, which consists of the following, with quick-dump-rinse (QDR) steps between each: (1) 10 minutes in pirhana at 120C (2) spin rinse dry (SRD) (3) 5 minutes in pirhana (4) 1 minute in HCl and (5) SRD. Once the wafers were in the final SRD step, I prepared a dedicated wafer bonding box which I previously wiped down with IPA-coated Tek wipes. I made sure to clean all surfaces the wafers would be near with IPA-coated Tek wipes, including my tweezers, all accessible surfaces within ˜3 meters of the wafer bonding tool, and the table and wall near where the wafers were transferred from the SRD to the wafer bonding box. To minimize the time the wafers spent in air, I transferred the teflon cassette directly from the SRD into the bonding wafer box, and closed the lid of that box. I then quickly walked to the wafer bonding tool, opened the chamber, placed and secured the wafers, and closed the chamber. The total time the wafers were exposed to air after SRD was less than 30 seconds, and this was sufficient to create consistently void-free wafers. I verified there were no voids in the bonded wafers by using a transmissive infrared camera.

For the bond process, I used aml-bonder (AML AWB-08 Wafer Bonder) radical-assisted fusion bonding. After loading the wafers into the chamber, I decreased their separation to 7mm, and then allowed the tool to pump down to $5 * 10^{-6}$ torr. I then engaged the direct

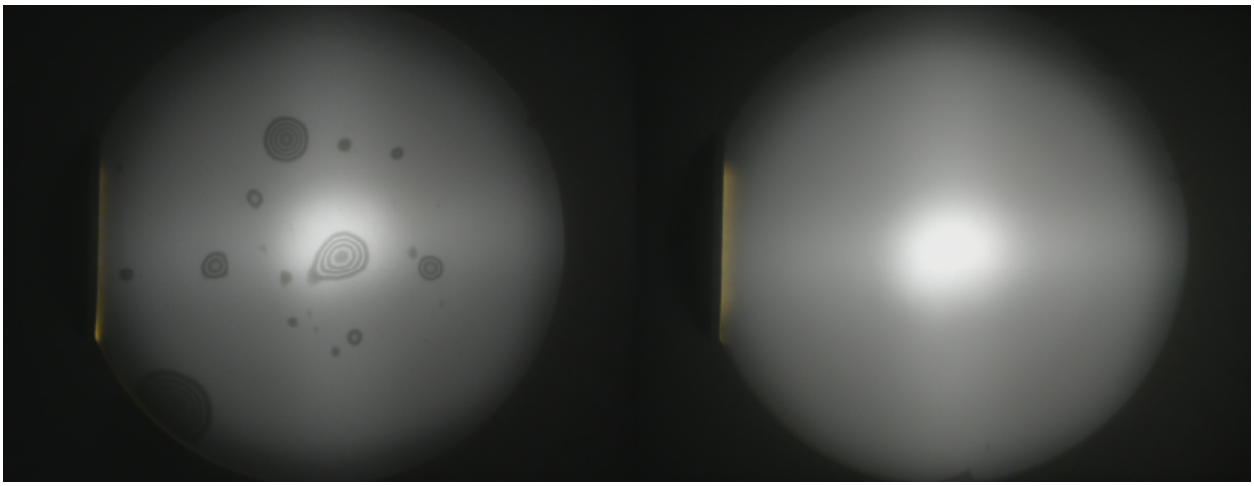

Figure 4.12: Bonded wafers before taking precautionary cleaning measures with several voids (left) and after precautionary measures with no voids (right)

bonding pin, forcing the center of the top wafer to bow outward. I injected O_2 into the chamber and set the pressure to 0.75 torr, and then applied 500V / 150mA of RF power to the platens to generate an O_2 plasma. After sustaining the plasma for 5 minutes, I shut it off, and injected DI water into the chamber, which stabilized after 5 minutes to a pressure of 6 torr. I then brought the wafers together and applied 300N of force, and allowed them to sit for 5 minutes. After this, I removed the wafer clamps, removed the wafers from the chamber, and annealed them for 4 hours at 300C in a vacuum oven with nitrogen atmosphere. Using the wafer blade test, I found this anneal dramatically improved the adhesion of the two wafers, causing the bond energy to increase from $200mJ/cm^2$ to $1400mJ/cm^2$, more than sufficient for subsequent processing.

Wafer etch-back

The second most challenging part of the layer transfer process was the etch-back after wafer bonding. All etch-back was done with either KOH or TMAH using ˜10L of both at msink14. I initially decided to use potassium hydroxide (KOH) to etch through the original handle wafer, and protect the handle wafer with a ˜100nm layer of silicon nitride. However, because the handle wafer was already coated in thermal SiO_2 for the wafer bonding step, I instead elected to use tetramethylammonium hydroxide (TMAH), which has a selectivity of ¿1000:1 to Si:SiO_2, far superior than the Si:SiO_2 selectivity. I grew this SiO_2 on a bare double-side polished Si wafer in a wet oxidation furnace at 1000C.

One challenge of working with TMAH is its extreme toxicity to humans and ability to cause sudden death [76], comparable to hydrofluoric acid (HF). TMAH is particularly dangerous because it can cause death even with immediate washing of the area, unlike HF,

which can be treated. I had a minor panic attack after initially being exposed to a dilute solution, after improperly rinsing my gloves when a drop of concentrated TMAH landed on them. After that, I decided against using elbow-length gloves because of the difficulty in decontaminating them, and used standard wrist-length chemical resistant gloves paired with a long plastic gown.

Another difficulty I discovered in wet etching is extreme sensitivity to any defects in the film. Any contamination prior to protective film deposition (in my case, thermal oxidation), will be found and etched by the TMAH. In figure 4.13 I show a couple expamples of these defects.

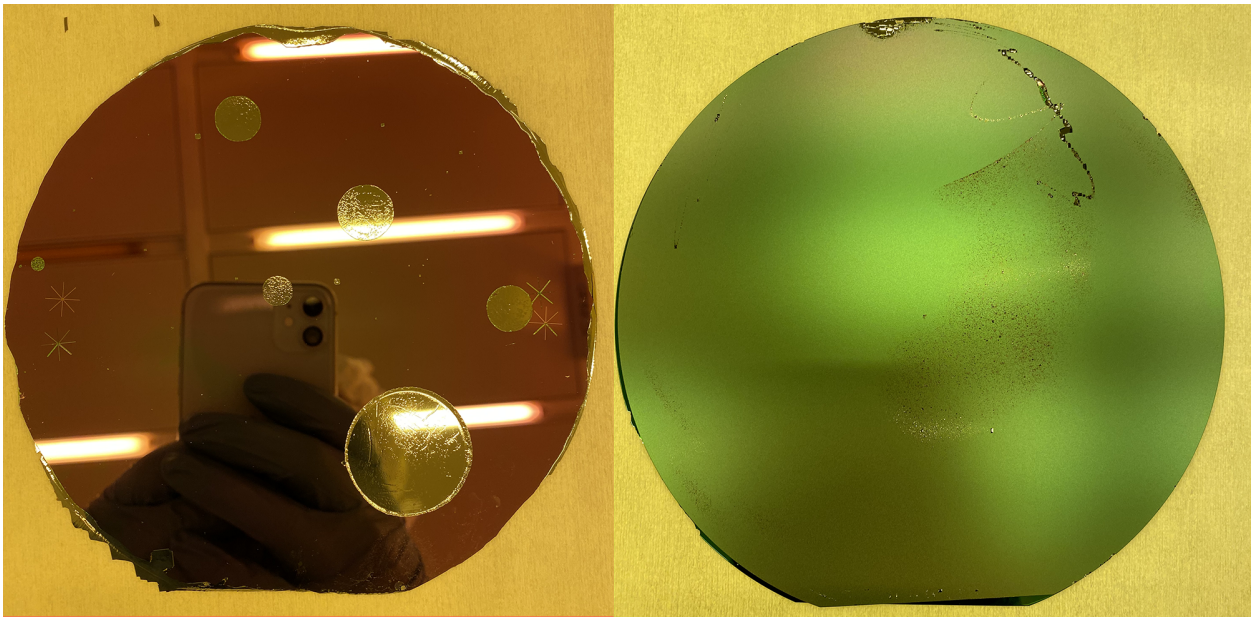

Figure 4.13: Two different bonded wafers after etch-back, with severe edge and front surface defects (left) and less severe edge and backside defects (right).

To fix this issue, I did not touch my handle wafers with vacuum wands when transferring them to the furnace for oxidation. Instead, I used a cleaned pair of tweezers. This eliminated the diffuse speckling defects I observed on the backside of the wafers. However, because the wafers must rest at two points on a surface for thermal oxidation, those points are masked from oxidation and can be attacked by the TMAH. To fix this issue, I split my oxide growth into two sequential oxide growth steps. Between the first and second steps I rotated my wafers by 90 degrees, being careful to use a wand touching only what would become the back surface of the handle wafer - this time using a wand on the backside surface to avoid scratching the wafers. Touching only the back surface helped to prevent difficulties in wafer bonding. I was able to use a wand at this point because surface contamination *after* the protective layer is deposited does not prevent the film from protecting the wafer.

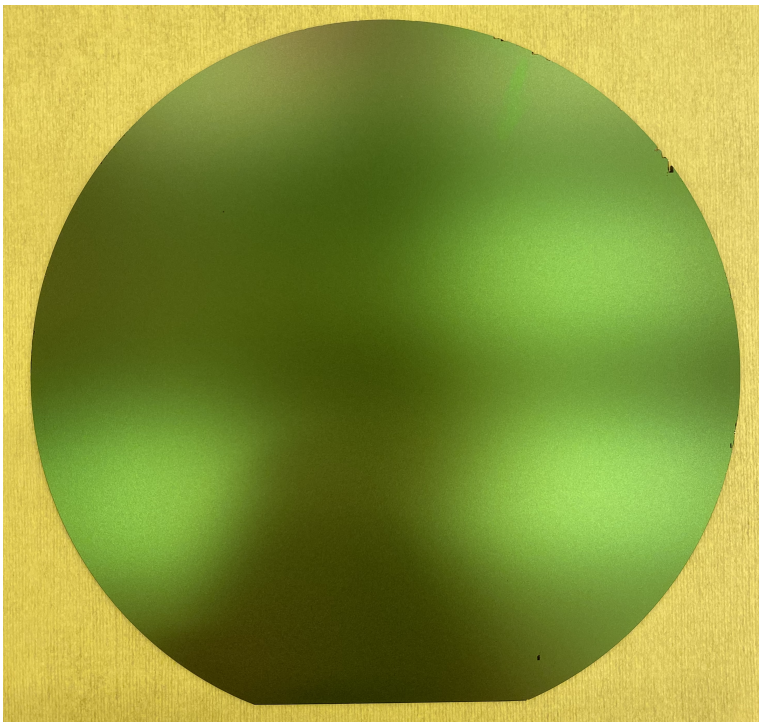

Figure 4.14: Backside of etched handle wafer after TMAH etching using careful handling and two-stage oxide growth. Note the small edge defect in the upper-right corner, but otherwise defect-free backside. This defect did not pose a problem in any of the equipment used subsequently.

After taking extreme care for wafers going into TMAH etch, I was able to eliminate diffuse backside defects and all major edge defects. Some ˜1mm edge defects were still present even after following the above steps.

Chapter 5

Conclusions and Future Work

"A work is never completed, but merely abandoned."
- Paul Valery

In this work, I have made substantial progress towards the goal of optical transduction of extracellular neural activity, as well as demonstrated the feasibility of using microfabricated optical voltage sensors for grid monitoring. I have introduced a figure of merit, the energy per quanta, or E_Q, which allows cross-modality comparison of the energy efficiency of different technologies which have a well-defined SNR, bandwidth, and power dissipation. This allows for quantitative comparison of different technologies to be made which target the same application, and demonstrates the power of optical methods in general.

5.1 The Roadmap to optical neural voltage sensing

While much work has been done on theory, device fabrication, and benchtop characterization, there are several future device-level improvements that will be required to achieve the desired device performance. I believe that direct measurement of neural activity with piezoelectric and/or pockels-effect based optical voltage sensors has potential to be useful in basic research and clinical application. The potential for wireless megascale neural recording exists and could in principle suprass the capabilities of traditional electronics.

5.2 Device-Level Improvements

Improved Piezoelectric Materials

The most significant area of improvement is use of a piezoelectric film with a higher d_{33} value. A higher d_{33} produces a higher device gain β, which gives a lower E_Q. Aluminum Nitride, while easy to deposit, has a relatively low piezoelectric constant ($d_{33} \approx 5pm/V$ [89]. Other materials, such as barium titanate and KNN, have orders of magnitude higher

piezoelectric coefficients (BTO 300pm/V [144]), and theoretical bandgaps comparable to AlN, which allows for essentially zero absorption in the near-IR range. The synthesis of thin films of these materials has been done previously, but doing so reliably may require equipment found outside the Marvell nanolab.

Modified Fabrication Process

As previously stated, most piezoelectric materials cannot be grown directly on an amorphous substrate. This poses a challenge - because materials which can be feasibly used to make Bragg mirrors in the near-IR range are, for the most part, amorphous. This necessitates the use of film transfer. However, for even better performance, the capacitive division caused by the Bragg mirrors on either side should be eliminated. Traditionally, such as in VCSELS [71], this is done by using conductive Bragg mirrors. I was not able to find a set of conductive materials with reasonably high index contrast compatible with the fabrication equipment I would need to complete this project over the course of my Ph.D. If such a solution could be found, that would eliminate this capacitive division and would make the fabrication process considerably easier.

Alternatively, by inserting thin layers of transparent conducting oxide, such as ITO or ZnO on either side of the piezoelectic layer between that layer and the Bragg mirrors, and then etching vias to that ITO layer, the capacitive division could be eliminated. This would require a very thin layer of conducting oxide due to the loss of this material in the near-IR. However, the problem is not as bad as it initially seems, because the optical mode in the region at the interface between the Bragg mirror and the piezoelectric layer is nearly zero. This means any loss caused by the conducting oxide would be greatly reduced.

For example, with 5nm of the ITO used in this work, the Q-factor is reduced to 1700 from 2900. With 10nm of ITO, this is reduced to 1200. This ITO, relatively speaking, had poor conductivity and was more optically lossy than typical ITO. With optimization, Q-factors of 3000-5000 should not be difficult to achieve.

Thin layers are also feasible because SF6, which etches SiO_2 and Si, the materials used here for the Bragg mirrors, does not etch ITO. In a short experiment I performed, I measured the etch selectivity at $> 1000 : 1$. Even 1-5nm of the film would be feasible to use as an etch stop layer.

Improved transparent conducting oxide

In this work, I did not do any optimization of the indium tin oxide (ITO) used as a transparent conducting topside layer. This was for purely practical reasons, as by the time I needed to optimize this film, our primary deposition tool broke. I have no reason to believe reducing the loss and improving the conductivity of the ITO layer would be unusually difficult. The lossy ITO resulted in substantial optical loss coupling into and out of the device (about 30% at our wavelength of interest), and an increase in device E_Q by roughly a factor of 2.

However, there is much room for improvement of the loss coefficient and conductivity of the film. Alternatively, other films such as ZnO may be used.

Miniaturization and Optical Alignment

All of the characterization in this work was done without the aid of a microscope and without much beam quality optimization or investigation. For this reason, we were not able to reliably align to devices smaller than $160\mu m$, and our measured device gain was more than an order of magnitude lower for the smaller $80\mu m$ devices, which I believe was caused by poor optical coupling. I did not even attempt to align to the $40\mu m$ devices because of the difficulty encountered in aligning to the $80\mu m$ devices. I had begun to adapt a microscope to provide both visual alignment and near-IR excitation of the device in question, but did not finish by the time I was ready to graduate. The modifications required were relatively straightforward.

In smaller devices, spatial mode profile matching starts to determine how much optical power can be coupled into the device, and can be calculated to first order by calculating the overlap integral of the incident mode and the device mode, as with optical fiber coupling [145]. This requires knowing exactly what the beam profile looks like, including how large it is at the device plane. I expect using a beam profiler will be necessary for this, along with filtering of the beam.

5.3 Remaining System-Level Challenges

In addition to the future device-level work, several practical challenges at the system level remain to be addressed and overcome.

Insertion Loss

Optical fibers were used in this system to allow for the device to be connected to multiple sensors (here an OSA, power meter, and photodiode). However, this is only necessary for characterization purposes, and direct coupling to a photodiode, along with using a polarization-sensitive beamsplitter could improve the insertion loss by just over ˜20dB, reducing the E_Q to an estimated $1.3\mu J$.

Optical Scattering

If these devices are to be placed underneath skin, skull, or dura mater, substantial scattering will happen between the source and the incident wavefront on the device surface. Scattering not only widens the incident beam, but also causes the angular spectrum reaching the devices to widen. Beyond a few scattering lengths, mode coupling any appreciable power to these devices is impractical. For this reason, the scattering must either be eliminated - for example

by replacing skull with a transparent window [148], or corrected for using adaptive optics techniques [156]. Scattering is probably the most serious of the practical challenges. However, this is an important problem in many domains, and is currently receiving a huge amount of attention in the academic community.

Optical Alignment

These devices require precise angular and spatial alignment. The angle of incidence of resonant cavities changes the resonant wavelength in planar thin films like 1 $cos(\theta)$, or approximately quadratic for small θ. For the devices described here a 2.5° angle of incidence corresponds to an 0.4nm resonant wavelength shift, and a 5° angle of incidence corresponds to 1.6nm. For smaller devices, the phase ramp caused by any angle deviation causes imperfect mode coupling, with the coupling coefficient $\langle\psi_{incident}|\psi_{device}\rangle$. For a $20\mu m$ device, this corresponds to a maximum possible deviation of 5° before coupling reaches half its initial value. If devices are free-floating and not mechanically stabilized or aligned, this poses substantial practical challenges. If multiple devices are used, ensuring the phase front is flat at every one of them poses even more of a challenge, and necessitates some type of spatial light modulation with closed-loop feedback to be used to deliver light to each of the devices.

Multiplexing

While collecting data from a single device, collecting data from many such devices is the ultimate goal. I believe the most feasible way to do this is with time-domain multiplexing. By applying short bursts of light to each device, and switching rapidly from one device to the next, then recording the resultant light with a single high-speed photodetector and de-multiplexing in time, it shuold be possible to record from many devices at once. This is feasible because the optical power for a single device is tiny compared to the power available from a typical optical source, and the bandwidth and noise performance available in a high-speed photodetector and amplification system is sufficient to multiplex up to millions of channels. The central challenge here is finding a way to rapidly raster a beam from one device to the next - traditional spatial light modulators provide nowhere near the bandwidth to do this.

Frequency-domain multiplexing is also an interesting option, and would require modulating the beam reaching each individual device at a different frequency. This would likely be much easier to do from an optical systems perspectives, but requires designing circuitry which can modulate many different channels.

I do not believe wavelength-division multiplexing is a feasible option, as it requires precise detuning of each device's resonant frequency, coupling of many monochromatic sources or precise filtering and de-coupling of a single source. In practice, every filter would likely have to be tunable, as the device's resonant wavelength cannot be tightly controlled.

5.4 Other potential avenues for future work

Photonic crystals

Instead of using Bragg mirrors to enhance the quality factor, it is possible to use a single film which is patterned in x- and/or y [103]. These structures, called photonic crystals, have interesting properties including resonances with arbitrarily high quality factors which are very sensitive to geometry. This could be exploited both with piezoelectric materials that have a high d_{33} or d_{31}. This is significant, because some materials (such as BTO) have d_{31} values that far exceed their d_{33}. Using a photonic crystal would greatly simplify the fabrication process, and if such a photonic crystal can be designed and reliably fabricated with photolithography, it has the potential to exceed the performance of film-based devices. The `rcwa` package I wrote and described in this dissertation and elsewhere was purpose-built for the simulation of photonic crystals. I expect this to be an area rich in unexplored possibility.

Integrated Photonics

Another potential avenue of exploitation is the use of integrated photonics for the direct measurement of neural activity. This has the advantage of being highly scalable. However, coupling into and out of integrated photonic waveguides is extremely difficult and requires micron-scale alignment, typically done with micromanipulators. I am uncertain whether the practical challenges posed by integrated photonics can be overcome to make it a feasible neural recording solution.

Both integrated photonics and photonic crystals are active areas of research, and I am confident that advances in these fields will make them more attractive for this applicaiton in the near- to mid-term future.

Competing Techniques

There are many ways to measure neural activity. This technique is likely to be most useful at scale - when many thousands or millions of channels are required. Optical sensing with genetically encoded voltage indicators (GEVIs) and similar techniques, are an alternative way to convert neural activity to optical information, and already represents a valuable tool in basic research, and for surface-level neurons, represents a superior technology in many ways.

Appendix A

A diversion with COVID-19

In March of 2020, COVID-19 began spreading rapidly in the United States. In response, shelter-in-place orders were enacted throughout California. Virtually all non-computational research activity was suspended. I, personally, was not able to access shared facilities or my group's laboratory. This was sudden, shocking, and scary. In response to the pandemic, UC Berkeley formed a COVID-19 response working group, of which my advisor, Michel, was a leading member. The goal of this working group was to direct the brains, resources, and firepower of the academy, currently dormant, to solving some of the problems arisen due to COVID-19. Some parts of the group worked on ventilators, which at the time were believed to be in short supply, some worked on remote access of ventilators, and our group worked primarily on adding airflow feedback to existing DIY ventilators in a way that could circumvent supply chain problems. While it is easy to look back on this work as a misguided response to a nonexistent problem, at the time the ventilator shortage problem was so feared that the FDA had issued guidance to doctors allowing them to use CPAP machines if typical ventilators were not available. The need for this work appeared urgent and dire. We are all very thankful the imagined worst-case scenario did not play out. The following is adapted from [40].

Abstract

COVID-19 has become a significant burden on the healthcare systems in the United States and around the world, with many patients requiring invasive mechanical ventilation (IMV) to survive. Close monitoring of patients is critical, with total volume per breath (tidal volume) being one of the most important data points. However, ventilators are complex and expensive devices, typically in the range of tens of thousands of US dollars, and are challenging to manufacture, typically requiring months. Solutions which could augment the ventilator supply rapidly and at low cost in the United States and elsewhere would be valuable. In this paper, we present a standalone tidal volume measurement system consisting of a D-Lite spirometer, pressure sensor, microcontroller, and tubing with a cost of parts less than $50 USD. We also provide a model to predict the error in tidal volume measurements

based on the pressure sensor used and the flow during ventilation. We validate this system and show that the tidal volume accuracy for flows above $10L/min$ was within 10%. We envision this system being used to increase the ventilator supply in resource-constrained settings.

A.1 Introduction

As of August 25, 2020, there have been over 23 million cases and 815,000 deaths from the COVID-19 disease caused by the SARS-CoV-2 virus[27], with hospitals across the world shouldering the burden for treatment in most severe cases. In the United States, of those who are admitted to the hospital for COVID-19, over 20% are subsequently admitted to the intensive care unit (ICU), although this number is decreasing as we learn how to better treat this disease. The majority of these patients (69%) require invasive mechanical ventilation (IMV) for a mean duration of 10 days [19].

Monitoring of mechanical ventilation is critically important to ensure patients are receiving enough air. Typically, this is done by delivering a set volume of air to the patient in a given breath (tidal volume) based on their body weight [97]. Insufficient volume can lead to hypoventilation and CO2 buildup, or atelectasis (lung collapse). From the onset of the pandemic, there were concerns about equipment shortages, including intensive care unit (ICU) ventilators [54, 118]. This led to some companies re-purposing facilities to participate in ventilator production [54], the U.S. president invoking the Defense Production act [94], as well as a number of emergency ventilator design efforts to bypass supply chain issues [137, 110, 84, 17]. The shortage was so severe governmental agencies released emergency guidelines, including an FDA guidance allowing clinicians to repurpose BiPAP (bilevel positive airway pressure) machines commonly used as a treatment for sleep apnea, anesthesia machines, and simpler ventilators for emergency use if needed [43, 132].

These events have made clear the importance of developing a low-cost, accurate, and easily deployed system that can supplement the existing IMV supply chain in future catastrophes. Most existing efforts to this end are focused on ventilators which are very cheap and but only provide pressure measurements [137, 110], or on designs that rely on external and expensive flow meters [110]. In this work, we present a middle ground to these two approaches: a guide to designing a cheap and accurate flow sensor for measurement of tidal volume, based around an easy-to-manufacture plastic spirometer which has been around for several decades. The advent of accessible microcontrollers and the rapidly falling cost of silicon electronics enable a low-cost, high-accuracy tidal volume measurement system to be built with interchangable parts which could supplement the functionality of existing minimal-feature ventilators or alternative respiratory devices, such as BiPAP machines. We envision such a device could be crucial in future pandemics and may also serve an important role in current low-resource settings.

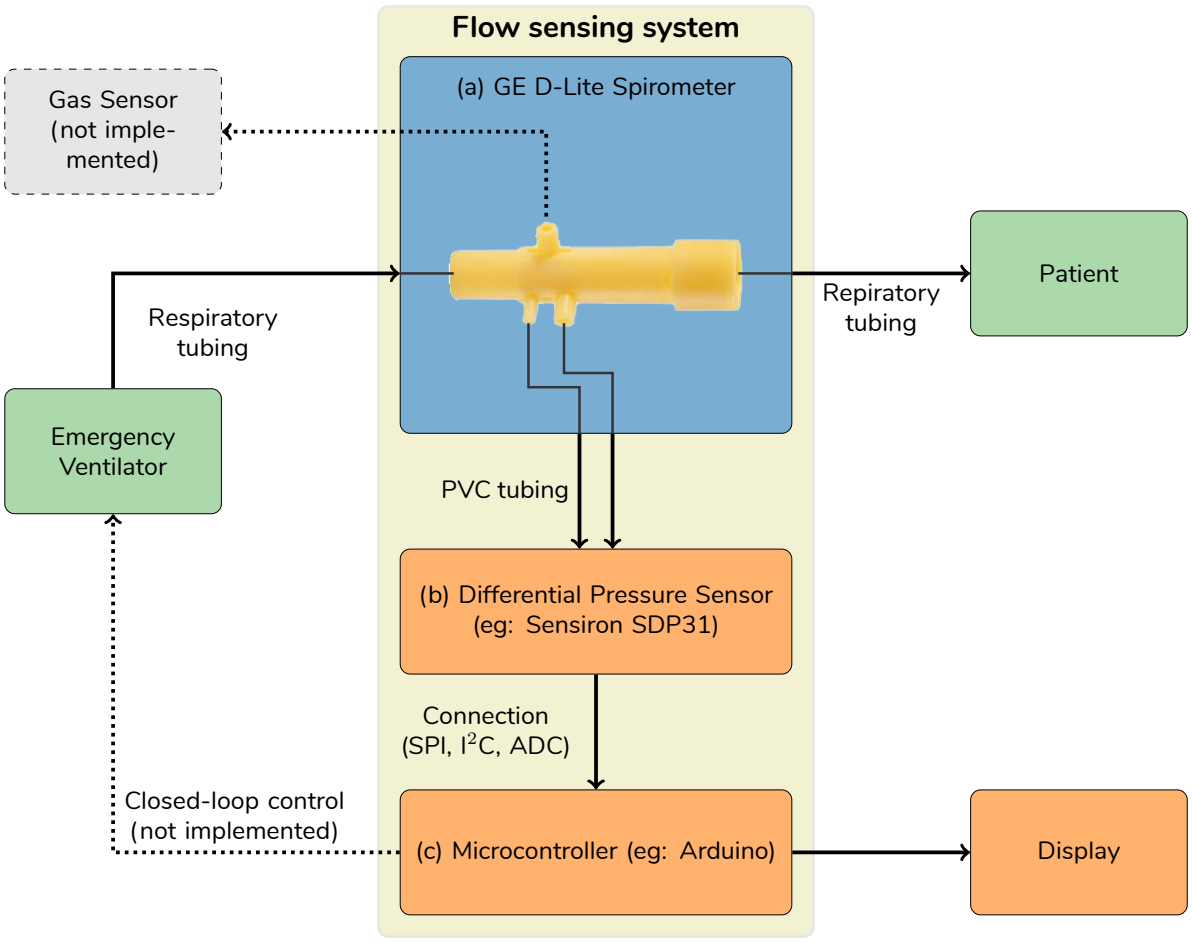

Figure A.1: Overview of flow sensing system and integration in a clinical setting. The D-Lite spirometer fits into sections of standard $22mm$ ventilator tubing (with an adapter at the smaller end). The dynamic pressure drop is sensed across the two oppositely-directed terminals by a pressure sensor (e.g. the Sensiron SDP31), and this information is sent to a central processing unit (e.g. an Arduino). This information can then be displayed or used in closed-loop feedback.

A.2 Methods

Our tidal volume meter consists of three main components: (a) a passive spirometer, (b) a differential pressure sensor, (c) a computational unit. Figure A.1 shows how these components interact with each other and a larger system.

Design Considerations

For the passive spirometer, we selected the General Electric D-Lite ++ spirometer [95, 55], which transduces flow velocity into a pressure difference. Although this part is available online, we developed a similar model that can be 3D printed (see Suplementary information). The device has two Pitot tubes, one upstream of a flow constriction and facing upstream, and the other downstream of the flow constriction and facing downstream, and a third port for gas sensing.

The dynamic pressure between between both Pitot tubes depends on the flow by:

$$\Delta P = kQ^2 \tag{A.1}$$

where ΔP is the pressure difference in cmH_2O, and Q is the flow in L/min. A proportionally constant k in $cmH_2O/(L/min)^2$ depends on air density, the spirometer channel geometry, and the Reynolds number [95]. Since the pressure difference also changes sign for reverse flow, there is a one-to-one correspondence between flow and pressure. However, due to the quadratic nature of the relationship, measurements at very low flows are challenging. Most regulations require reporting of tidal volume to a given relative accuracy, such as 10 % [132]. In order to achieve this, a sufficiently accurate pressure sensor must be selected, and the error should be quantified. The error in volume measurements can be found integrating equation (1) over time, and assuming some absolute pressure accuracy. We show this accuracy in Figure A.2 as a function of sensor pressure accuracy and flow.

Among electronic pressure sensors, the key features involve communication protocol, pressure range, wide commercial availability, and accuracy, as a percentage of full scale. We selected digital pressure sensors to minimize part count and the need for a PCB, and settled on the Sensiron SDP31 ($0.001cmH_2O$ absolute accuracy) and a less accurate Honeywell Honewell SSCMRRN060MDSA5 (-$6kPa$ to +$6kPa$, 0.25% of full scale accuracy).

A microcontroller or similar computational unit is needed to compute flow and integrate it over time for tidal volume. Many emergency ventilators already have a comparable device for control and sensing [84, 17]. We chose an Arduino for its well-supported software ecosystem and wide availability, and these can be purchased off-the-shelf for as low as $5. The Arduino runs a finite state machine (see Sumplementary Materical) to read the pressure, convert it into flow, detect breath onset with a simple threshold, and compute tidal volume, with a sampling and computational frequency of 167 Hz. This is then reported to a computer for display or further analysis. The spirometer was calibrated by comparing pressure measured with both sensors against flow, where the flow was measured with a gold standard Alicat M-Series 1000 SLMP mass flow meter.

Pressure Sensor and Ventilator Characterization

We connected the Honewell differential pressure sensor to an Arduino using SPI and measured data at $166Hz$. To ensure the Honeywell sensor was operating as expected, we measured its pressure reading against that of a water-column manometer in the range of 0-4

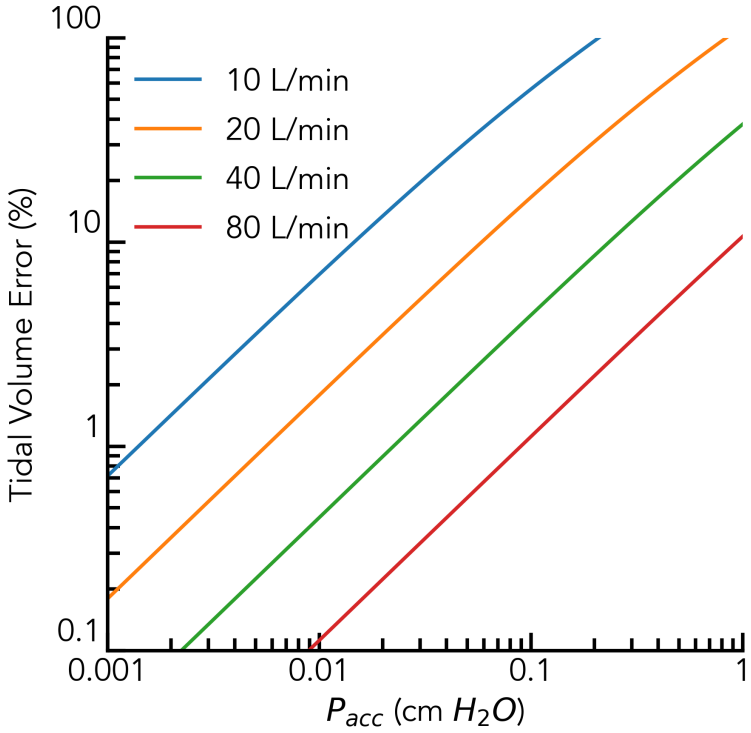

Figure A.2: Tidal volume error vs. pressure sensor absolute accuracy at varying flows. Each curve corresponds to a different instantaneous flow (in which case the results are exact), or average flow over a sensing period, in which case they are approximate. The Sensiron SDP31 has an absolute accuracy of $0.001 cmH_2O$, while the Honeywell SSCMRRN060MDSA5 has an accuracy of $0.01 cmH_2O$. Note that the error increases with worsening accuracy and lower flow

cmH_2O (Supplementary Figure S1) by varying the flow through the spirometer. This information was used to correct an offset of $-0.02 cmH_2O$. Having validated the sensor was behaving as expected, we characterized the pressure of the D-Lite spirometer as a function of flow using both the Honeywell sensor ($0.01 cmH_2O$ absolute accuracy) and a more accurate sensor (Sensiron SDP31, $0.001 cmH_2O$ water absolute accuracy). We then offset-corrected data from the Honeywell and the Sensiron and fitted the resulting flow vs. pressure to a quadratic function (Figure A.3). We additionally corrected the Sensiron data for the pressure drop from the measurement tubes (1/16in ID, 1 meter length) using a company-provided datasheet with correction factors [3].

We investigated transient behaviour for typical ventilator breaths by connecting the spirometer to a Bird Mark 7 pressure controlled ventilator and a rubber test lung (Figure A.5) with the Honeywell sensor. Similarly, we measured the transient behaviour for a constant flow emergency ventilator (an early prototype of the Stanford Rapid Response ven-

tilator) [119] and a Dua-Adult Training and Test Lung Model 1600 (Michigan Instruments) (Figure A.4). The device interrupts pressurized flow, with a constant-flow valve, which drops a regulated 50 psi of pressure across it. The spirometer was connected in several different locations: the inspiratory limb, expiratory limb, and near the test lung such that it recorded inspiration only, expiration only, and both inspiration and expiration for each breath. We tested nominal flows (set by a needle valve and determined by a rotameter) of 10, 30, and $50 L/min$ as well as lung compliances of 10, 30, and $150 mL/(cmH_2O)$, with at least 138 breaths in each condition. Each breath lasted 4 seconds with a 1:3 inspiration to expiration ratio (I:E).

A.3 Results

We found the proportionality constant k to be $6.97{*}10^{-4} cmH_2O/(L/min)^2$ for the Honeywell with a standard error of the regression of $0.06 cmH_2O$, and $6.43 * 10^{-4} cmH_2O/(L/min)^2$ for the Sensiron with a standard error of the regression of $0.03 cmH_2O$ prior to pressure drop correction. After this correction, the proportionality constant increased to $6.7{*}10^{-4} cmH_2O/(L/min)^2$.

Once the proportionality constant was known, we used this to measure flow from the Stanford Rapid-Response ventilator at various flows (Figure A.4). The measurements were very repeatable, with maximum flow standard deviations of 0.5% and and min/max differences of 2% of the expected volume (see Figure S2). To ensure our thresholding method was accurate, we verified it agreed with offline numerical integration (Figure S3). We also characterized an older model pressure-controlled ventilator, the Bird Mark 7, and show a typical breath curve in both pressure and volume (Figure A.5). Finally, we characterized the accuracy of our system at measuring tidal volume using a 500mL syringe, varying the breath time and volume (Figure A.6). The accuracy of these measurements is compared to our accuracy model (Figure A.7).

A.4 Discussion

We designed a low-cost differential-pressure based system to measure inhaled and exhaled flow based around the D-Lite differential pressure spirometer, which consists of a microcontroller, differential pressure sensor, D-Lite spirometer, and connecting wires and tubing. We used this system to characterize the D-Lite spirometer, and report the measured differential pressure vs. flow for two different pressure sensors. We measured the proportionality constant k to be between $6.43 * 10^{-4}$ and $6.97 * 10^{-4} cmH_2O/(L/min)^2$. We also developed a model to estimate the error in converting the differential pressure measurement to flow, depending on the pressure sensor used, and used this to predict the relative volume error vs. flow (Figure 7). The accuracy of our measurements of tidal volume exceeded the expected error of our sensor model, indicating there are unaccounted-for sources of experimental error which are not due to the sensor alone. The main source is the error in the proportionality

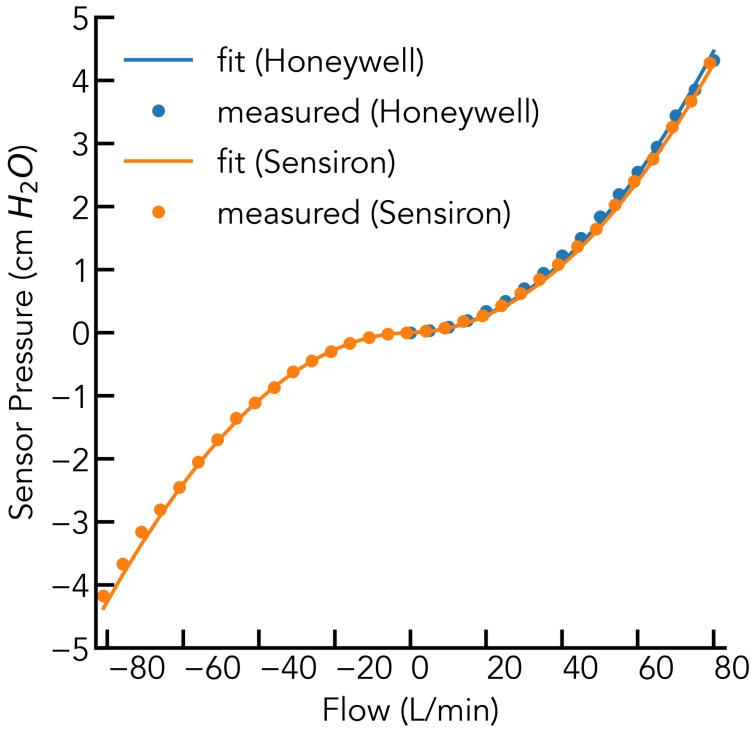

Figure A.3: Pressure vs. flow for the D-Lite spirometer as measured by Alicat mass flow meter. Honeywell data is corrected for zero offset prior to fitting to quadratic. Uncertainty in flow is$< 3L/min$ at all flows, and uncertainty in pressure is $< 0.01cmH_2O$

constant, as any error in that constant is (for small errors) linearly proportional to the error of a volume measurement. This level of accuracy is sufficient for adult breath measurement clinical use cases where flows tend to be larger ($> 10L/min$). With a cost of components of <\$50, such a system could be a practical add-on to otherwise simple mechanical ventilators in emergency or low-resource settings.

If high accuracy is needed at both low and high (relative to full-scale) flows, more than one pressure sensor may be used, one with a small full-scale range to ensure accuracy at low flows, and the other with a large full-scale to capture larger flows. This, however, increases the cost of the system, as the pressure sensor is the most expensive component.

As a proof of concept, we used the D-Lite spirometer and SDP31 differential pressure sensor to characterize the operation of the Stanford emergency response ventilator [119] (a flow-controlled ventilator), and the Bird Mark 7 . As can be seen in Figure 6, at flows near zero the flow discretization noise becomes very large (the derivative of the square root function becomes very large near zero). Even a zero offset of a few least significant bits (a very good pressure sensor) can correspond to a flow offset large enough to significantly affect

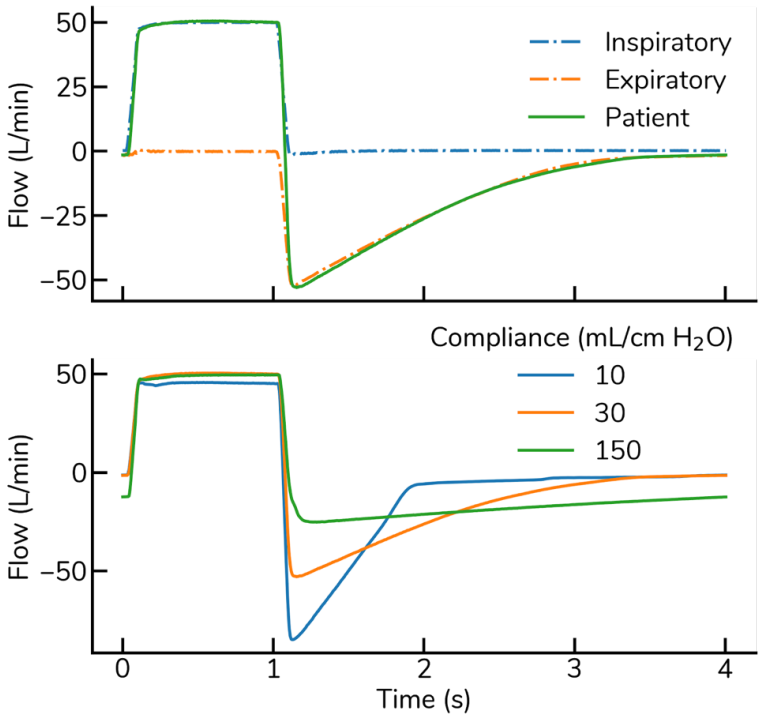

Figure A.4: Flow vs. time over a single 4-second breath cycle for the Stanford Rapid Response ventilator. Data was obtained from 138 breaths and averaged into a single plot (top). Test lung compliance was varied at a fixed flow, and the resultant flow profiles are plotted versus time (bottom).

the measured tidal volume.

We also demonstrated that our simple finite state machine based thresholding for calculating tidal volume on the Stanford rapid response ventilator was nearly identical to the performance of offline integration, with deviations of at most 2% from the integrated value, with a standard deviation of 0.5%. This was true across settings for typical flows (10, 20, and $50 L/min$) and lung compliances (10 - $150 mL/(cmH_2O)$ per lung) for a typical lung resistor (Rp20). As long as the sensor is sufficiently accurate so that the zero offset is below the minimum expected measurable flow, this should yield accurate tidal volumes for flow-controlled ventilators.

Taken together, these results demonstrate that a simple, low-cost system can be used to accurately determine tidal volume independently of the ventilator used. This could supplement the functionality of existing ventilators or alternative respiratory devices, and may help increase the ventilator supply in future pandemics or in low-resource settings.

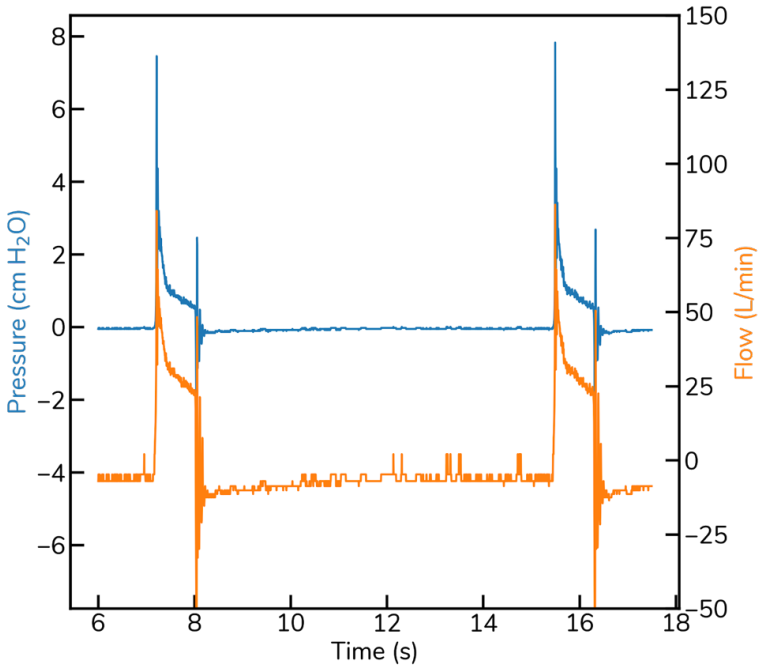

Figure A.5: Pressure and derived flow over two sample breaths given by a Bird Mark 7 to a simple test lung, as recorded by the Honeywell differential pressure sensor.

A.5 Conclusion

We designed a differential pressure based flow measurement system around the D-Lite spirometer. This system was shown to be accurate to within 15% for measurement of tidal volumes for all conditions tested (4L/min - 20L/min), and was used to characterize the Stanford rapid response ventilator. The data and methodology we present here could be used to design simple, low-cost flow measurement systems.

A.6 Acknowledgements

We would like to thank the UCSF COVID-19 Response Fund for supporting this work, as well as the Chan-Zuckerberg Biohub and Hertz Foundation for supporting this work. We would also like to thank members of the Stanford Linear Accelerator (SLAC), especially Professor Aaron Roodman and his team for useful conversations, in addition to members of 219 Design.

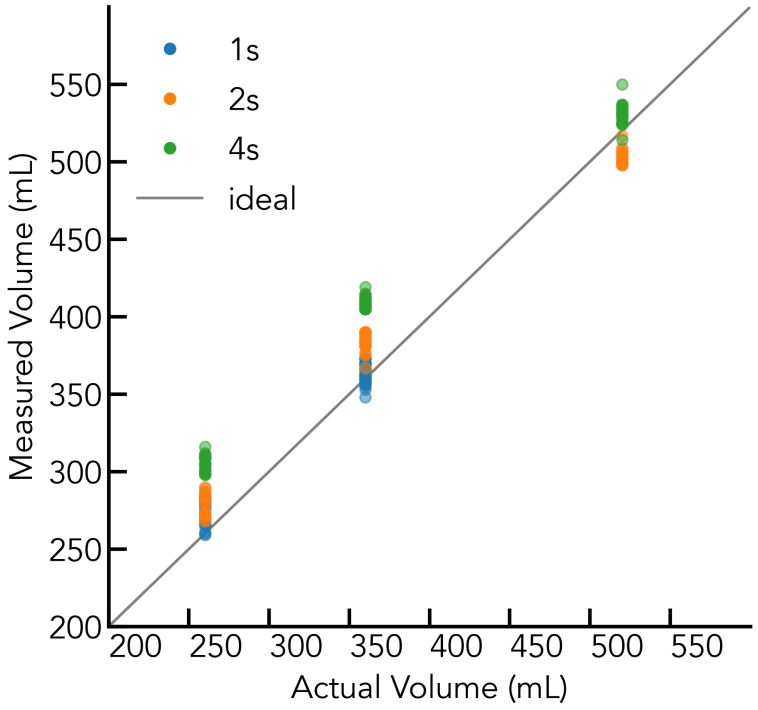

Figure A.6: Measured volume from syringe experiments using 1s, 2s, and 4s flow durations. Error increases at lower flows, with a maximum error of MAX at the lowest volume/flow rate ($3.75 L/min$), much lower than typical flow for an adult ventilator. Syringe plunges were performed visually, and the error increases with decreasing flow. Expected maximum error from device alone is 6% at the lowest flow, and the mean of the lowest-flow measurements are 10% away from the mean of the highest-flow measurements in the same condition.

A.7 Suplementary Material

Code and designs

Code for the Arduino and data processing is available at `https://github.com/edmundsj/TidalVolumeArduino`. Raw data is available at `https://drive.google.com/drive/u/0/folders/1wQxoaZd3DSQLKDs2eLh8qEjYdGccdzsO`.

Open source 3D printed designs for spirometry can be found at `https://3dprint.nih.gov/discover/3dpx-013632`.

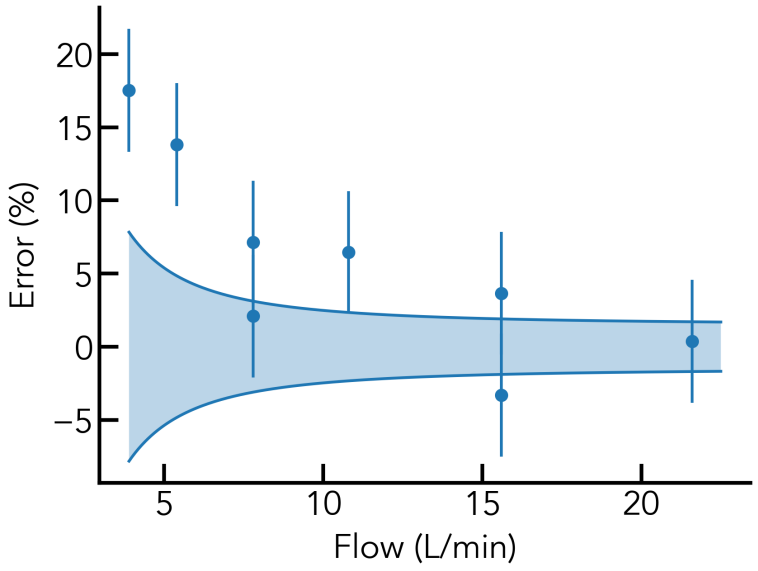

Figure A.7: Error vs. Measured flow taking the mean of inhaled and exhaled breaths (10 trials each, 20 total measurements per condition)

Table S1: Bill of Materials

Part Name	Source	Cost	Quantity	Total Cost
Arduino Due *	Amazon	$22	1	$22 ($5)
Arduino Nano	Amazon	$5	1	$5
D-Lite Spirometer	4MDMedical	$8ea	1	$8
Pressure Sensor (SDP31)	Digikey	$31	1	$31
Medical-grade tubing	US Plastics	$1/ft	1ft	$1
Total				$62 ($45)

* The Arduino Due can be subsituted for the Arduino Nano at a cost of $5.

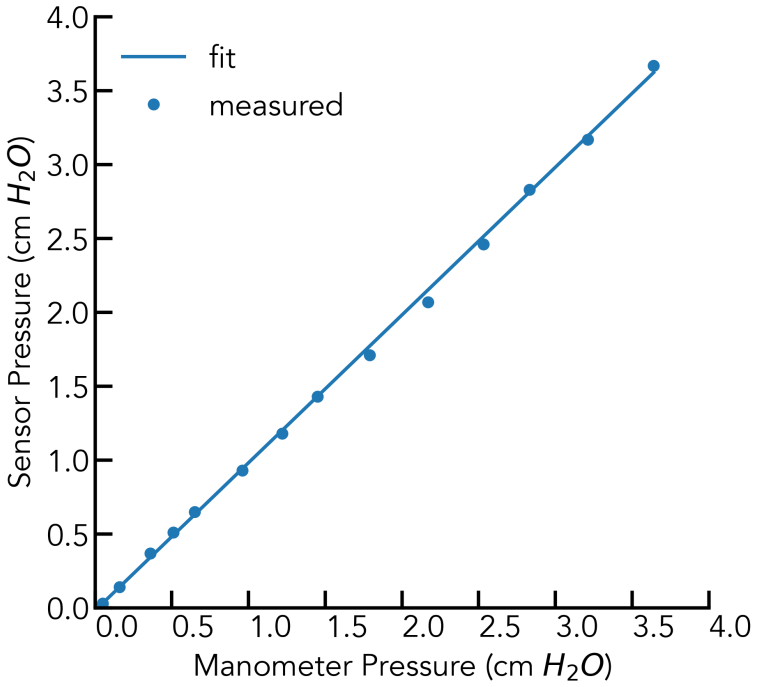

Figure S1: Measured pressure sensor data vs. data taken from water column manometer. Fit line is $y = 1.000x - 0.0185$, where x is the measured manometer pressure and y is the sensor pressure. Pictures of the water column manometer were taken using a stationary iPhone 11 at 2.7x optical zoom, and analyzed using ImageJ to find the height of the water.

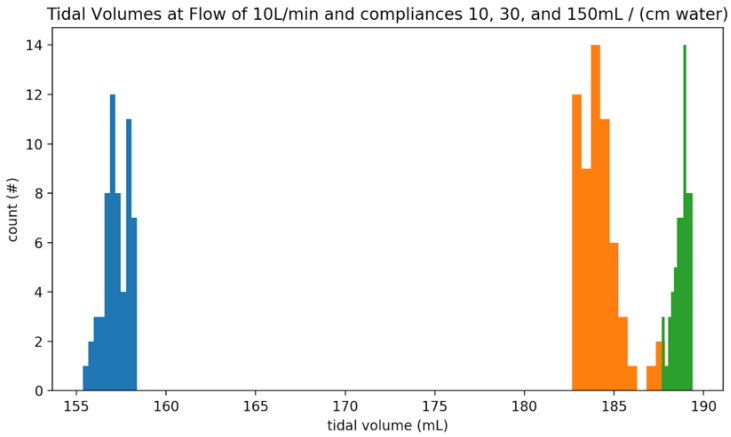

Figure S2: Histogram of measured tidal volumes of Stanford rapid-response ventilator at compliances of 10, 30, and $150mL/(cmH_2O$ at the lowest flow characterized ($10L/min$). The intra-compliance variability was <2% min/max with standard deviation <0.5%

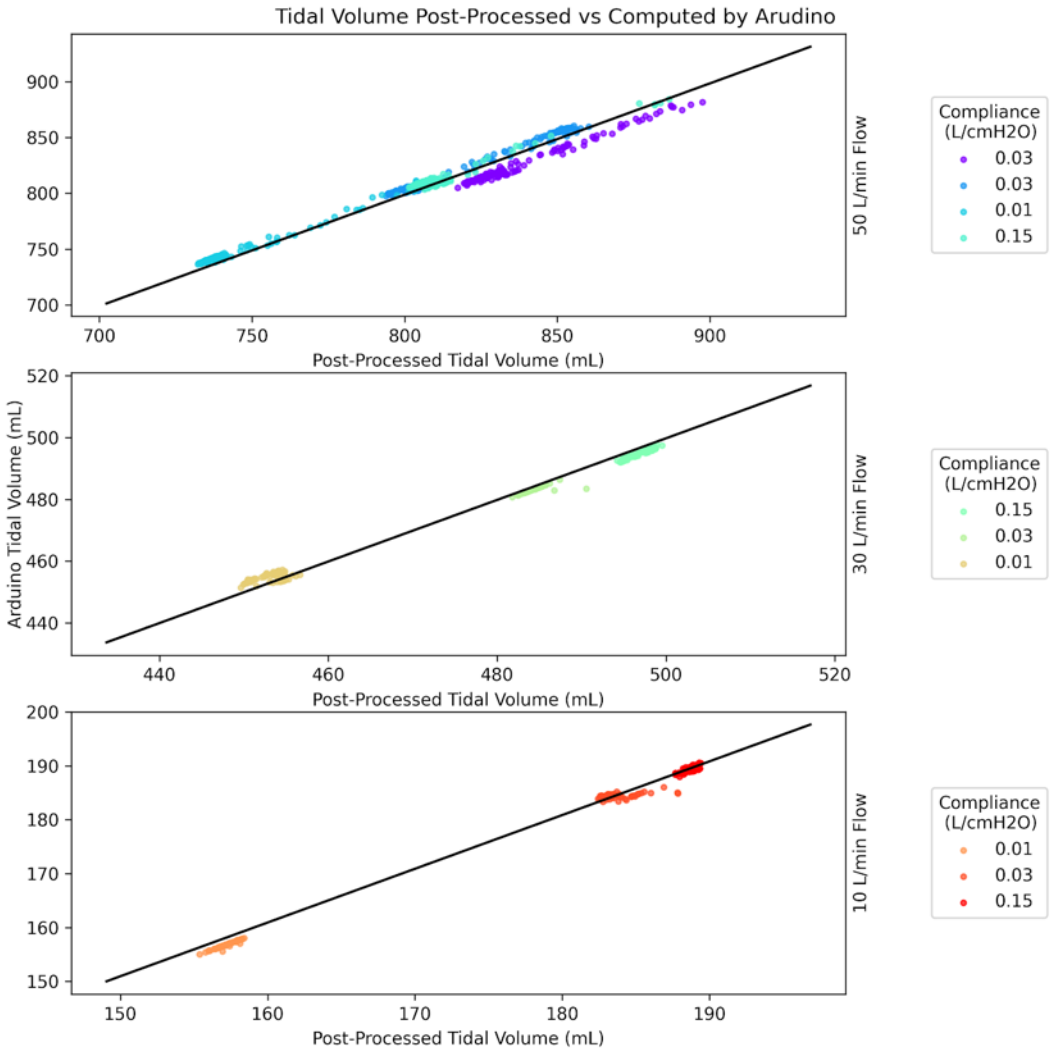

Figure S3: Tidal volumes computed by offline numeric integration vs. tidal volumes computed using online thresholding algorithm. Graphs are for three different flows (50L/min, 30L/min, and 10L/min).

Appendix B

A detour with electromagnetics simulation and software

B.1 Introduction

In about the middle of my Ph.D., before I pivoted to working on piezoelectric-based optical modulators, I briefly investigated the use of photonic crystals for optical voltage detection. However, I almost immediately ran into a stumbling block - at the time, there was very little software available to simulate the reflectance spectra photonic crystals. The software that did exist at the time was either very difficult to use and install, requiring compilation from source (such as S4 [87]), or did not provide the desired functionality. Arguably the best technique for studying these structures is one called Rigorous Coupled Wave Analysis (RCWA).

To better understand this method, and to make the lives of future researchers easier, I decided to write my own software package to implement rigorous coupled wave analysis. I wanted to write the software in pure python to enable cross-platform operation and ease of use and install, and to further my own python development skills.

B.2 Theory of the Transfer Matrix Method

Fundamentally, all the Transfer Matrix Method does is solve Maxwell's Equations for a particular set of cases. It is extremely efficient at doing this, but only works for a small subset of problems.

That subset includes all problems which have continuous symmetry along two dimensions - infinite in in x and y and have a finite number of layers with arbitrary thicknesses and material properties along z (i.e. a thin film of SiO_2 resting on an Si substrate).

The mathematics behind the transfer matrix method are described extensively here [66].

B.3 Theory of Rigorous Coupled Wave Analysis

The math behind rigorous coupled wave analysis is described here [96], and its implementation in software in terms of scattering matrices is described here [122].

B.4 Design Goals

I had a few main goals in writing this software

- Ease of installation - a simple `pip install rcwa` should work.
- Ease of use - Only take a few lines of code to run a simulation
- Trustworthiness - Be heavily tested with known-correct data
- Cross-Platform - Work on all major platforms that support python
- Object-oriented - Clear representation of simulation ideas by objects
- Easy to modify - Auto-deploy to `PyPi` when changes are made

B.5 Unit Testing

One fundamental problem of writing academic software as a one-off script is that it is difficult to trust and modify. Usually, researchers will poke and tweak a script until it gives them the output they are expecting in a particular case, or set of cases. However, if (god forbid) these scripts need to be re-used at a later date or (even worse) modified to have different functionality, this can be very difficult to do. Modifying a script, more often than not, breaks it, and sometimes the researcher in question may not be able to fix their script. While version control, such as git, allows researchers to revert a script or package to a previous version if they cannot find the source of the error, it does not solve the problem of how to easily modify a script - enter unit tests.

Unit tests formalize the process of answering the question "does this script do what I am expecting?". They are fundamentally very simple - they are a set of scripts that are run which *test* different functionality of *units* of the script or software in question. Typically, they will apply known inputs to a function, and run that function to see if the desired output is returned. An example of a unit test for a piece of electromagnetics software is ensuring that when you pass a plane wave through an interface which has the same refractive index on either side, no reflection should occur - the transmittivity (T) of that system should be 1 and the reflectivity (R) should be 0. If this simple test fails, you can definitively say the software is producing a nonphysical result. Even with a small number of such tests, you can be reasonably sure that the software is working. With a large number of tests covering many different cases, you can be increasingly confident that the software can be trusted. Even

better, if you want to make a change to the software in the future or add functionality, you can make the change and then re-run the tests. If the tests pass, the added functionality at least did not break anything you knew was working. If it did break something, you can quickly identify what broke and fix the issue. Without unit tests, nonphysical results can be difficult to spot, much less fix. If additional functionality is added, however, additional tests should also be added verifying that the new functionality works as expected. Unit tests are a way of demonstrating, via induction, that software works as expected.

After building up a set of tests (often called a "test suite") these can serve as a proxy for whether software is ready to use. If the tests passed, you can be confident in using the software. When writing this package, I have written (to date) 160 tests covering several hundred test cases, including reflection off interfaces, thin films, multi-layer stacks, diffraction gratings, and photonic crystals, using previously-published and known-correct data. I originally wrote these tests using the python `unittest` module. However, once I began to recognize the limitations of this software, especially in regards to parameterization and the need to manually add additional tests to the suite of those to be run, I switched over to using `pytest`, and would recommend this framework to all academic developers. It's easier to use, requires less boilerplate code, auto-detects tests, produces more readable code, and has much more flexible parameterization. It also can run unit tests already written in the `unittest`. About half of the tests currently are still written with `unittest` and half with `pytest`.

B.6 Continuous Deployment

Another major downfall of writing one-off scripts is that they are difficult to share. While I can certainly e-mail my script to other researchers (who doesn't love receiving someone else's MATLAB script and then trying to get it to work). When using python, this can be fixed by ensuring the software can be run on a remote *build server* and then *deploying* the software to a package index such as the Python Package index (PyPi). This package index comes with a built-in package mananger called `pip`, so that anyone who has python installed can simply run `pip install desiredPackage`. For the unfortunate souls (myself included) running windows, the command will be `python -m pip install desiredPackage`. Figuring out how to deploy packages was (for me) pretty difficult to do, and still represents a high bar to clear for researchers who are not developers. However, once you do it once, you can set up a repository on Github (or other build servers) to automatically re-deploy the software every time you make a change that passes your test suite. This process is referred to as "continuous deployment".

I did this with my RCWA package, following the instructions here, using GitHub's built-in GitHub Actions. Now, each time I push to the `master` branch of the RCWA repository, Github Actions spins up a virtual machine, installs dependencies, and runs the full test suite. If the test suite passes, GitHub Actions then re-deploys the RCWA package to the python package index so others can install the latest version. That software is then downloaded

from the python package index using `pip` and the tests are re-run, verifying that installation works. All this is done by a simple `git push` command on my local machine. Additionally, I created what is called a "hook" on my local machine to cause the "git commit" command to also run the test suite locally. This means that no commits can even be made to my local repository that causes my unit tests to fail. While I sometimes have a legitimate need to get around this, it's been an overwhelmingly positive practice I would highly recommend.

B.7 Software Architecture and Use

The latest version of the documentation for my RCWA package can be found here. Here I will discuss how it was designed as of when I am writing this in August 2022, and what changes might happen in the future.

Guiding Principles

1. Provide reasonable defaults
2. Keep it Simple Stupid
3. Represent physical objects as software objects
4. Represent physical processes as object methods or functions

Design Architecture

While there are many different schools of programming (procedural, functional, and object-oriented to name a few), object-oriented programming is unreasonably effective at representing things in the physical world, precisely because the physical world is made up of objects. For example, a plane wave is typically thought of as an "object" (though we don't typically use that term as physicists). A thin film is also an "object", and a semi-infinite region is also an "object". crystal lattices are "objects", diffraction gratings are "objects", "materials" are objects, and so on. In my RCWA package there are two fundamental objects:

- `Source`, used to represent plane waves of different wavelengths and angles of incidence
- `Layer`, used to represent a region of space which is infinite in x and y and finite in z.

The `Layer` object itself has properties which are represented by their own objects.

- `Material`, used to represent material properties of a `Layer`
- `Crystal`, used to represent a periodic lattice of a particular `Layer`

Layers have other properties as well which are scalars rather than objects, including a thickness (which is zero for semi-infinite layers).

Finally, these layers can be linked together in a particular order by the `LayerStack` object, which fully describes the simulation domain.

At this point, a simulation can be fully described, and the remainder of the code could be procedural, and describe how these objects interact with each other. However, in the interest of consistency and ease of use, I decided to make the interaction itself object-orinted, by introducing the `Solver` object. This object links together the `Source` and `LayerStack` objects and by running `Solver.solve()`, the simulation is run and the results can be obtained. A bare-bones simulation is shown below describing reflection off an interface from a material with n=1 to a material with n=2:

```
from rcwa import Layer, LayerStack, Source, Solver

source = Source(wavelength=1)

incident_layer = Layer(n=1)
transmission_layer = Layer(n=2)

stack = LayerStack(incident_layer=incident_layer,
    transmission_layer=transmission_layer)

solver = Solver(stack, source)
results = solver.solve()
```

The user-facing API is simple, intuitive, represents physical objects as objects in software, and provides reasonable defaults. I have used this software extensively for modeling thin-film based devices, for calculating reflection spectra of thin films and resonant cavities, including highly lossy materials and metals. I have found the unpatterned results to be in good agreement with physical reality, though I have yet to personally test the results from modeling photonic crystals against the real world.

Appendix C

A practical tutorial using JMP for Statistically Designed Experiments

C.1 Introduction

Often in my Ph.D., I wanted to obtain a film with certain properties, such as roughness, stress, or optical constants. Or I wanted an etch with a particular sidewall profile. Approaching this from a first-principles approach is rarely possible - the systems involved are so complex that any attempts to model them will usually consume a great deal of time while yielding little - or will require an entire Ph.D. worth of work just to understand a process under a single set of conditions. The typical response to this is to throw up one's hands after many hours of fruitless labor, or to tweak different variables one at a time semi-randomly until you get something to work. Fortunately, there is a better way.

In the early 1900's a man named Ronald Fisher encountered a very similar problem: trying to maximize crop yields at the Rothamsted Experimental Station. There are many variables to work with when growing crops, and they can interact with each other in complex ways. Below is an example of what the traditional approach of optimizing variables one-by-one looks like.

C.2 An Example: Growing Cacti

Suppose Fisher was growing a cactus. From previous generations of cactus-growers, he knows that a moderate amount of sunlight and a moderate amount of water can produce cacti with a decent growth rate. But he wants to grow bigger, juicier cacti, and so decides to vary the water and sunlight given to the plant. First, he tries varying the water given to the plant, and tries a low, moderate, and high amount of water. I'll denote these as (-), (0), and (+) for low, moderate, and high. He finds, to his delight, that at lower water amounts the cacti grow a little faster! We can draw his findings in a diagram. Lower numbers mean

less growth, which I have drawn in more reddish color. Higher numbers mean more growth, which I have draws in green:

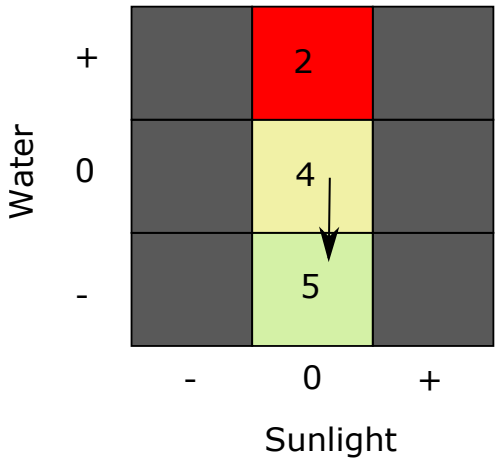

Figure S1: Experimental matrix after running step to optimize water.

Now, he tries varying the sunlight, while setting the water at its optimal (low) value. He then finds that the best growing condition is low sunlight, and believes he is done.

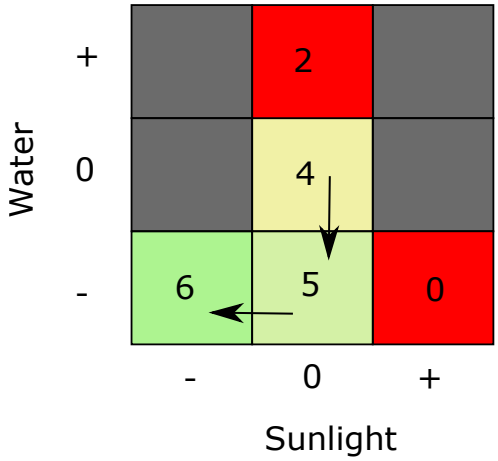

Figure S2: Experimental matrix after running step to optimize sunlight.

This type of experimental design is referred to as One Factor At a Time (OFAT), because Fisher changed one variable (factor) at a time. However, something looks funny about this diagram of experiments. It's got some holes in it - some experiments Fisher never ran (the boxes colored gray). In particular, Fisher never checked to see what would happen at high

levels of sunlight and water. Suppose he does the rest of the experiments, filling out what remains of the grid:

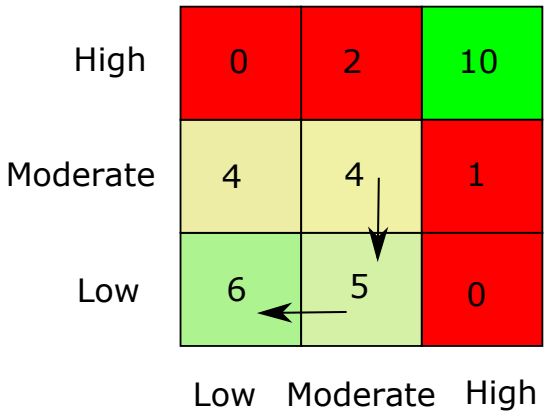

Figure S3: Experimental matrix after filling in the missing experiments.

Aha! It looks like at high values of sunlight and high water, Fisher's cactus grows dramatically better than in any of our previous experiments. These variables "interact" with each other - it's not that high sunlight is beneficial on its own, but if the plant has high sunlight, then it can take advantage of high amounts of water without drowning. This is what's called a "two-factor interaction", because it involves a dependence on two factors together. Simply running one variable at a time, Fisher missed this completely, and would have caused England to starve.

This interaction among variables tends to be the rule rather than the exception in complex systems. By systematically exploring the set of possible experiments we can do, we can identify complex interactions among variables. The type of experimental design shown in this example is called a "three-level, two-factor, full factorial" design. "three-level" because we attempted three different "levels" of each variable: low (-), moderate (0), and high (+). "two-factor" because we had two different factors: water and sunlight. "Full factorial" means we tried all possible combination of these two factors.

C.3 Factorial Explosion and Sparsity of Effects

However, typically in nanofabrication we don't just have two variables, we might have six or seven. For example, when I was attempting to optimize the sidewall angle of aluminum nitride etch in *centura-met*, I had seven factors to adjust: Plasma power, bias power, pressure, Cl_2 flow, H_2 flow BCl_3 flow, and Ar flow. If I were to try a two-level seven-factor design it would require $2^7 = 128$ separate experiments. I don't have time for that. I'm trying to get my Ph.D. before the end of the century.

Fortunately, we can make one incredibly powerful assumption which makes this more tractable. We can assume that there are no "high-order-interactions". For example, plasma power and pressure might interact with each other, and plasma power and Cl_2 flow might interact with each other, but plasma power *and* pressure *and* Cl_2 flow probably don't interact with each other simultaneously. In other words, increasing Cl_2 flow might do one thing at a low power and a different thing at a high power, but changing the bias power probably doesn't affect how different powers affect what Cl_2 does. In practice, this is usually the case. Third-order interactions may occur, but they are rare and usually not the dominant effects in a system. Fourth-order interactions are even rarer, and fifth-order rarer still. Typically it is sufficient to only be able to estimate second-order interactions.

If we assume that there are no third-order or higher interactions, and we want to be able to *estimate all possible* single- and two-factor interactions with reasonable accuracy, then for a two-level 7-factor experiment we actually only need 29 runs. There will be some what is called "partial confounding", but, in practice, usually this is not an issue, and I will not discuss it further here. For a reference see [135]. If we are even more aggressive, and assume that there are no two-factor interactions at all, then we only need to perform 8 runs. That's a whole lot less than 128. For the mathematically inclined, this is equivalent to fitting the data to a 7-dimensional hyperplane.

C.4 A Real-World Example: Optimizing an Aluminum Nitride Etch for Sidewall Angle in centura-met

Recently I needed to develop an aluminum nitride etch to have as close to vertical sidewalls as possible. The staff-suggested starting point struck a plasma, but gave me a sidewall angle of roughly 30 degrees - not great. To whet your appetite for what follows, I was able to achieve an 84 degree sidewall angle within 2 days of starting the experiment, with no literature review and minimal guidance from staff (in particular - which gases might affect the etch). I could have optimized it further but was content with the results after the second round of optimization. I should note I do not have a strong chemistry background. Here I will walk through the full process of setting up, running, and analyzing this experiment.

As discussed previously, this particular tool has seven independent knobs I was able to turn. This is a lot of knobs. Because of this, and because I didn't have any data to work from previously, I decided to conduct a two-level, main-effects-only experiment. This is usually a good starting point with a large number of factors (more than three). The goal here is not to try and fit an accurate model, but to see which factors have any effect on the process at all. In JMP 16, this is under the "DOE" Menu → "Classical" → "Two Level Screening" → "Screening Design". This type of experiment is a good choice when you don't have much information to go on and each "run" is relatively expensive.

The most subjective and trickiest part was the next step: choosing ranges over which

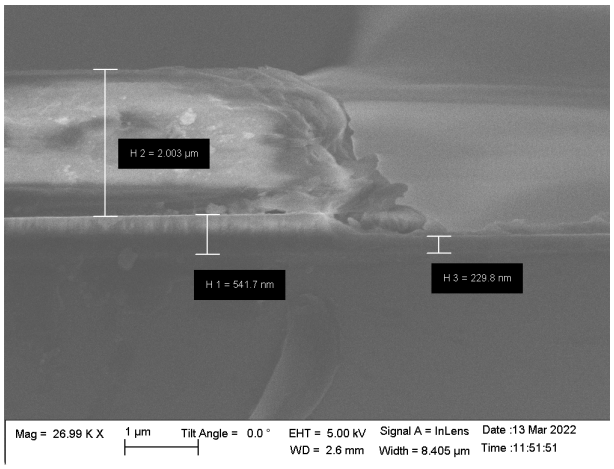

Figure S4: First attempt at an Alimunum Nitride Etch. Approximate sidewall angle was 30 degrees.

to vary each variable. I like to follow the wise words of Queen Latifah in this regard - "Be Bold." Typically, I allow practical considerations to constrain the minimum and maximum value of each variable. For example, you cannot flow more gas than the flow controller is rated for. For reactant gases, obviously a flow of zero is probably not a great idea. Pressure and power are often bounded from below by the need to sustain a plasma, and so on. Smaller increments have the advantage of giving a higher-resolution picture of what is going on - for example power might initially improve the sidewall angle, then start to make it worse if it gets too high. Typically, however, working with larger increments gives a greater chance you will find something interesting, even if for the wrong reason, and I tend to favor this approach.

Choosing the Power Range

I found that I was able to sustain a plasma with at least 400W of power at a pressures between 8 and 14mtorr, so I set 400W as my minimum pressure value. The tool power only goes up to 1kW, so I elected to use 800W as my upper bound. In general, I think it is good practice to set your maximum value a little below what the tool can handle in a screening experiment. If you find that a higher power is beneficial, you can always increase it further in a follow-up experiment.

Choosing the Pressure Range

I found it difficult to sustain a plasma consistently below 8mtorr, so I chose this as my minimum pressure value. At pressures above 14mtorr, the pressure control valve was almost

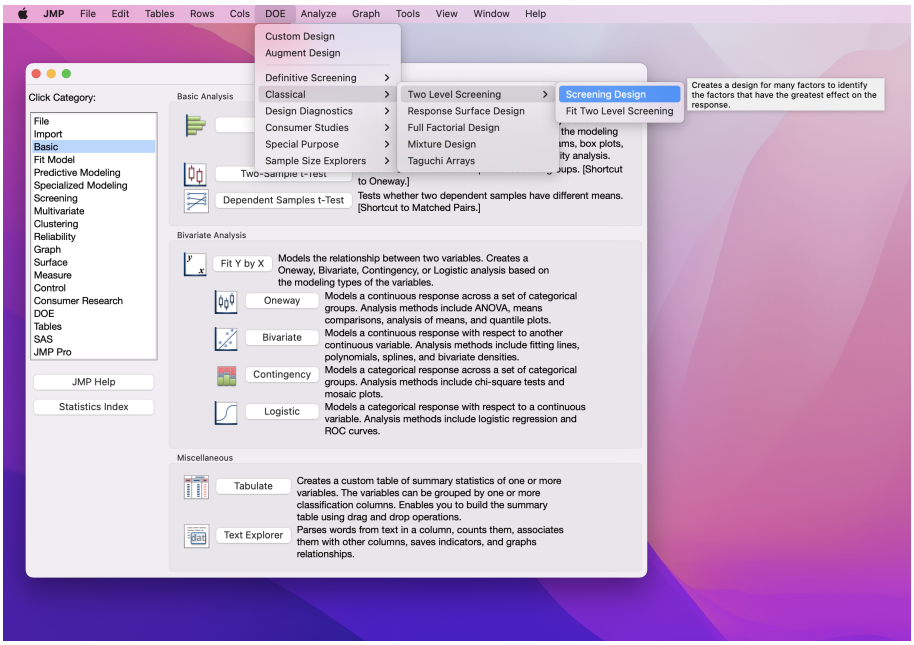

Figure S5: Menu in JMP 16 to set up a two-level screening design.

fully closed, and the pressure started to become more difficult to control, so I chose this as my upper bound.

Choosing the bias power

This was also relatively easy - the tool was not able to provide stable bias powers below 50W, and its maximum value was 100W, so I chose the range of 50-100W as my range. If I were to follow my own advice, I might have chosen 50-90W instead.

Choosing Ar Flow

In order for this particular equipment to control the pressure in the range I wanted, I found a total flow of 200-300sccm worked well. So, instead of varying the Ar flow directly, I created a different variable called "Total Flow", which I varied between 200 and 300sccm.

Choosing H2 Flow

The MFC on the H_2 only went up to 50sccm, and I believed that at least some H_2 would be needed, so I somewhat arbitrarily set the range to 20-50sccm. This was among the least-informed of my decisions.

Choosing BCl_3 flow

Staff suggested that only a small amount of BCl_3 was necessary to break through the native oxide formed on aliminum nitride. The minimum amount the flow controller could reliably provide was 5sccm, so I set this as my lower bound. The flow controller went up to 100sccm, but because I expected only a small amount to be needed, I set the upper limit at 25sccm.

Choosing Cl_2 flow

This choice was almost completely arbitrary. I noticed that the flow controller went up to 200sccm, so I set my upper bound at 150sccm. At least some Cl_2 would be necessary to drive the reaction, as this is a Cl_2 based etch, so I arbitrarily set my minimum flow at 50sccm.

Design setup in JMP

After choosing all my factor ranges using the above reasoning, I added Once I selected the "Screening Design", I added seven continuous factors (by hitting the "Continuous" button seven times), named them, and added their minimum-maximum ranges.

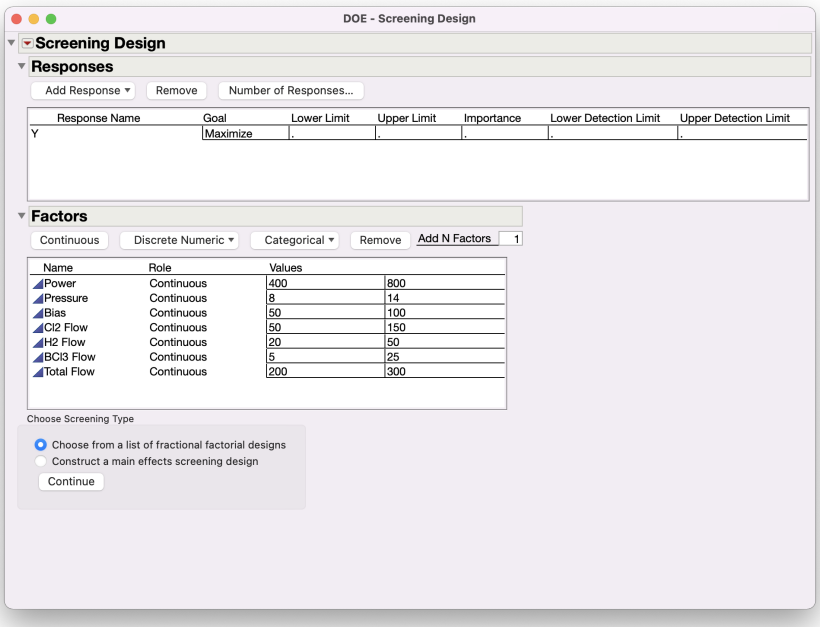

Figure S6: All seven factors listed with their minimum and maximum values.

Then I hit the "Continue" button underneath the factor list and selected "Choose from a list of fractional factorial designs". JMP gives us a list of possible experiments it can

generate, and the main way in which they differ is how well they can estimate multi-factor interactions. More runs usually leads to a better ability to estimate multi-factor interactions. Because I could prepare wafers in batches of 6, and I wanted to give myself room to screw up, I chose the 8-run "Main Effects only" fractional factorial experiment, and hit "Continue".

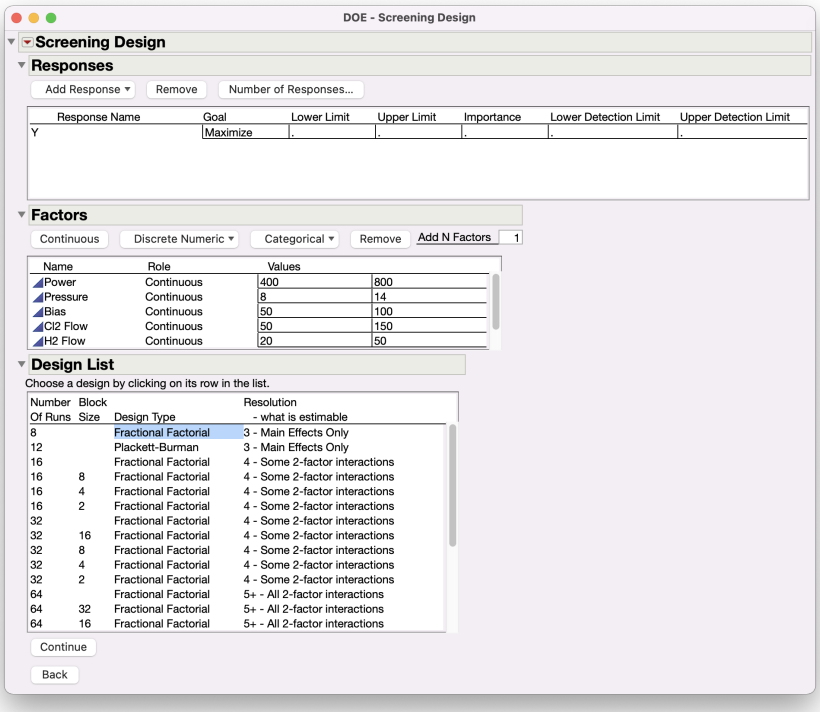

Figure S7: All seven factors listed with their minimum and maximum values.

At this point, we are almost done designing the experiment. By default, JMP will randomize the run order. *This is absolutely critical.* Randomizing run order ensures that you do confuse tool drift for the effect of one of your variables - if I were to start at a low power and increase to higher and higher powers in later runs, I would not be able to tell apart the effect of power and the effect of tool drift. Randomizing run order "smears out" the effect of tool drift onto all variables, turning drift into noise. In this context, some noise is better than being completely unable to estimate the effect of one of your factors.

I also always choose to run a center point - which will add a considerable amount of information. Typically, I run two center points to see how much the tool drifts, or what the noise in my process conditions is. In this particular experiment, I only ran a single center point, with no replicates. JMP

To generate the design table I then hit "Make Table", and JMP generates a table which looks something like this:

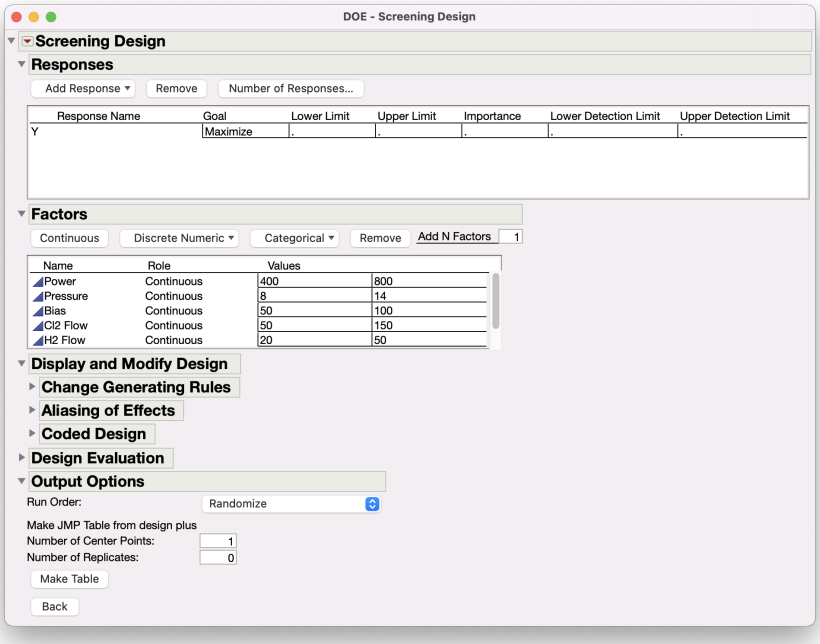

Figure S8: All seven factors listed with their minimum and maximum values.

Once the screening design was complete, I needed to generate the Ar flow variable, as the equipment doesn't actually have a field for "total flow". To do this I right-clicked on the column labeled "Y" and hit "Insert Column":

I then right-clicked on the top of the column that said "Column 9" and selected "Formula", then created a formula for Ar flow (Total Flow - all the other gases). Once I did this, I noticed some of the values for Ar flow were negative. Doh! That's clearly not possible. Fortunately, the mistake wasn't egregious. The total flow variable would only need to be slightly modified to fix this. In the cases that would require a negative argon flow, I increased the total flow so that it would be zero. This will make modeling of the "total flow" variable a little less optimal, but it shouldn't dramatically affect the design. This type of thing often happens - you may have one or two conditions, for example, where you need to slightly increase or decrease one variable to satisfy physical constraints or the constraints of the tool after making your design. That is OK and often unavoidable. It's not *optimal*, but it will only make your effects harder to estimate, not impossible.

Run the Experiment, Analyze the Data

For all the experiments I ran, I ran the etch for 180s. I prepared the wafers with 500nm of AlN deposited in aln2. I expected my etch rate not to vary by more than a factor of 2

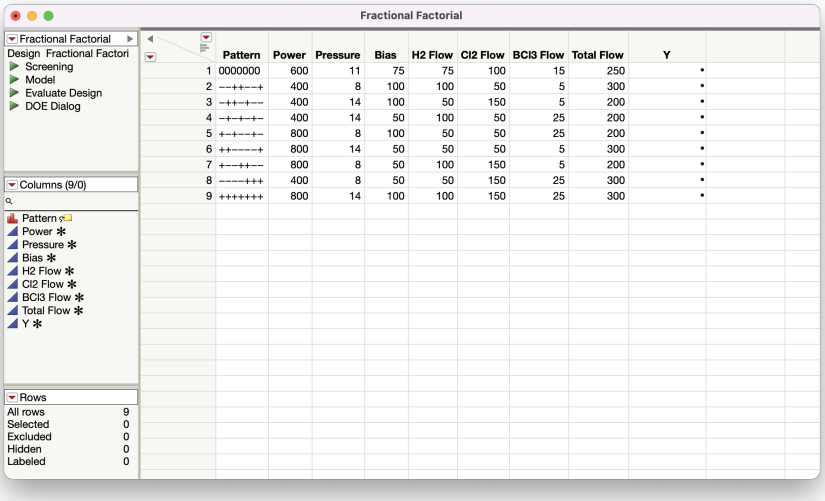

Figure S9: Randomized main-effects screening design with seven factors, two levels, and one center point.

or so based on prior experience with chlorine etches, and 180s this would allow me to get through enough material for a meaningful cross-sectional SEM, without breaking through the entire film. I took some cross-sectional SEM's and measured the following sidewall angles (in degrees): 24, 67.5, 48, 60, 50, 62, 64.3, 63, 50.9. Looking at that list of numbers should excite you, because there is substantial variance. While none of the values were close to my target of 90 degrees, there was considerable variability, which means one (or more) of my variables is doing something interesting, and we have something to work with. The *single worst* thing that can happen in a statistically designed experiment is when there is no variance in the data.

Now that I had my data, I fitted a model in JMP by clicking "Analyze" → "Fit Model".

Since I designed a "main effects" design, I am only able to fit the "main effects" - the linear, independent effects of each variable. I added "sidewall angle" to the "Role Variables" section by selecting it and then clicking the "Y" button. Then I selected all of the variables that I wanted to fit (everything except Ar Flow) and pressed the "Add" button in the "Construct Model Effects" area.

Once I hit the "Run" button, the following screen shows up:

The most important figure on this screen is the "Actual by Predicted Plot". This plots your actual data versus the predicted data from a linear model with the variables we selected previously. It looks like a straight line - but the model variance (the shaded red area) is huge. The next step is eliminating terms from this model until you get something that looks more reasonable. One way to do this is called "backwards stepwise elimination" - successively

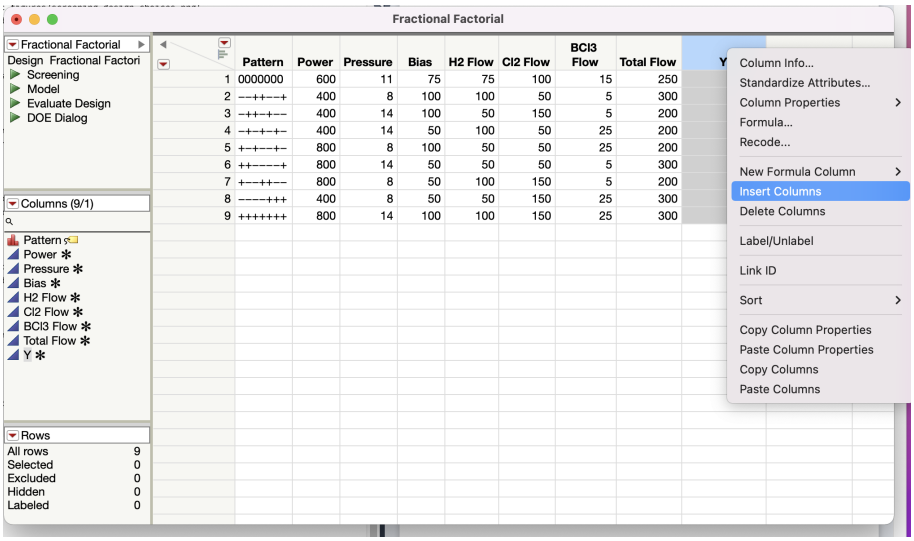

Figure S10: Randomized main-effects screening design with seven factors, two levels, and one center point.

removing effects until all the ones that remain are significant. I usually start with the least-significant effect (the one with the highest P-value". I selected it, and then hit "Remove" to remove it. I kept doing this until I was satisfied. Typically, I will use a threshold p-value of 0.1-0.2 for leaving in effects for an initial screening experiment, and discard the rest.

Analyzing the Model

With that in mind, I eliminated variables until I was left with only four: Pressure, Bias, H_2 flow, and total flow. All four appear to be important. Next, I like to play around with the "Prediction Profiler" (select "Factor Profiling" $\rightarrow$ "Profiler" if it doesn't show up on the little red triangle near the top). It looks like a high H_2 flow, high total flow, high bias, and low pressure give the best sidewall angle.

Of the four variables, the two which appear to have the strongest effect *and* the most statistically significant. effect are pressure and bias. Unfortunately, high total flows prevent achieving low pressures, so in practice I will choose the highest possible total flow that can give me a lower pressure, so this won't be used as an explicit factor in follow-up experiments. Second, since H_2 flow was already maxed out and I cannot increase it further, I will drop it as a factor as well and set it in the future. That leaves me with two factors to investigate further: pressure and bias. While not significant, lower flows of BCl_3 improved sidewall angle. Increasing power and increasing Cl_2 flow did not have any significant effect on the sidewall angle, but substantially increased the etch rate. In my case, I wanted a slower etch rate, so I decided to use a lower amount of Cl_2. I would have used a lower power as well,

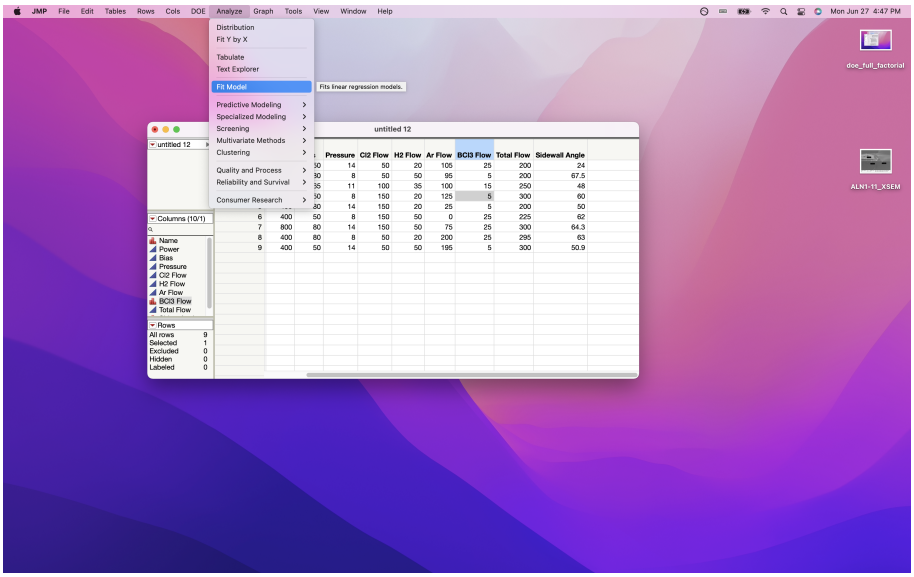

Figure S11: Standard model-fitting menu.

but decided against it for reasons you will see in the next section.

Designing a Follow-Up Experiment

After playing around with the pressure, power, and flow a little more I found I was able to sustain a plasma at 800W and 5mtorr, and the maximum total flow I could use in this case was 175sccm. So in the follow-up experiment I would set my power to 800W, total flow to 175sccm, H_2 flow to 50sccm, BCl_3 flow to 5sccm, Cl_2 to 50sccm and Ar flow to 70sccm. Since I was left with only two variables to adjust, I decided to use what is called a "Response surface" design. This allows more complex (quadratic) models to be fit to the data, instead of just linear models, and allows for two-factor interactions to be estimated. I could have done a full factorial experiment and got similar information, but a response surface experiment turns out to have nicer statistical properties. This can be found under "DOE" $\rightarrow$ "Classical" $\rightarrow$ "Response Surface"

The two variables I want to vary are the bias and pressure.

Choosing the range for Pressure

While I want to go as low as possible, the lowest I was able to sustain a plasma was 5mtorr, so I chose this as my lower bound. For my upper bound, I decided to choose 11mtorr. Lower than my previous upper bound, but would give some overlap with my previous experiment.

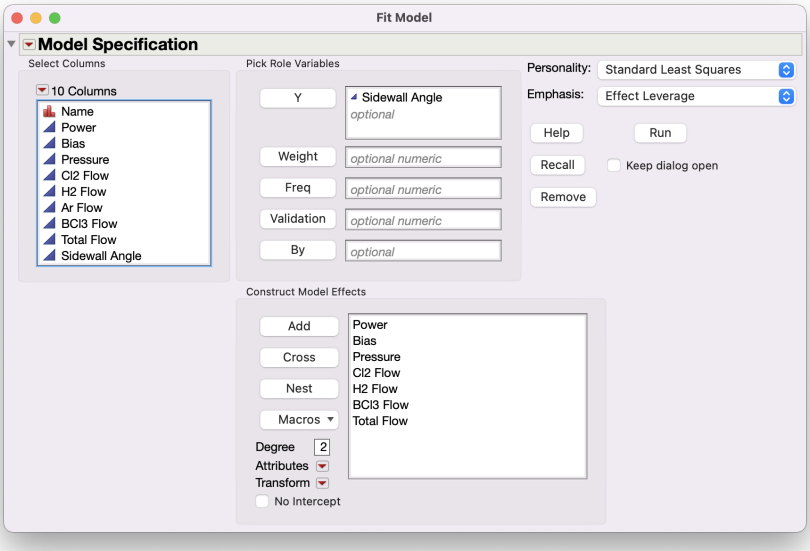

Figure S12: Standard model-fitting menu.

Choosing the range for bias

This was also easy an set by practical constraints - the tool could not go higher than 100W bias, and I want to go as high as possible, so my upper bound was 100W. For some reason I cannot remember, I decided to choose 70W as my lower bound.

Creating and running the experiment

I chose the "Central composite Design", because I wasn't interested in running more wafers than absolutely necessary. As before, I chose a single center point and a randomized run order, generated the table with the "Make Table" button, and then ran the experiment in the order specified by JMP and collected the data. This time, the values for sidewall angle I obtained were (in degrees): 61, 29.6, 83.4, 83.9, 77.7, 63.0, 73.1, 83.2, 78.3. This was much better than my previous run, and tells me that I am on the right track.

In fact, at this point I was satisfied with the results - I didn't even bother going through the process of fitting a model to the data, and just directly used the recipe that got me a sidewall angle of 83.9 degrees. All told, it took a few hours checking which conditions would strike a plasma, a few hours to prepare the 18 total wafers, a few hours to analyze the data, and about 6 hours to make the recipes and run the actual experiment - in total, about 2 days.

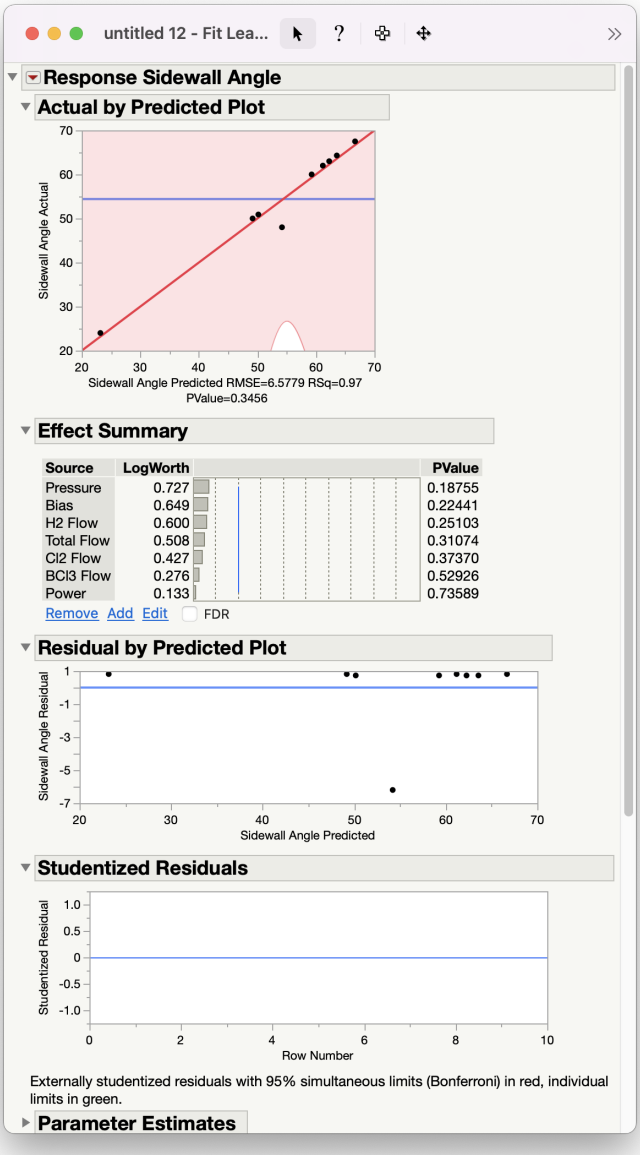

Figure S13: Initial model after hitting the "Run" button.

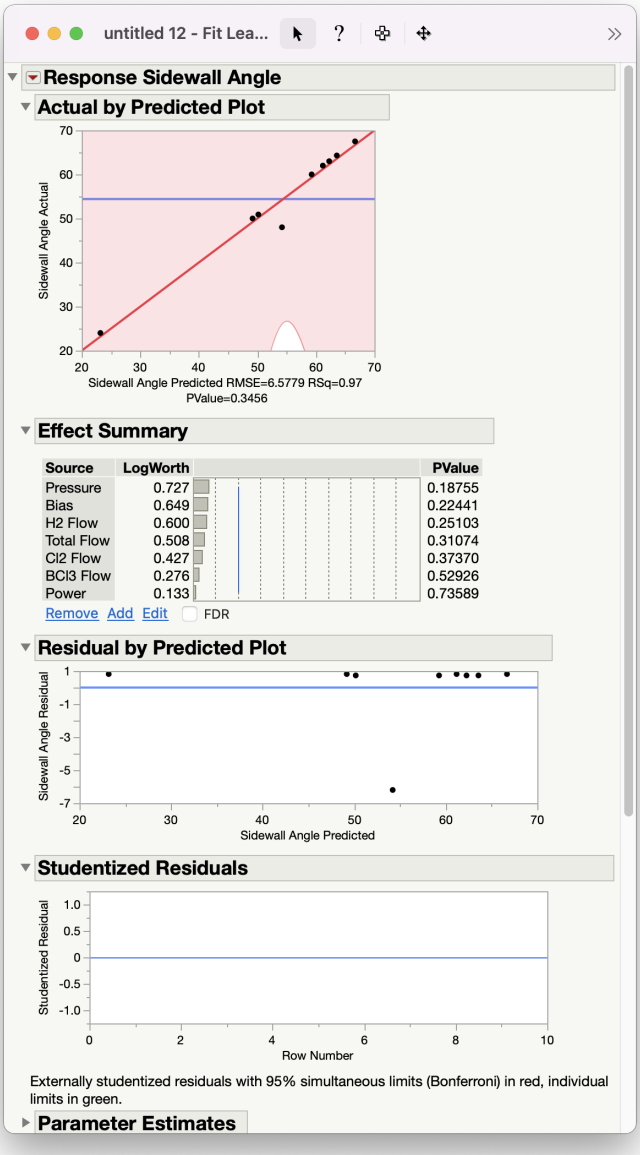

Figure S14: Initial model after hitting the "Run" button.

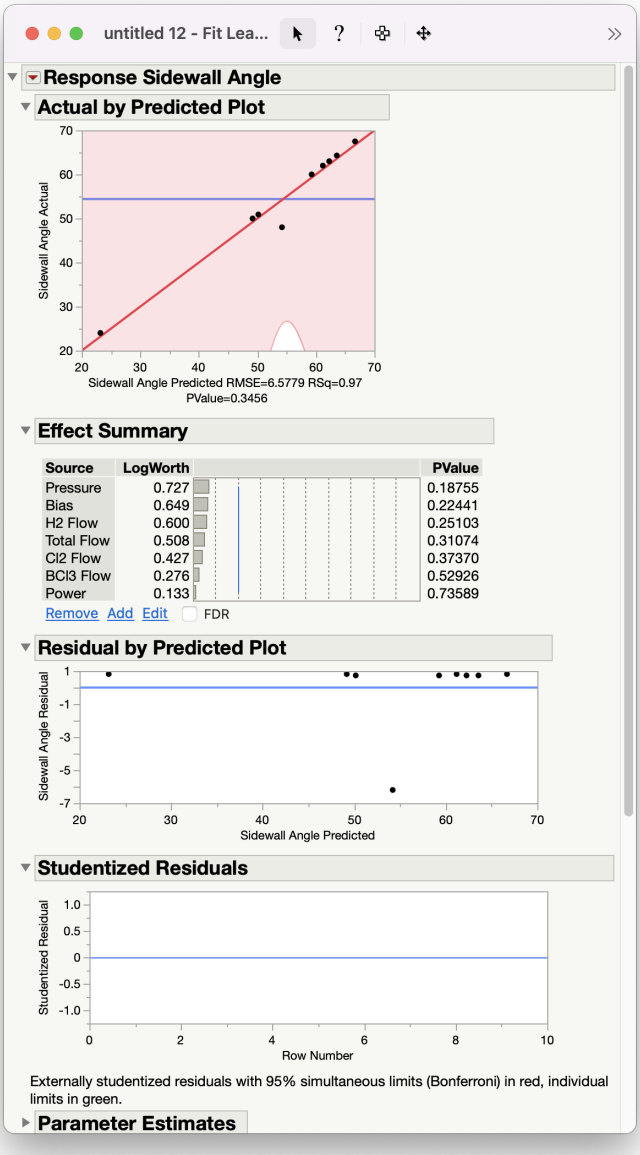

Figure S15: Initial model after hitting the "Run" button.

Bibliography

[1] D. Michael Ackermann et al. "Designing the Optical Interface of a Transcutaneous Optical Telemetry Link". In: *IEEE Transactions on Biomedical Engineering* 55.4 (2008), pp. 1365–1373.

[2] Taner Akkin et al. "Detection of neural activity using phase-sensitive optical low-coherence reflectometry". In: *Optics Express* 12.11 (2004), pp. 2377–2386.

[3] *Application Note for SDP3x Series differential pressure sensors Compensation of pressure drop in a hose.*

[4] Sahar Ayazian and Arjang Hassibi. "Delivering optical power to subcutaneous implanted devices". In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 2874–2877.

[5] Claudio Babiloni et al. "Fundamentals of electroencefalography, magnetoencefalography, and functional magnetic resonance imaging". In: *International review of neurobiology* 86 (2009), pp. 67–80.

[6] Bradley J Baker et al. "Imaging brain activity with voltage-and calcium-sensitive dyes". In: *Cellular and molecular neurobiology* 25.2 (2005), pp. 245–282.

[7] Yuki Bando et al. "Comparative evaluation of genetically encoded voltage indicators". In: *Cell reports* 26.3 (2019), pp. 802–813.

[8] Mark W Barnett and Philip M Larkman. "The action potential". In: *Practical neurology* 7.3 (2007), pp. 192–197.

[9] Christoph Baumgartner and Johannes P Koren. "Seizure detection using scalp-EEG". In: *Epilepsia* 59 (2018), pp. 14–22.

[10] Farid Belhachemi, Stéphane Rael, and Bernard Davat. "A physical based model of power electric double-layer supercapacitors". In: *Conference Record of the 2000 IEEE Industry Applications Conference. Thirty-Fifth IAS Annual Meeting and World Conference on Industrial Applications of Electrical Energy (Cat. No. 00CH37129)*. Vol. 5. IEEE. 2000, pp. 3069–3076.

[11] Corbett Bennett, Sergio Arroyo, and Shaul Hestrin. "Subthreshold mechanisms underlying state-dependent modulation of visual responses". In: *Neuron* 80.2 (2013), pp. 350–357.

[12] Jean M Bennett and EJ Ashley. "Infrared reflectance and emittance of silver and gold evaporated in ultrahigh vacuum". In: *Applied Optics* 4.2 (1965), pp. 221–224.

[13] William Biederman et al. "A Fully-Integrated, Miniaturized (0.125 mm^2) 10.5 µW Wireless Neural Sensor". In: *IEEE Journal of Solid-State Circuits* 48.4 (2013), pp. 960–970.

[14] J Blake, P Tantaswadi, and RT De Carvalho. "In-line Sagnac interferometer current sensor". In: *IEEE Transactions on Power Delivery* 11.1 (1996), pp. 116–121.

[15] Joseph A Boales, Shyamsunder Erramilli, and Pritiraj Mohanty. "Measurement of nonlinear piezoelectric coefficients using a micromechanical resonator". In: *Applied Physics Letters* 113.8 (2018), p. 083501.

[16] KSA Butcher and TL Tansley. "Ultrahigh resistivity aluminum nitride grown on mercury cadmium telluride". In: *Journal of Applied Physics* 90.12 (2001), pp. 6217–6221.

[17] J Buytaert et al. "The HEV Ventilator". In: *arXiv preprint arXiv:2007.12012* (2020).

[18] Papichaya Chaisakul et al. "Recent progress on Ge/SiGe quantum well optical modulators, detectors, and emitters for optical interconnects". In: *Photonics*. Vol. 6. 1. MDPI. 2019, p. 24.

[19] Raymond Chang et al. "COVID-19 ICU and mechanical ventilation patient characteristics and outcomes-A systematic review and meta-analysis". In: *medRxiv* (2020).

[20] Ting Chia Chang et al. "End-to-End Design of Efficient Ultrasonic Power Links for Scaling Towards Submillimeter Implantable Receivers". In: *IEEE Transactions on Biomedical Circuits and Systems* 12.5 (2018), pp. 1100–1111.

[21] Ritchie Chen, Andres Canales, and Polina Anikeeva. "Neural recording and modulation technologies". In: *Nature Reviews Materials* 2.2 (2017), pp. 1–16.

[22] Minkyu Cho et al. "Amorphous Si/SiO2 distributed Bragg reflectors with transfer printed single-crystalline Si nanomembranes". In: *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena* 34.4 (2016), p. 040601.

[23] Alejandro J. Cortese et al. "Microscopic sensors using optical wireless integrated circuits". In: *Proceedings of the National Academy of Sciences* 117.17 (2020), pp. 9173–9179.

[24] Alex Dante et al. "A temperature-independent interrogation technique for FBG sensors using monolithic multilayer piezoelectric actuators". In: *IEEE Transactions on Instrumentation and Measurement* 65.11 (2016), pp. 2476–2484.

[25] M Jamal Deen, John Ilowski, and Ping Yang. "Low frequency noise in polysilicon-emitter bipolar junction transistors". In: *Journal of applied physics* 77.12 (1995), pp. 6278–6288.

[26] Jean Delbeke, Sebastian Haesler, and Dimiter Prodanov. "Failure modes of implanted neural interfaces". In: *Neural Interface Engineering*. Springer, 2020, pp. 123–172.

[27] Ensheng Dong, Hongru Du, and Lauren Gardner. "An interactive web-based dashboard to track COVID-19 in real time". In: *The Lancet infectious diseases* 20.5 (2020), pp. 533–534.

[28] Wolfgang Drexler. "Ultrahigh-resolution optical coherence tomography". In: *Journal of biomedical optics* 9.1 (2004), pp. 47–74.

[29] Xiao-hong Du et al. "Crystal orientation dependence of piezoelectric properties of lead zirconate titanate near the morphotropic phase boundary". In: *Applied physics letters* 72.19 (1998), pp. 2421–2423.

[30] Yukun Duan and Bin Liu. "Recent advances of optical imaging in the second near-infrared window". In: *Advanced materials* 30.47 (2018), p. 1802394.

[31] Jordan Edmunds et al. *Dataset 1 - Modulation output spectra*. figshare (2021), `https://osapublishing.figshare.com/s/a300ca8f06d07eea104a`.

[32] Jordan Edmunds et al. *Dataset 2 - Modulation output at variable voltage*. figshare (2021), `https://osapublishing.figshare.com/s/24b3d04259a504ccf953`.

[33] Jordan Edmunds et al. *Dataset 3 - Modulated output for nonlinearity characterization*. figshare (2021), `https://osapublishing.figshare.com/s/989cb9c2190d738f1d75`.

[34] Jordan Edmunds et al. *Dataset 4 - Modulated output vs. temperature*. figshare (2021), `https://osapublishing.figshare.com/s/aef076d85f95fed1742a`.

[35] Jordan Edmunds et al. *Dataset 5 - Dielectric breakdown IV curve*. figshare (2021), `https://osapublishing.figshare.com/s/698240ba617312041952`.

[36] Jordan Edmunds et al. *Dataset 6 - Cumulative charge to failure*. figshare (2021), `https://osapublishing.figshare.com/s/f91b0f06f96fd1a8502e`.

[37] Jordan Edmunds et al. *Dataset 7 - Leakage current IV curves various temperatures*. figshare (2021), `https://osapublishing.figshare.com/s/db666d69d9674c56a358`.

[38] Jordan Edmunds et al. *Dataset 8 - Impedance spectra*. figshare (2021), `https://osapublishing.figshare.com/s/15592cfcacb02d7ee5a4`.

[39] Jordan Edmunds et al. *Dataset 9 - Measured reflectance spectra*. figshare (2021), `https://osapublishing.figshare.com/s/71a5ac7cbbdb2f4a7784`.

[40] Jordan Edmunds et al. "Low-Cost Differential Pressure Spirometry for Emergency Ventilator Tidal Volume Sensing". In: *medRxiv* (2020).

[41] Jordan L Edmunds et al. "Optical voltage sensor based on a piezoelectric thin film for grid applications". In: *Optics Express* 29.21 (2021), pp. 33716–33727.

[42] Lingling Fan, Zhixin Miao, and Miao Zhang. "Subcycle Overvoltage Dynamics in Solar PVs". In: *IEEE Transactions on Power Delivery* 36.3 (2020), pp. 1847–1858.

[43] US Food, Drug Administration, et al. "Enforcement Policy for Ventilators and Accessories and Other Respiratory Devices During the Coronavirus Disease 2019 (COVID-19) Public Health Emergency: Guidance for Industry and Food and Drug Administration Staff". In: *United States. Food and Drug Administration*. United States. Food and Drug Administration. 2020.

[44] Peng Fu et al. "Structure and electrical properties of (1- x)(Na0. 5Bi0. 5) 0.94 Ba0. 06TiO3–x BiAlO3 lead-free piezoelectric ceramics". In: *Materials & Design* 46 (2013), pp. 322–327.

[45] D Gerlich, SL Dole, and GA Slack. "Elastic properties of aluminum nitride". In: *Journal of Physics and Chemistry of Solids* 47.5 (1986), pp. 437–441.

[46] Mohammad Meraj Ghanbari et al. "A sub-mm 3 ultrasonic free-floating implant for multi-mote neural recording". In: *IEEE Journal of Solid-State Circuits* 54.11 (2019), pp. 3017–3030.

[47] Carl Gold et al. "On the origin of the extracellular action potential waveform: a modeling study". In: *Journal of neurophysiology* 95.5 (2006), pp. 3113–3128.

[48] Marceli Nunes Gonçalves and Marcelo Martins Werneck. "A temperature-independent optical voltage transformer based on FBG-PZT for 13.8 kV distribution lines". In: *Measurement* 147 (2019), p. 106891.

[49] P Gräupner et al. "Electro-optical effect in aluminum nitride waveguides". In: *Journal of applied physics* 71.9 (1992), pp. 4136–4139.

[50] C Gui et al. "The effect of surface roughness on direct wafer bonding". In: *Journal of applied physics* 85.10 (1999), pp. 7448–7454.

[51] Mohamad HajjHassan, Vamsy Chodavarapu, and Sam Musallam. "NeuroMEMS: neural probe microtechnologies". In: *Sensors* 8.10 (2008), pp. 6704–6726.

[52] Matti Hämäläinen et al. "Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain". In: *Reviews of modern Physics* 65.2 (1993), p. 413.

[53] John G Hartnett, Michael E Tobar, and Jerzy Krupka. "Dependence of the dielectric permittivity of single-crystal quartz on thermal deformation at cryogenic temperatures". In: *Journal of Applied Physics* 102.7 (2007), p. 074103.

[54] Center for Health Security. "Ventilator Stockpiling and Availability in the US". In: *John's Hopkins School of Public Health* (2020).

[55] GE Healthcare. *D-lite++ Patient Spirometry Sets*. Accessed: 9-11-2020.

[56] Eugene Hecht and Alfred Zajac. *Optics*. Vol. 4. Addison Wesley San Francisco, 2002.

[57] Izumi Hirano et al. "Time-dependent dielectric breakdown (TDDB) distribution in n-MOSFET with HfSiON gate dielectrics under DC and AC stressing". In: *Microelectronics Reliability* 53.12 (2013), pp. 1868–1874.

[58] Phillip CD Hobbs. "Reaching the shot noise limit for $10". In: *Optics and Photonics News* 2.4 (1991), pp. 17–23.

[59] Brad W Hoff et al. "Characterization of AlN-based ceramic composites for use as millimeter-wave susceptor materials at high temperature: Dielectric properties of AlN: Mo with 0.25 vol% to 4.0 vol% Mo from 25 to 550 C". In: *Journal of Materials Research* 34.15 (2019), pp. 2573–2581.

[60] Roarke Horstmeyer, Haowen Ruan, and Changhuei Yang. "Guidestar-assisted wavefront-shaping methods for focusing light into biological tissue". In: *Nature photonics* 9.9 (2015), pp. 563–571.

[61] Li Hui et al. "Tracking algorithm for the gain of the phase modulator in closed-loop optical voltage sensors". In: *Optics & Laser Technology* 47 (2013), pp. 214–220.

[62] Steven L Jacques. "Optical properties of biological tissues: a review". In: *Physics in Medicine & Biology* 58.11 (2013), R37.

[63] Daniel Johnston et al. "Active dendrites, potassium channels and synaptic plasticity". In: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 358.1432 (2003), pp. 667–674.

[64] M Jueptner and C Weiller. "does measurement of regional cerebral blood flow reflect synaptic activity? Implications for PET and fMRI." In: *Neuroimage* 2.2 (1995), pp. 148–156.

[65] Kazuhiko Kano et al. "Temperature dependence of piezoelectric properties of sputtered AlN on silicon substrate". In: *Sensors and Actuators A: Physical* 130 (2006), pp. 397–402.

[66] Charalambos C Katsidis and Dimitrios I Siapkas. "General transfer-matrix method for optical multilayer systems with coherent, partially coherent, and incoherent interference". In: *Applied optics* 41.19 (2002), pp. 3978–3987.

[67] Gerd Keiser. *Optical fiber communications*. Vol. 2. McGraw-Hill New York, 2000.

[68] Adam Khalifa et al. "The Microbead: A 0.009 mm^3 Implantable Wireless Neural Stimulator". In: *IEEE Transactions on Biomedical Circuits and Systems* 13.5 (2019), pp. 971–985.

[69] Il Yong Kim and Oliver L De Weck. "Adaptive weighted-sum method for bi-objective optimization: Pareto front generation". In: *Structural and multidisciplinary optimization* 29.2 (2005), pp. 149–158.

[70] Fabian Kloosterman et al. "Micro-drive array for chronic in vivo recording: drive fabrication". In: *JoVE (Journal of Visualized Experiments)* 26 (2009), e1094.

[71] Fumio Koyama. "Recent advances of VCSEL photonics". In: *Journal of Lightwave Technology* 24.12 (2006), pp. 4502–4513.

[72] Jerzy Krupka et al. "Complex permittivity of some ultralow loss dielectric crystals at cryogenic temperatures". In: *Measurement Science and Technology* 10.5 (1999), p. 387.

[73] ILAN Lampl and YOSEF Yarom. "Subthreshold oscillations of the membrane potential: a functional synchronizing and timing device". In: *Journal of neurophysiology* 70.5 (1993), pp. 2181–2186.

[74] Günter Last and Mathew Penrose. *Lectures on the Poisson process*. Vol. 7. Cambridge University Press, 2017.

[75] Mariya Lazebnik et al. "Functional optical coherence tomography for detecting neural activity through scattering changes". In: *Optics letters* 28.14 (2003), pp. 1218–1220.

[76] Chung-Hsun Lee et al. "Toxicity of tetramethylammonium hydroxide: review of two fatal cases of dermal exposure and development of an animal model". In: *Toxicology and industrial health* 27.6 (2011), pp. 497–503.

[77] Dong-Joon Lee and John F Whitaker. "Optimization of sideband modulation in optical-heterodyne-downmixed electro-optic sensing". In: *Applied optics* 48.8 (2009), pp. 1583–1590.

[78] Ethan S Lee et al. "Time-dependent dielectric breakdown under AC stress in GaN MIS-HEMTs". In: *2019 IEEE International Reliability Physics Symposium (IRPS)*. IEEE. 2019, pp. 1–5.

[79] Jihun Lee et al. "Neural recording and stimulation using wireless networks of microimplants". In: *Nature Electronics* 4.8 (2021), pp. 604–614.

[80] Jun Su Lee et al. "Meeting the electrical, optical, and thermal design challenges of photonic-packaging". In: *IEEE Journal of selected topics in quantum electronics* 22.6 (2016), pp. 409–417.

[81] Sunwoo Lee et al. "A 250 μm$\times$ 57 μm microscale opto-electronically transduced electrodes (MOTEs) for neural recording". In: *IEEE transactions on biomedical circuits and systems* 12.6 (2018), pp. 1256–1266.

[82] Sunwoo Lee et al. "A 250 m $\times$ 57 m Microscale Opto-electronically Transduced Electrodes (MOTEs) for Neural Recording". In: *IEEE Transactions on Biomedical Circuits and Systems* 12.6 (2018), pp. 1256–1266.

[83] Fei Li et al. "Determination of temperature dependence of piezoelectric coefficients matrix of lead zirconate titanate ceramics by quasi-static and resonance method". In: *Journal of Physics D: Applied Physics* 42.9 (2009), p. 095417.

[84] Hongquan Li et al. "Utah-Stanford Ventilator (Vent4US): Developing a rapidly scalable ventilator for COVID-19 patients with ARDS". In: *medRxiv* (2020).

[85] Hui Li et al. "An analysis on the optimization of closed-loop detection method for optical voltage sensor based on Pockels effect". In: *Journal of lightwave technology* 32.5 (2014), pp. 1006–1013.

[86] Patricia Yang Liu et al. "Cell refractive index for cell biology and disease diagnosis: past, present and future". In: *Lab on a Chip* 16.4 (2016), pp. 634–644.

[87] Victor Liu and Shanhui Fan. "S4: A free electromagnetic solver for layered periodic structures". In: *Computer Physics Communications* 183.10 (2012), pp. 2233–2244.

[88] Lan Luan et al. "Ultraflexible nanoelectronic probes form reliable, glial scar–free neural integration". In: *Science advances* 3.2 (2017), e1601966.

[89] CM Lueng et al. "Piezoelectric coefficient of aluminum nitride and gallium nitride". In: *Journal of applied physics* 88.9 (2000), pp. 5360–5363.

[90] Adam H Marblestone et al. "Physical principles for scalable neural recording". In: *Frontiers in computational neuroscience* 7 (2013), p. 137.

[91] Travis L Massey et al. "A high-density carbon fiber neural recording array technology". In: *Journal of neural engineering* 16.1 (2019), p. 016024.

[92] Edwin M Maynard, Craig T Nordhausen, and Richard A Normann. "The Utah intracortical electrode array: a recording structure for potential brain-computer interfaces". In: *Electroencephalography and clinical neurophysiology* 102.3 (1997), pp. 228–239.

[93] John I McCool. *Using the Weibull distribution: reliability, modeling, and inference*. Vol. 950. John Wiley & Sons, 2012.

[94] *Memorandum on Order Under the Defense Production Act Regarding the Purchase of Ventilators*. URL: `https://www.whitehouse.gov/presidential-actions/memorandum-order-defense-production-act-regarding-purchase-ventilators/`.

[95] Pekka Meriläinen, Hannu Hänninen, and Lauri Tuomaala. "A novel sensor for routine continuous spirometry of intubated patients". In: *Journal of clinical monitoring* 9.5 (1993), pp. 374–380.

[96] MG Moharam and TK Gaylord. "Rigorous coupled-wave analysis of planar-grating diffraction". In: *JOSA* 71.7 (1981), pp. 811–818.

[97] Stefan Möhlenkamp and Holger Thiele. "Ventilation of COVID-19 patients in intensive care units". In: *Herz* (2020), p. 1.

[98] Alyosha Christopher Molnar et al. "Nanoliter-Scale Autonomous Electronics: Advances, Challenges, and Opportunities". In: *2021 IEEE Custom Integrated Circuits Conference (CICC)*. IEEE, 2021, pp. 1–6.

[99] Abigail Morrison et al. "Exact subthreshold integration with continuous spike times in discrete-time neural network simulations". In: *Neural computation* 19.1 (2007), pp. 47–79.

[100] Muhammad Mujeeb-U-Rahman et al. "Optical power transfer and communication methods for wireless implantable sensing platforms". In: *Journal of Biomedical Optics* 20.9 (2015), pp. 1–9.

[101] Rikky Muller, Simone Gambini, and Jan M Rabaey. "A $0.013mm^2$, $5\mu W$, DC-Coupled Neural Signal Acquisition IC With 0.5 V Supply". In: *IEEE Journal of Solid-State Circuits* 47.1 (2011), pp. 232–243.

[102] Elon Musk et al. "An integrated brain-machine interface platform with thousands of channels". In: *Journal of medical Internet research* 21.10 (2019), e16194.

[103] Rajesh V Nair and Ramarao Vijaya. "Photonic crystal sensors: An overview". In: *Progress in Quantum Electronics* 34.3 (2010), pp. 89–134.

[104] Ryuichi Nakajima and Bradley J Baker. "Mapping of excitatory and inhibitory post-synaptic potentials of neuronal populations in hippocampal slices using the GEVI, ArcLight". In: *Journal of physics D: Applied physics* 51.50 (2018), p. 504003.

[105] Donald A Neamen. *Semiconductor physics and devices: basic principles*. McGraw-hill, 2003.

[106] Abdulmalik Obaid et al. "Massively parallel microwire arrays integrated with CMOS chips for neural recording". In: *Science Advances* 6.12 (2020), eaay2789.

[107] Kwi-Il Park et al. "Piezoelectric BaTiO3 thin film nanogenerator on plastic substrates". In: *Nano letters* 10.12 (2010), pp. 4939–4943.

[108] J Pastrňák and L Roskovcová. "Refraction index measurements on AlN single crystals". In: *physica status solidi (b)* 14.1 (1966), K5–K8.

[109] Paras R Patel et al. "Chronic in vivo stability assessment of carbon fiber microelectrode arrays". In: *Journal of neural engineering* 13.6 (2016), p. 066002.

[110] Americo Pereira et al. "Proof-of-concept of a minimalist pressure-controlled emergency ventilator for COVID-19". In: *arXiv preprint arXiv:2004.00310* (2020).

[111] Kiryl D Piatkevich et al. "A robotic multidimensional directed evolution approach applied to fluorescent voltage reporters". In: *Nature chemical biology* 14.4 (2018), pp. 352–360.

[112] David K. Piech et al. "A wireless millimetre-scale implantable neural stimulator with ultrasonically powered bidirectional communication". In: *Nature Biomedical Engineering* 4 (2020), pp. 207–222.

[113] Miroslav Poljak and Boris Bojanić. "Method for the reduction of in-service instrument transformer explosions". In: *European transactions on electrical power* 20.7 (2010), pp. 927–937.

[114] Valentina Quaresima and Marco Ferrari. "Functional near-infrared spectroscopy (fNIRS) for assessing cerebral cortex function during human behavior in natural/social situations: a concise review". In: *Organizational Research Methods* 22.1 (2019), pp. 46–68.

[115] F Rahmatian and PP Chavez. "SF/sub 6/-free 550 kV combined optical voltage and current transducer system". In: *2003 IEEE PES Transmission and Distribution Conference and Exposition (IEEE Cat. No. 03CH37495)*. Vol. 1. IEEE. 2003, pp. 379–382.

[116] Farnoosh Rahmatian and James N Blake. *Sensors and sensing methods for three-phase, gas insulated devices*. US Patent 7,633,285. 2009.

[117] Farnoosh Rahmatian, Patrick P Chavez, and Nicolas AF Jaeger. "138 kV and 345 kV wide-band SF/sub 6/-free optical voltage transducers". In: *2002 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No. 02CH37309)*. Vol. 2. IEEE. 2002, pp. 1472–1477.

[118] Megan L Ranney, Valerie Griffeth, and Ashish K Jha. "Critical supply shortages—the need for ventilators and personal protective equipment during the Covid-19 pandemic". In: *New England Journal of Medicine* 382.18 (2020), e41.

[119] Samuel J. Raymond et al. "A low-cost, rapidly scalable, emergency use ventilator for the COVID-19 crisis". In: *medRxiv* (2020).

[120] Antoni Rogalski. "Infrared detectors: an overview". In: *Infrared physics & technology* 43.3-5 (2002), pp. 187–210.

[121] C Rossel et al. "Temperature dependence of the transverse piezoelectric coefficient of thin films and aging effects". In: *Journal of Applied Physics* 115.3 (2014), p. 034105.

[122] Raymond C Rumpf. "Improved formulation of scattering matrices for semi-analytical methods that is consistent with convention". In: *Progress In Electromagnetics Research B* 35 (2011), pp. 241–261.

[123] SL Rumyantsev et al. "Low frequency noise and long-term stability of noncoherent light sources". In: *Journal of applied physics* 96.2 (2004), pp. 966–969.

[124] Patrick Ruther and Oliver Paul. "New approaches for CMOS-based devices for large-scale neural recording". In: *Current opinion in neurobiology* 32 (2015), pp. 31–37.

[125] Joseph W Salatino et al. "Glial responses to implanted electrodes in the brain". In: *Nature biomedical engineering* 1.11 (2017), pp. 862–877.

[126] GA Sanders et al. "Commercialization of fiber-optic current and voltage sensors at NxtPhase". In: *2002 15th Optical Fiber Sensors Conference Technical Digest. OFS 2002 (Cat. No. 02EX533)*. IEEE. 2002, pp. 31–34.

[127] Charles E Seeley et al. "Packaging and Performance of a Piezo-Optic Voltage Sensor". In: *ASME International Mechanical Engineering Congress and Exposition*. Vol. 42991. 2007, pp. 287–296.

[128] Dongjin Seo et al. "Wireless recording in the peripheral nervous system with ultrasonic neural dust". In: *Neuron* 91.3 (2016), pp. 529–539.

[129] CJR Sheppard. "Approximate calculation of the reflection coefficient from a stratified medium". In: *Pure and Applied Optics: Journal of the European Optical Society Part A* 4.5 (1995), p. 665.

[130] Soner Sonmezoglu et al. "A Method and Analysis to Enable Efficient Piezoelectric Transducer-Based Ultrasonic Power and Data Links for Miniaturized Implantable Medical Devices". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 68.11 (2021), pp. 3362–3370.

[131] Soner Sonmezoglu et al. "Monitoring deep-tissue oxygenation with a millimeter-scale ultrasonic implant". In: *Nature Biotechnology* 39 (2021), pp. 855–864.

[132] *Specification for ventilators to be used in UK hospitals during the coronavirus (COVID-19) outbreak*. `https://www.gov.uk/government/publications/specification-for-ventilators-to-be-used-in-uk-hospitals-during-the-coronavirus-covid-19-outbreak`. 2020.

[133] Nicholas A Steinmetz et al. "Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings". In: *Science* 372.6539 (2021), eabf4588.

[134] Christian Stelling et al. "Plasmonic nanomeshes: their ambivalent role as transparent electrodes in organic solar cells". In: *Scientific reports* 7.1 (2017), pp. 1–13.

[135] Ajit C Tamhane. *Statistical analysis of designed experiments: theory and applications*. Vol. 609. John Wiley & Sons, 2009.

[136] Michal Teplan et al. "Fundamentals of EEG measurement". In: *Measurement science review* 2.2 (2002), pp. 1–11.

[137] J Tharion et al. "Rapid Manufacturable Ventilator for Respiratory Emergencies of COVID-19 Disease". In: *Transactions of the Indian National Academy of Engineering* (2020), pp. 1–6.

[138] Jordan Thimot and Kenneth L. Shepard. "Bioelectronic devices: Wirelessly powered implants". In: *Nature Biomedical Engineering* 1.3 (2017), p. 0051.

[139] Anish Thukral et al. "Soft Neural Interfaces for Ultrathin Electronics". In: *IEEE Nanotechnology Magazine* PP (Jan. 2018), pp. 1–1. DOI: `10.1109/MNANO.2017.2781290`.

[140] Lin Tian, S Andrew Hires, and Loren L Looger. "Imaging neuronal activity with genetically encoded calcium indicators". In: *Cold Spring Harbor Protocols* 2012.6 (2012), pdb–top069609.

[141] Rodney S Tucker. "Green optical communications—Part I: Energy limitations in transport". In: *IEEE Journal of selected topics in quantum electronics* 17.2 (2010), pp. 245–260.

[142] Alexei Verkhratsky and Vladimir Parpura. "History of electrophysiology and the patch clamp". In: *Patch-clamp methods and protocols*. Springer, 2014, pp. 1–19.

[143] JF Verweij and JH Klootwijk. "Dielectric breakdown I: A review of oxide breakdown". In: *Microelectronics Journal* 27.7 (1996), pp. 611–622.

[144] Mirjana Vijatović, Jelena Bobić, and Biljana D Stojanović. "History and challenges of barium titanate: Part II". In: *Science of Sintering* 40.3 (2008), pp. 235–244.

[145] Richard Edward Wagner and Wilfred J Tomlinson. "Coupling efficiency of optics in single-mode fiber components". In: *Applied optics* 21.15 (1982), pp. 2671–2688.

[146] Robert H Walden. "Analog-to-digital converter survey and analysis". In: *IEEE Journal on selected areas in communications* 17.4 (1999), pp. 539–550.

[147] Robert H Walden. "Analog-to-digital converter technology comparison". In: *Proceedings of 1994 IEEE GaAs IC Symposium*. IEEE. 1994, pp. 217–219.

[148] J Wang et al. "An innovative transparent cranial window based on skull optical clearing". In: *Laser physics letters* 9.6 (2012), p. 469.

[149] Max L. Wang et al. "Wireless data links for next-generation networked micro-implantables". In: *2018 IEEE Custom Integrated Circuits Conference (CICC)*. 2018, pp. 1–9. DOI: 10.1109/CICC.2018.8357096.

[150] Woradorn Wattanapanitch, Michale Fee, and Rahul Sarpeshkar. "An energy-efficient micropower neural recording amplifier". In: *IEEE Transactions on Biomedical Circuits and Systems* 1.2 (2007), pp. 136–147.

[151] Patrick D Wolf and WM Reichert. "Thermal considerations for the design of an implanted cortical brain–machine interface (BMI)". In: *Indwelling Neural Implants: Strategies for Contending with the In Vivo Environment* (2008), pp. 33–38.

[152] RA Wolf and S Trolier-McKinstry. "Temperature dependence of the piezoelectric response in lead zirconate titanate films". In: *Journal of applied physics* 95.3 (2004), pp. 1397–1406.

[153] Helen H Yang and François St-Pierre. "Genetically encoded voltage indicators: opportunities and challenges". In: *Journal of Neuroscience* 36.39 (2016), pp. 9977–9989.

[154] Qing Yang et al. "An optical fiber Bragg grating and piezoelectric ceramic voltage sensor". In: *Review of Scientific Instruments* 88.10 (2017), p. 105005.

[155] Pavel S Yarmolenko et al. "Thresholds for thermal damage to normal tissues: an update". In: *International Journal of Hyperthermia* 27.4 (2011), pp. 320–343.

[156] Seokchan Yoon et al. "Deep optical imaging within complex scattering media". In: *Nature Reviews Physics* 2.3 (2020), pp. 141–158.

[157] Toshihiko Yoshino et al. "Fiber-optic Fabry-Perot interferometer and its sensor applications". In: *IEEE Transactions on Microwave Theory and Techniques* 30.10 (1982), pp. 1612–1621.