A Miniature, Implantable Wireless Neural Stimulation System

by

Scott K. Arfin

Bachelor of Science in Electrical Engineering
Columbia University, 2004

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2006

Author .....
Department of Electrical Engineering and Computer Science
May 12, 2006

Certified by .....
Rahul Sarpeshkar
Associate Professor of Electrical Engineering
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

A Miniature, Implantable Wireless Neural Stimulation System

by

Scott K. Arfin

Submitted to the Department of Electrical Engineering and Computer Science
on May 12, 2006, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering

Abstract

In this thesis, I present the design of a wireless neural stimulation system. The system consists of an external transmitter, controllable through a computer interface, and a miniature, implantable wireless receiver and stimulator. The implant is tailored for use in zebra finches – small birds weighing just 12-15g – as part of ongoing research into the neural mechanisms of sequence generation and learning. The implant, assembled on a miniature printed circuit board, contains a receiver coil, battery, electrodes, and a custom integrated circuit for data demodulation and neural stimulation. The chip, fabricated in a standard 0.5μm CMOS process, is capable of delivering biphasic current pulses to 4 addressable electrode sites at 16 selectable current levels ranging from 100μA to 1mA. Additionally, the biphasic pulses may be inverted. The entire implant weighs less than 1.5g and occupies a footprint smaller than 1.5cm^2.

A miniaturized neural stimulator such as this one also has applications in neural prostheses for blindness, Parkinson's disease, and paralysis.

Thesis Supervisor: Rahul Sarpeshkar
Title: Associate Professor of Electrical Engineering

Acknowledgments

There are many people whom I would like to thank for supporting me while I did the research for this thesis.

I would like to thank my adviser, Professor Rahul Sarpeshkar for his thoughtful advice and inspirational teaching. I have learned an enormous amount about circuit design and how to conduct research from him. Also, thanks to Professor Michale Fee in the Department of Brain and Cognitive Sciences for providing the vision for this project.

I would also like to thank my parents, Paula and Ronald, and my brother Daniel. My family has been there for me from the beginning. Their level of support is unsurpassed.

Next, I would like to thank the past and present members of the Analog VLSI & Biological Systems Group who helped and supported me with this research. They are Christopher Salthouse, Ji-Jon Sit, Keng-Hoong Wee, Lorenzo Turicchia, Woradorn Wattanapanitch, Micah O'Halloran, Michael Baker, Serhii Zhak, Soumyajit Mandal, Tim Lu, Heemin Yang, Maziar Tavakoli-Dastjerdi, and Andrew Selbst. I would also like to thank some past and present members of Professor John Wyatt's group who have aided me. They are Mariana Markova, Luke Theogarajan, and Dr. Shawn Kelly. I also owe thanks to Rose Englund and Gretchen Jones, administrative assistants for AVNSL, and my academic adviser, Professor Jae Lim.

On a lighter note, I would like to thank the Boston Red Sox for their once-in-a-lifetime feat, which occurred during my first month as a student and Boston area resident here at MIT.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Neural Mechanisms of Sequence Generation and Learning

The focus of my thesis research is on the design and implementation of a miniature, implantable wireless neural stimulation system. The purpose of the device is to study the neural mechanisms of sequence generation in small songbirds. However, this work will also benefit research into the neurological behavior of larger animals such as rats, and even monkeys and humans, since the techniques explored here are also more generally applicable. In particular, hopes of curing serious neurological disorders in humans such as blindness, Parkinson's disease, and even paralysis, will be aided by wireless neural stimulation systems.

Adult male zebra finches (songbirds) are a popular subject for neurological study. Their song reproductions are an excellent example of the generation and learning of complex sequences in the brain. Further, avian brain areas involved in song learning are closely related to mammalian brain areas involved in motor learning [5]. Thus, the study of song generation in birds may teach us more general principles about sequence generation and learning in humans.

In previous experiments, researchers placed songbirds inside a cubic cage approximately 20cm on an edge inside an acoustically isolated chamber. Then, using several

carefully positioned electrodes implanted into the bird's brain and a microphone, the researchers simultaneously recorded sounds of the bird singing and the neural activity in its brain. However, they could not avoid wires coming out of the implant, impeding the bird's motion. To mitigate this difficulty, they used a bundle of fine wires connected to a motorized, commutating post. The motor ran automatically under a feedback control to reduce the torque experienced by the bird due to the bundle of wires [6]. However, an even better solution would be to replace the bundle of wires altogether with a wireless data link.

In this thesis, I implemented a wireless neural stimulation system rather than a recording system. Researchers plan to use wireless neural stimulation in the brain of the bird to attempt to influence and control the bird's song generation. For instance, a microphone and digital signal processor may be used to detect when the bird begins singing. When this happens, a computer program initiates wireless commands to stimulate neurons in the bird's brain.

1.2 System Description

The wireless system, shown in Figure 1-1 consists of two parts: an external transmitter, controllable through a computer interface, and a miniature, implantable wireless receiver and stimulator. The implant is assembled on a miniature printed circuit board, and contains a receiver coil, battery, electrodes, and a custom integrated circuit for data demodulation and neural stimulation. The chip is fabricated in standard 0.5μm CMOS process. It is capable of delivering biphasic current pulses to 4 addressable electrode sites at 16 selectable current levels ranging from 100μA to 1mA. Additionally, the biphasic pulses may be inverted. The entire implant weighs less than 1.5g and occupies a footprint smaller than 1.5cm^2. The system supports a data rate of 25kbps. Data rates of 1Mbps are also possible. Power is supplied from a 3.0V, 5mAh lithium-manganese rechargeable battery. The implant can run up to approximately 3 days on a single charge, depending on stimulation rate. Future improvements in static power dissipation will increase the operation time to approximately 20 days.

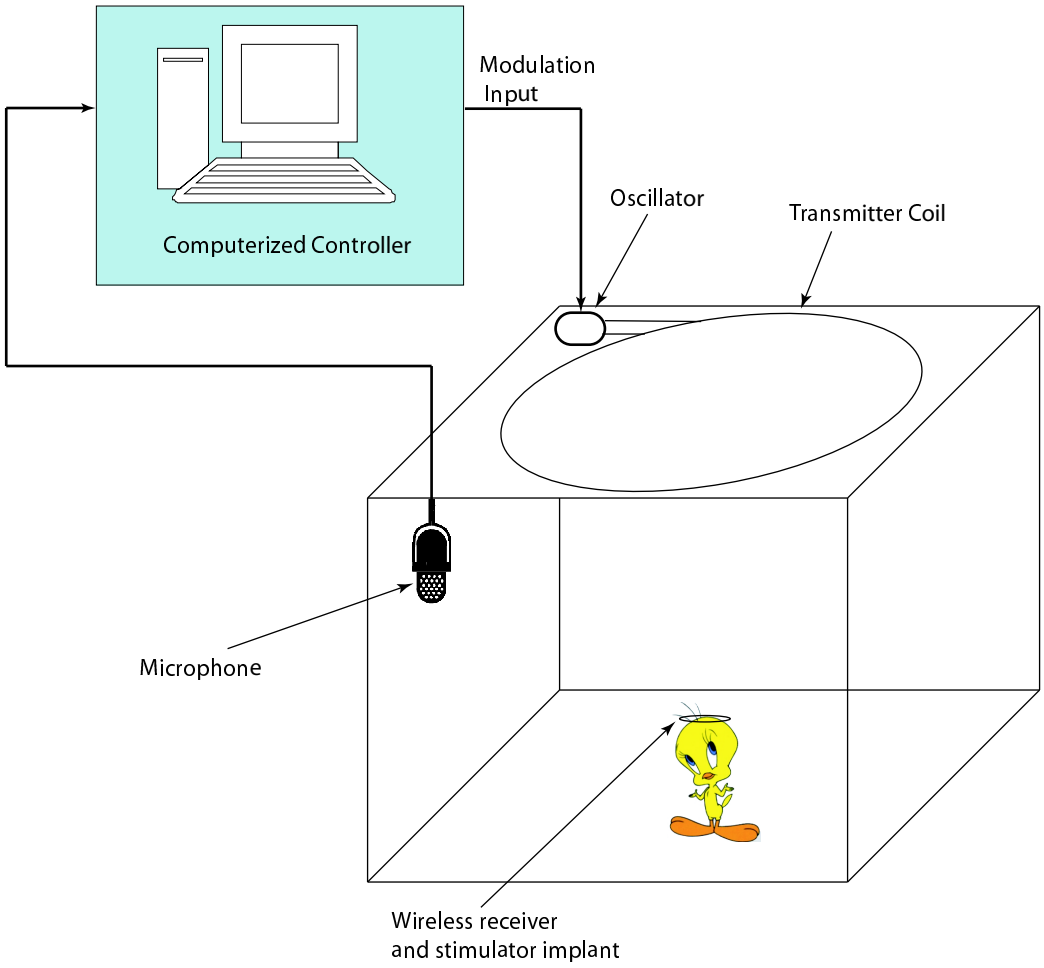


Figure 1-1: Overview of entire wireless system.

The overall size of the implant is restricted by the physical size of the zebra finch itself. Figure 1-2 shows a male zebra finch. Zebra finches generally weigh 12-15 grams. As a result, these small songbirds can only carry up to 2g on the cranium [6]. The implantable portion of the stimulation system must also fit within an area of 1.5cm^2 in order to fit adequately on the bird's head. A major challenge of this project was designing the implanted portion of the system to meet these size and weight specifications.

Figure 1-2: Male Zebra Finch.

1.3 Applications to Neural Prosthetics

One application of wireless neural stimulation systems, apart from studying the neural mechanisms of sequence generation, is cortical visual prostheses for the blind. In 1929, Foerster successfully studied the effects of electrical stimulation of the occipital lobe of the human cortex. A patient who received this stimulation reported "seeing" a spot of light [8]. Since then, the idea of using neural stimulations at a single site to produce spots of light has evolved into using many sites simultaneously to produce entire images. A major breakthrough in cortical prostheses occurred in 1968 when Dr. Brindley and Dr. Lewin implanted a device with an array of 80 electrodes into a blind patient. They activated the sites using a wireless interface in which they placed the transmitting coil of an oscillator tuned to the appropriate frequency on the scalp directly over the electrode site. The patient reported seeing spots that corresponded approximately to the position of the electrode in the visual cortex [2].

Figure 1-3 shows an example of a practical, chronically implantable cortical visual prosthesis. An external light sensor captures and encodes video and passes this data to a digital signal processor, which maps the light information into neural stimulation data. The stimulation data is then transmitted across a transcranial wireless link, formed from a network of coupled coils placed on both sides of the skull, to a chip implanted inside the brain. The chip demodulates the wireless data and stimulates in the visual cortex. The wireless link in this thesis, designed at 13.5MHz, is also

compatible with human tissue. Loss due to attenuation from the skin becomes a significant problem at frequencies greater than 30MHz [15]. A practical visual prosthesis would also require many more stimulation sites than the 4 implemented in this thesis. However, due to the digital programmability of electrode selection, adding additional stimulation sites needed for vision would be easily accommodated.

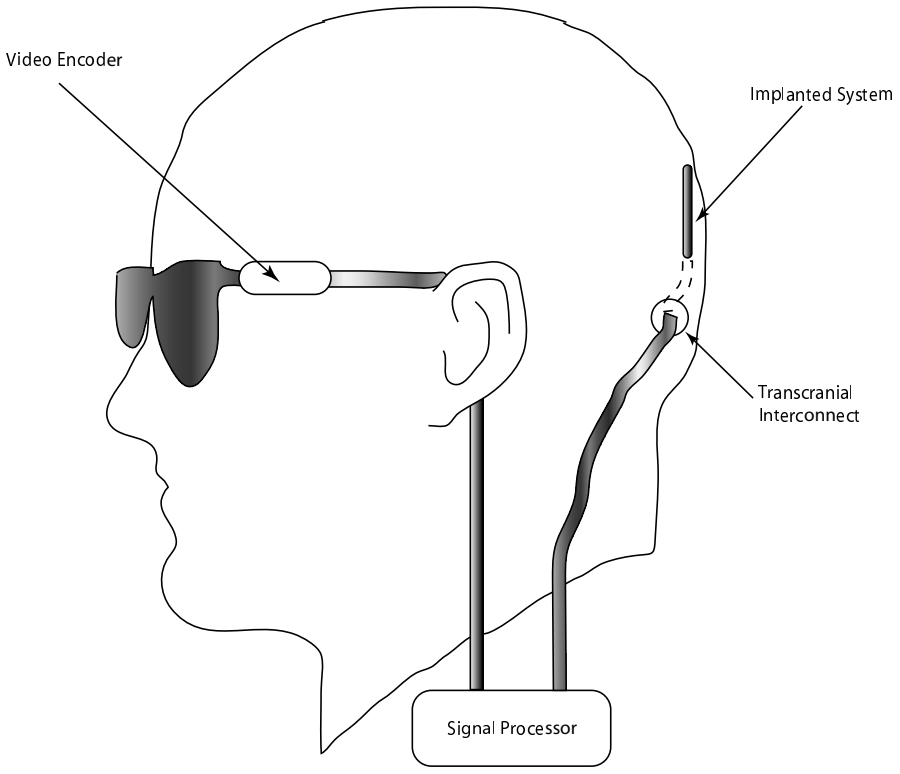


Figure 1-3: Overview of a cortical visual prosthesis. The transcranial interconnect uses inductively coupled coils to wirelessly transmit data and power to the implanted system containing a VLSI chip and an array of electrodes.

Chapter 2

Coupled Coils

2.1 Introduction

The wireless neural stimulation system presented in this thesis employs a pair of weakly coupled coils for sending data between the transmitter and the implant. The coils can be modeled as a transformer with a very small coupling coefficient, k. In this regard, the transmitter coil shall be referred to as the "primary" and the receiver coil as the "secondary."

A number of competing constraints resulted in coupled coils as the mechanism for data transmission. The factors considered were the size and weight of the implant, data rate, operating distance, and power consumption. If we are constrained to fit the implant into an area not exceeding 1.5cm^2 and with a mass not exceeding 1.5g, then it will not be possible to build an efficient antenna at any standard radio frequency that meets these considerations. An efficient antenna typically has some dimension equal to a certain fraction of the wavelength it is designed receive. For example, radio waves at 2.4GHz in free space have a wavelength of 12.5cm. An efficient antenna design at this frequency that also meets the area and weight specifications would be difficult to achieve. The problem only becomes more difficult at lower frequencies. Further, a wireless receiver at 2.4GHz would consume a significant amount of power – easily several milliwatts. The bandwidth afforded by such a high operating frequency well exceeds the 25kbps that is required. Shifting to a lower frequency to reduce

power consumption is desirable since the life of the implant depends on its power consumption. The entire implant needs to be powered by a miniature battery which should last the length of an experiment without having to be replaced. Fortunately, the distances over which we must transmit are quite small. The bird enclosure is a cube shaped box 20cm on each edge. Therefore, with the bird present in the cage and wearing the implant on its head, transmission distances may be approximately 5-15cm. At this range, selecting an operating frequency in which we are still in the magnetic near-field enables us to use a pair of coupled coils as the link between the transmitter and receiver [20].

For this application, we choose 13.56MHz. This unlicensed band is designated by the FCC for Industrial/Scientific/Medial (ISM) use. It also provides adequate bandwidth for our application. Even if demands on data rate increase, this band should still be sufficient. With a wavelength in air of approximately 22m, magnetic coupling is a viable option in the range of 5-15cm.

2.2 Mutual Inductors

When two inductors are placed near each other, or are coupled via a core with a high magnetic permeability, magnetic flux produced by either inductor is coupled to the other. The linked flux then induces a voltage in the opposite coil. We will exploit this property of coupled coils in our design.

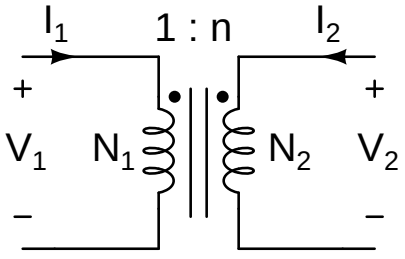


Figure 2-1: Schematic representation of an ideal transformer.

In an ideal transformer comprised of two windings on a common, infinitely permeable magnetic core, driving the primary side produces a time varying magnetic flux, of

which all is coupled to the secondary side. Figure 2-1 is a schematic representation of an ideal transformer. N_1 is the number of turns of the primary and N_2 is the number of turns of the secondary. We can then define the turns ratio, n, to be N_2/N_1. It can be readily shown that the terminal voltages and currents are related by:

$$\frac{V_2}{V_1} = n \tag{2.1}$$

$$\frac{I_2}{I_1} = -\frac{1}{n} \tag{2.2}$$

To deal with the case where not all of the flux created by the primary is linked to the secondary, one must introduce the concept of mutual inductance. After some labor, one arrives at the terminal relations [4]:

$$V_1 = L_1\frac{dI_1}{dt} + M\frac{dI_2}{dt} \tag{2.3}$$

$$V_2 = L_2\frac{dI_2}{dt} + M\frac{dI_1}{dt} \tag{2.4}$$

where

$$M = k\sqrt{L_1 L_2} \tag{2.5}$$

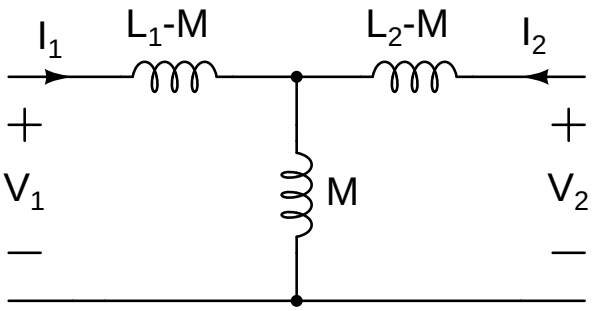


Figure 2-2: Equivalent T-model of mutual inductors.

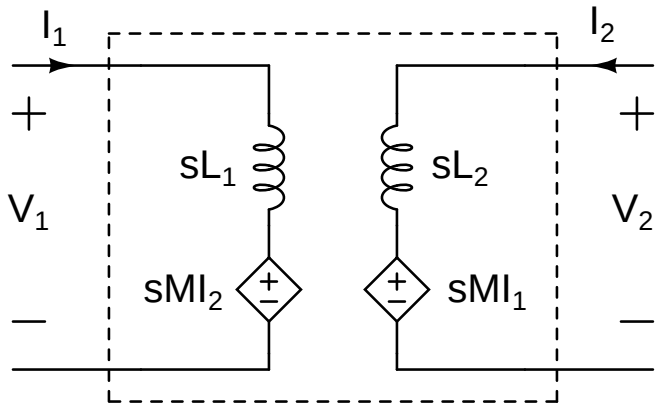


Figure 2-3: Two-port view of mutual inductors.

The coupling coefficient, k, is a measure of the amount of coupling between the two coils. A coefficient of $k = 1$ corresponds to the ideal case where all of the flux links both coils. For any physical system, $0 < k < 1$. In our wireless data link, we expect the exact value of k to be difficult to calculate and likely to change during the course of an experiment as the bird moves inside the cage, or alters the angle of its head. However, we can be certain that $k \ll 1$, and will use this approximation to help simplify the design.

2.3 Analysis of Mutual Inductors Using Feedback

Equations 2.3 and 2.4 are not especially useful by themselves for understanding or analyzing circuits with mutual inductance. However, the two two-port models in Figures 2-2 and 2-3 are equivalent representations of these differential equations. The two-port model in Figure 2-3 arises directly from the differential equations. The T-model of mutual inductors in Figure 2-2, although it may contain negative inductances, is externally equivalent to the circuit in Figure 2-3. The T-model can also be verified by solving for the circuit's $z-$parameters. The equivalent two-port model in Figure 2-3 is then obtained.

To further aid in our understanding, we would like to construct a feedback model of the system of coupled inductors shown in Figure 2-4. We have taken the two-port model of Figure 2-3 and added a voltage source at port one. We have also replaced

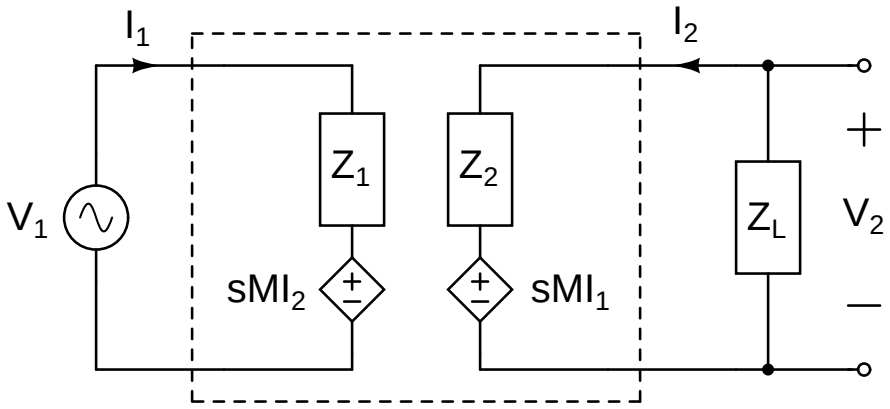


Figure 2-4: Two-port view of mutual inductors with source and load.

the inductances with general impedances Z_1 and Z_2 and connected a load, Z_L, to port two. The resulting feedback model is shown in Figure 2-5.

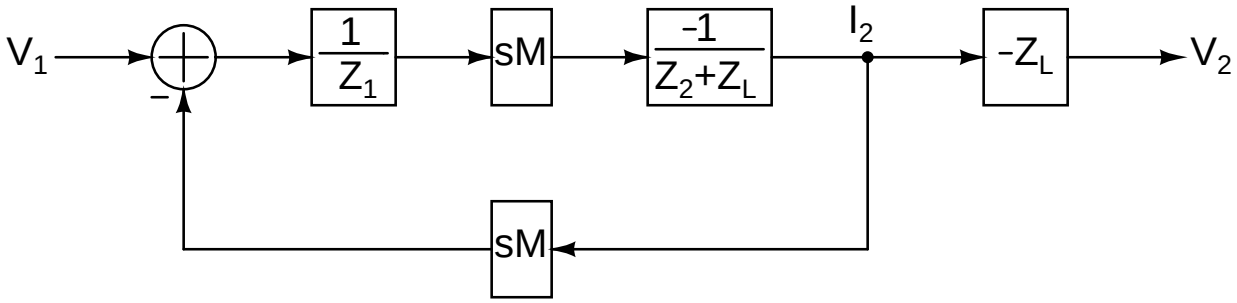


Figure 2-5: Block diagram view of mutual inductors.

$$\frac{V_2}{V_1} = \frac{sMZ_L}{Z_1\left(Z_2 + Z_L\right) - s^2M^2} \tag{2.6}$$

$$\frac{V_2}{V_1} = \frac{MZ_L}{Z_L L_1 + sL_1L_2 - sM^2} \tag{2.7}$$

Applying Black's Formula to the feedback block diagram, we obtain Equation 2.6. We also now assume that both inductors L_1 and L_2 are of high quality factor, and replace them with their respective impedances sL_1 and sL_2, obtaining Equation 2.7. Next, we substitute in the denominator $M^2 = k^2L_1L_2$, and obtain Equation 2.8.

$$\frac{V_2}{V_1} = \frac{MZ_L}{Z_L L_1 + sL_1L_2\left(1 - k^2\right)} \tag{2.8}$$

We now recognize that in the situation where the two coils are quite far apart and the received signals are small ($k \ll 1$), the k^2 term in the denominator of Equation 2.8 may be eliminated. The last step is to replace the load impedance, Z_L, with a parallel RC network. The C is chosen to resonate with L_2 at the primary coil driving frequency. Upon making this final substitution, we obtain the familiar form of the resonant low-pass filter, multiplied by a scaling term, β. This result is shown in Equation 2.9.

$$\frac{V_2}{V_1} = \beta \cdot \frac{1}{s^2 L_2 C + \frac{sL_2}{R} + 1} \tag{2.9}$$

$$\beta = k\sqrt{\frac{L_2}{L_1}} \tag{2.10}$$

The results from this analysis can be explained intuitively. If k is small, then this is equivalent to opening the feedback loop, or eliminating the dependent source on the primary side of Figure 2-4. The dependent source on the secondary side, equal to sMI_1, is also equal to $k\sqrt{\frac{L_2}{L_1}}$. This arises from substituting $k\sqrt{L_1 L_2}$ for M and $\frac{V_1}{sL_1}$ for I_1. The resulting equivalent secondary coil circuit is shown in Figure 2-6. Here we can clearly see how the resonant low-pass filter arises. The prefactor, $\beta = k\sqrt{\frac{L_2}{L_1}}$, results from the non-ideal transformer pair formed by the two coils.[1] Figure 2-7 shows the Bode magnitude and phase plots for the transfer function $\frac{V_2}{\beta V_1}$ of the resonant low-pass filter. In this example, the tank has a resonant frequency of 13.5MHz and a Q of 25.

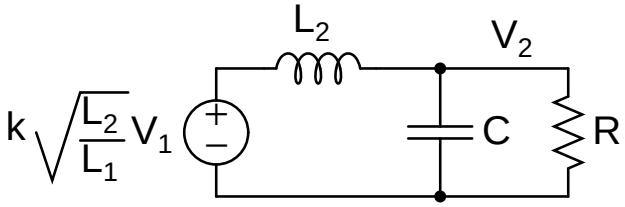


Figure 2-6: Equivalent secondary coil circuit.

[1]The ratio $\sqrt{\frac{L_2}{L_1}}$ is sometimes written in the case of a transformer as the turns ratio, $n = \frac{N_2}{N_1}$. This is equivalent to $\sqrt{\frac{L_2}{L_1}}$ because a coil of N turns nominally has an inductance $L \propto N^2$.

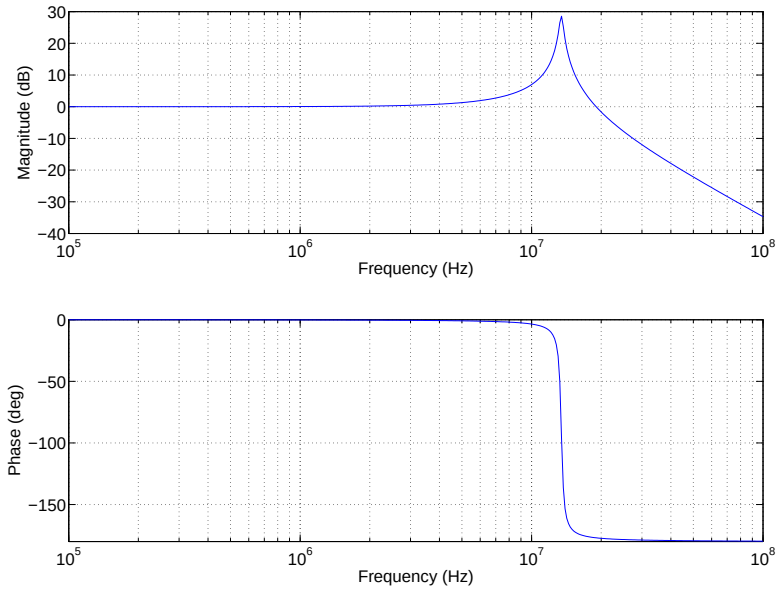


Figure 2-7: Bode gain/phase plot for a high-Q resonant low-pass filter.

2.4 Coil Design and Measurements

For this thesis project, both the primary and secondary coils need to be custom designed. The precise value of their inductance, size, and shape are important for the success of the wireless system. Figure 2-8 shows an overview of the functions of the primary and secondary coils in this system. As indicated, both coils will operate as part of tuned circuits. Use of resonance improves the spectral cleanliness of the output at the transmitter. At the receiver, resonance directly increases the voltage of the received signal by a factor of Q, the quality factor of the resonant tank.

2.4.1 Secondary Coil

First, let us examine the design of the secondary coil in this wireless system. Its design is more difficult than the primary coil because of the space constraint. The secondary coil must fit on the implant, which may not occupy more than 1.5cm^2 in area or weigh more than 1.5g. The implant must also fit the receiver chip, wire-bonds, battery, and surface-mount components. Because the size of the secondary coil directly affects receiver performance, a significant portion of this area should be

allocated to the receiver coil. I chose to build the secondary coil as a spiral trace on the receiver printed circuit board. Printed circuit board traces are robust to handling and fabricate reliably. Alternative types of inductors include wound coils and commercially available surface-mount inductors. Commercial inductors, however, usually do not make good receiver coils because they are designed to control their flux inside a packaging so as not to disturb other circuits. They may also contain heavy ferrite cores which could exceed the implant's weight specification. A hand wound coil may achieve higher inductance than a printed spiral because it could be wound tighter than a printed spiral, but the printed spiral has the advantage of reliable inductance and ease of fabrication, since the receiver must already use a printed circuit board to mount all parts. Accurate, predictable inductance values are essential for tuned circuits.

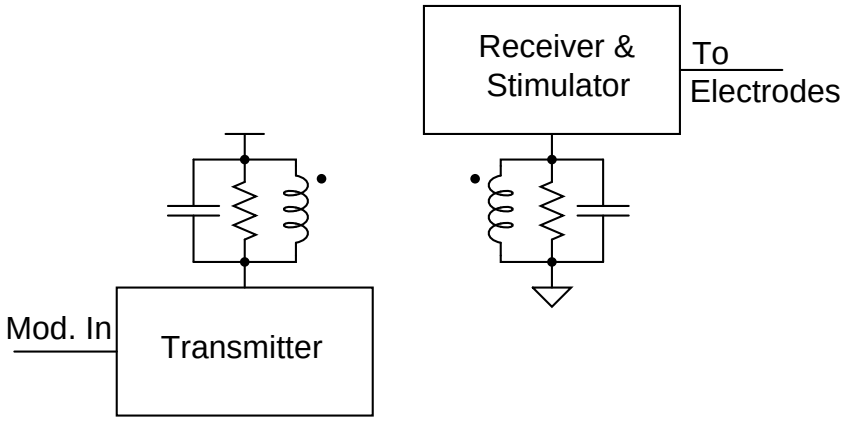


Figure 2-8: Overview of the primary and secondary coil functions.

Figure 2-9 shows two printed spiral inductors. The approximate dimensions of the spirals are 0.5 × 0.5cm for the smaller spiral and 0.5 × 1.0cm for the larger spiral. Measurements of the inductance and quality factor of each spiral as a function of frequency is shown in Figure 2-10. As it turned out, there was sufficient space for the larger printed coil in the final implant design. From the measurements, we see that the larger coil has approximately 140nH of inductance and a quality of factor of about 30 at 13.5MHz. A discussion of how to approximate the inductance of planar spirals appears in [12].

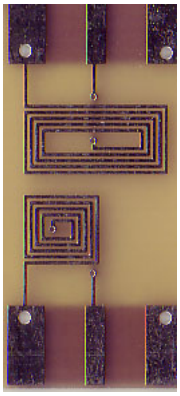


Figure 2-9: Two printed inductors.

2.4.2 Primary Coil

Given the restrictions on the design of the secondary coil, what is the optimal design of the primary coil? In many other types of neural recording and/or stimulation implants, such as cochlear implants and retinal implants, a transcutaneous inductive link is used to transmit both data and power to the implant. In this particular arrangement, the primary and secondary coils are placed as close together as possible[2], typically achieving a coupling coefficient, k, between 0.08 and 0.24 [18]. Further, the primary and secondary coils always operate at a fixed, known distance apart. In this project, the primary and secondary coils operate at a varying distance apart.

Let us consider the case where the primary and secondary coils are on axis. Equation 2.11 gives an approximation for the k value for on-axis loops of wire of radii r_1 and r_2 separated by a distance x [7]. Such explicit calculation of k is only possible for very simple geometries.

$$k(x) \approx \frac{r_1^2 \cdot r_2^2}{\sqrt{r_1 \cdot r_2} \cdot \left(\sqrt{x^2 + \max(r_1, r_2)^2}\right)^3} \tag{2.11}$$

If we plot the k factor as a function of separation x for primary loops of various radii r_1 while holding the secondary loop fixed with $r_2 = 1$cm, we obtain the plot in Figure 2-11. The results indicate that unity coupling is only possible for loops of wire

[2]In cochlear implants, the coils are separated by the thickness of the skin, typically about 7mm.

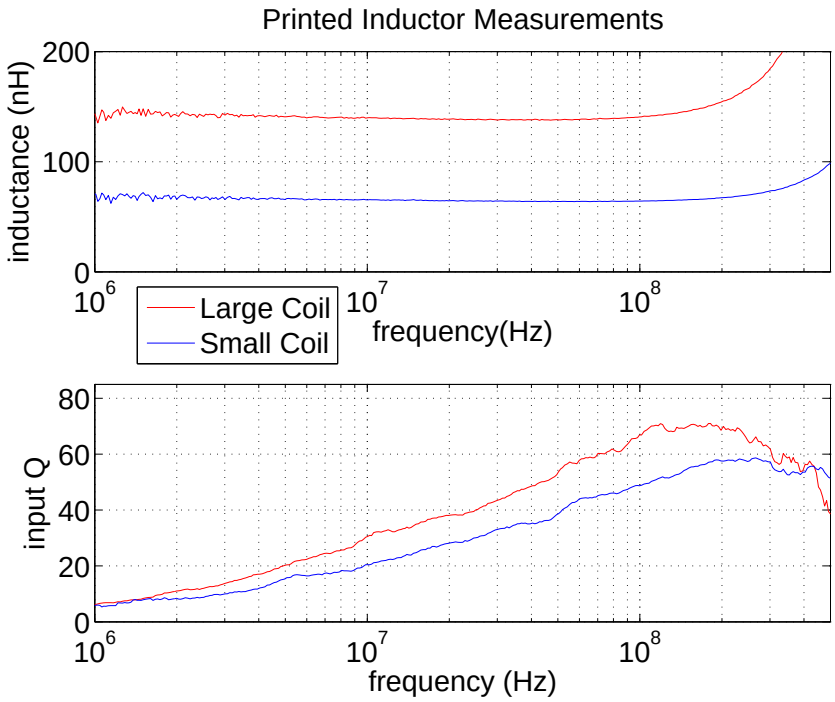


Figure 2-10: Measurements for small and large printed inductors.

of identical radius separated by 0 distance. This is intuitive because at 0 separation, the field generated by the primary is approximately uniform. Therefore, the coupling factor is just the ratio of the areas of the two coils. A more interesting effect occurs at larger separations, which is closer to where the transmitter-receiver pair in this thesis are likely to operate. At large separations, loops which are mismatched in geometry performed better than loops of similar size. This can be explained as a reduction in flux density, i.e. spreading of the magnetic flux to occupy a larger region. Smaller primary loops have their flux concentrated over a smaller region, and hence have a larger k factor at small separations. However, k falls rapidly as separation increases. In contrast, the flux density of large loops of wire decays gracefully. This property will be useful in this design where the coil separation can change by several centimeters.

One can observe empirically from Figure 2-11 that if one is constrained to a system with a fixed, small secondary size, then the radius of the primary loop that maximizes the coupling factor at a given distance is approximately equal to that operating distance. Assuming that this observation holds for other coil configurations besides single loops of wire, such as spiral inductors, we now have a simple criterion

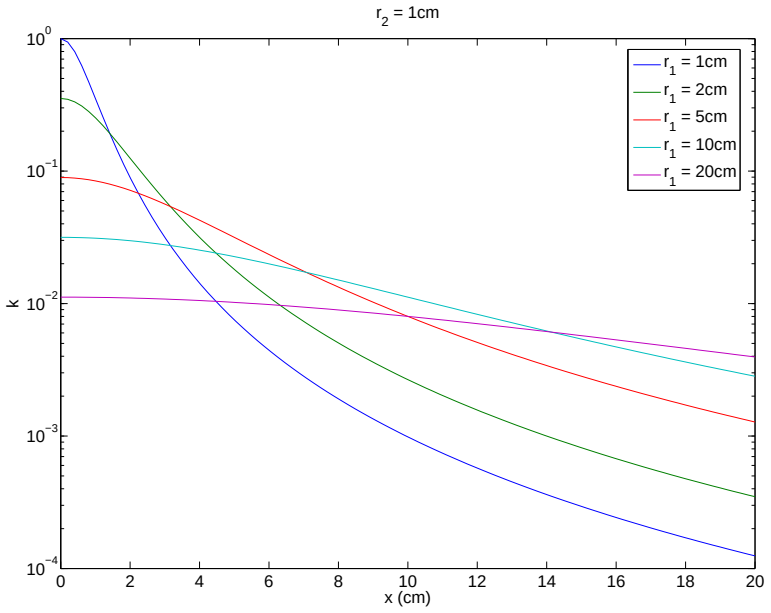


Figure 2-11: Theoretical coil coupling curves for two on-axis loops of wire of radius r_1 and r_2 as a function of separation x. The radius r_2 is fixed at 1cm while r_1 is varied.

to design our primary coil. If we assume that the average operating distance between the primary and secondary coils is about 10cm, then we should choose a primary coil radius of 10cm. Due to practical constraints on the size of the primary coil, the actual primary coil radius used in this thesis is 8.6cm. Figure 2-12 shows the measured inductance and quality factor of the transmitter coil as a function of frequency.

Table 2.1 summarizes the important features and design parameters of the primary and secondary coils.

Table 2.1: Summary of primary and secondary coil parameters.

Coil	**Radius**	**Inductance**	Q **at 13.5MHz**	**Self-Resonance Frequency**
Primary	8.6cm	700nH	65	70MHz
Secondary	0.2cm (equivalent)	140nH	30	350MHz

2.4.3 Coupling Factor Measurements

In order to validate the analysis in this chapter that k is very weak, we must substantiate that claim with some laboratory measurements. The following measurements

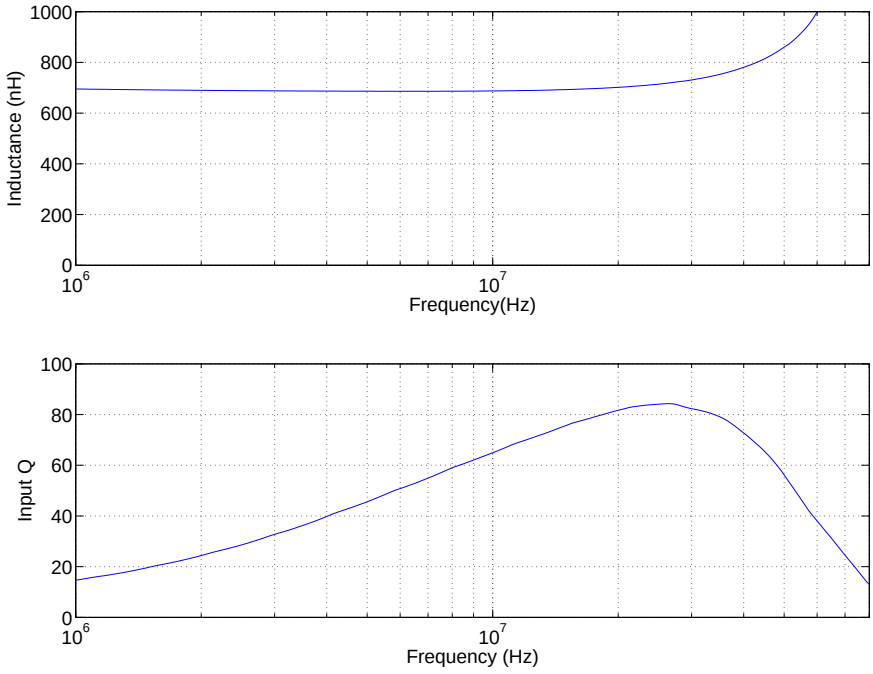


Figure 2-12: Measured inductance and quality factor for the transmitter coil as a function of frequency.

for k were taken at the operating frequency of 13.5MHz. Figure 2-13 shows the experimental setup for measuring the coupling coefficient. The small secondary coil, denoted as the probe in the figure, was positioned centrally over the transmitter coil. The amplitude of the transmitter output was measured independently to define V_1, the voltage level across the transmitter coil. The probe was then moved through the magnetic field generated by the transmitter. Its position is characterized by three parameters: r, the radial distance of the probe from the transmitter coil's geometric center, h, the height of the probe above the transmitter coil, and θ, the angle the probe makes with transmitter coil. An angle of $\theta = 0$ corresponds to both coils having identical, parallel orientation.

The amplitude of the signal received at the probe was measured with an Agilent 4395A spectrum analyzer. Although the transmitter was designed to be resonant to transmit at a single frequency, the probe in this experiment was designed not to be resonant. In theory, resonance contributes a certain amount of gain that can be

calibrated out. In practice, however, this calibration would be unreliable because in high Q systems, if the resonance of the probe is slightly off from the transmitter's resonance, this could result in a large error.

If we denote the voltage on the primary as V_1 and the voltage on the secondary as V_2, then we can use Equation 2.9 with one modification to determine k. Since the resonance of the secondary coil has been removed, Equation 2.9 simplifies to:

$$\frac{V_2}{V_1} = k\sqrt{\frac{L_2}{L_1}} \tag{2.12}$$

which can be solved for k once the ration $\frac{V_2}{V_1}$ has been measured.

The results of the field measurements, performed by sweeping r at several different heights for two different probe angles are shown in Figure 2-14. Immediately we see that the $k \ll 1$ assumption is valid, since k ranges from about 2×10^{-4} to 2×10^{-3} depending on height, radius, and angle. We also see that this coil arrangement is actually relatively insensitive to all 3 parameters. This is because of the significant size mismatch between the transmitter and probe coils. The large transmitter coil generates a relatively uniform magnetic field. Because the probe is so small compared to the transmitter, any movements of the probe do not significantly alter the strength of the magnetic coupling. The received signal strength curve, as a function of radius, shifts towards smaller values of k as the probe height increases. This can be explained intuitively by a near-far argument. At greater heights, radial changes as a fraction of probe height decrease. The "view" of the transmitter coil as a function of r, as seen by the probe, does not change significantly at larger heights. The coupling factor's relative insensitivity to probe position and tilt is an important feature of our design because the bird is free to move about the cage.

One interesting observation is the peaking in the strength of the coupling in Figure 2-14(a) for $h = 1$cm. The peaking around $r = 8$cm occurs because the transmitter coil radius is 8.6cm. The magnetic probe is very close to the turn of wire on the transmitter coil. Once the probe is moved past the transmitter radius, the magnetic field strength falls off sharply. This peaking effect only occurs for small values of h. The

peaking is a purely local effect that can also be masked by tilting the probe. The peaking effect has also been noted in [15] and [19].

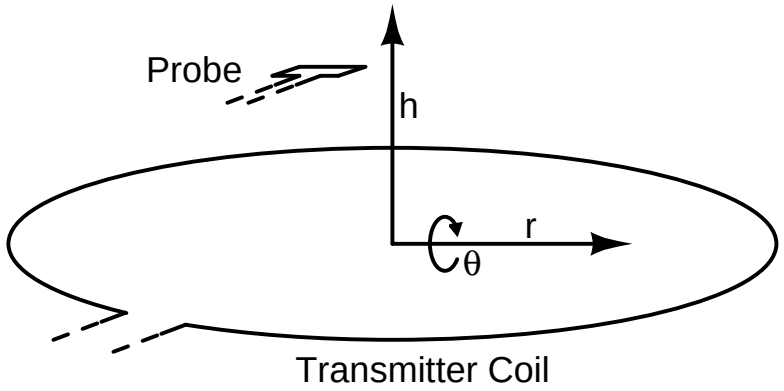
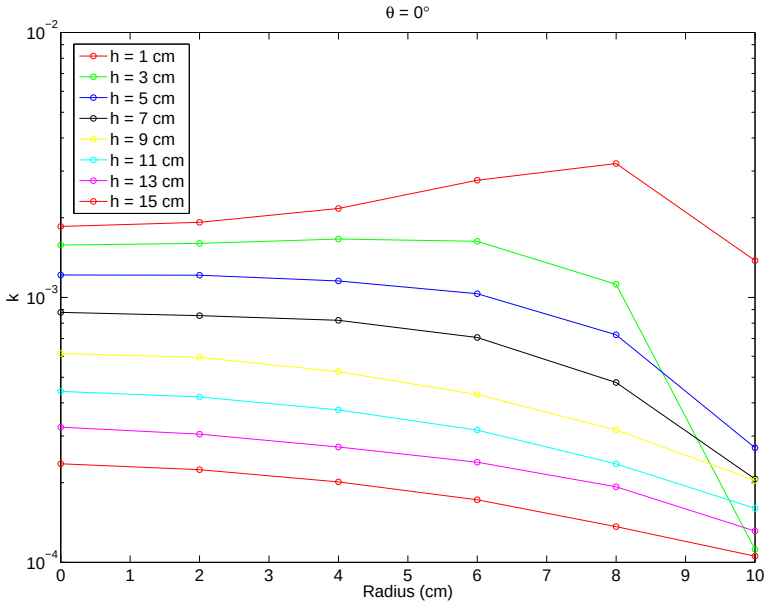
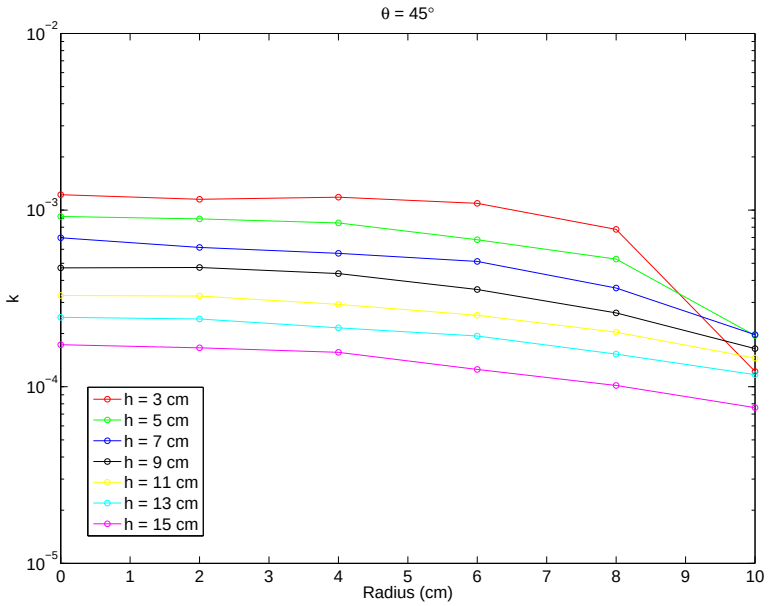


Figure 2-13: Testing arrangement for field measurements.



(a) Magnetic coupling measurements for $\theta = 0°$.



(b) Magnetic coupling measurements for $\theta = 45°$

Figure 2-14: Magnetic coupling measurements as a function of radius for different heights and tilt angles.

Chapter 3

Wireless Transmitter

As discussed earlier in the chapter on coupled coils, wireless communication occurs via a system of magnetically coupled coils. The transmitter coil is large and sits outside the bird cage. Its job is to produce a time varying magnetic flux which is gathered by the secondary coil on the receiver. We would like to transmit with a large voltage swing across the primary coil since the voltage swing on the secondary coil, and therefore the transmitter range, is directly proportional to the swing on the primary. The voltage received at the secondary will be directly amplified and demodulated.

From these design considerations, we realize that we need an amplification circuit capable of swinging large voltages across the transmitter coil. To this end, we will be interested in resonant amplifier circuits since transmission will take place at a single carrier frequency. Resonant circuits are capable of providing the voltage amplification that we need to make this wireless transmission work.

3.1 Class C Power Amplifier

The Class C power amplifier shown in Figure 3-1 is a very simple, single device power amplifier designed for efficient operation. This circuit is the same as a Class A amplifier. The only difference is the conduction angle, Φ, or the angular fraction of a period when the current through the power device is non-zero. In small signal amplifiers, the

device is biased at a particular DC operating point, and then driven with a gentle AC signal to avoid the nonlinearities in the device from causing distortion at the output. As long as the input drive is kept small, the voltage swing across the load remains proportional to the voltage swing at the input. In power amplifier applications, Class A does not refer to the linearity of the amplifier, but to the device conduction angle, – 360°. Since the load is tuned, it filters the drain current extensively. Even under large input voltage drives which would not normally be considered small signal, the output voltage of the tank may remain sinusoidal with an amplitude that is still reasonably proportional to the input drive amplitude. However, because there is a large DC current being burned in the amplifier which is not being delivered to the load, a large amount of the available power is wasted. The maximum Class A amplifier efficiency is 50%.

In this thesis, to call the transmitter a power amplifier is actually somewhat of a misnomer. In contrast to more conventional RF power amplifier applications, where the amplifier is designed to deliver a certain amount of power to a resistive load in the form of an antenna, this transmitter does not have an antenna.[1] Rather, the coil *stores* energy in a magnetic field. Loading is caused by the inductor's finite Q, or loss due to resistance in the inductor. Also, a small amount of energy is captured by the secondary coil and used for data recovery. This is how data is transmitted through near-field magnetic coupling. Any far field radiation is negligible. Both of these losses are small which results in a high load impedance seen by the power device. Hence, the power amplifier is mainly a large signal voltage amplifier.

Our specifications require large voltage swings at the output. Rough estimations suggest that 30-40V_{pp} is needed to obtain reliable transmission at 15cm separation. If we use the circuit in Figure 3-1, we expect that to get this level of output, we will need a very large input drive. Further, the device may be destroyed if it is biased with a large DC current. Since the amplifier will be operated in simple on/off fashion,

[1]In some designs, the transmitter coil may be driven at the frequency where it radiates, and hence loads the resonant tank with some radiation impedance. The coil therefore acts as both an inductor and an antenna. Given the driving frequency of 13.5MHz, one would not ordinarily expect a coil 17.2cm in diameter to radiate at this frequency.

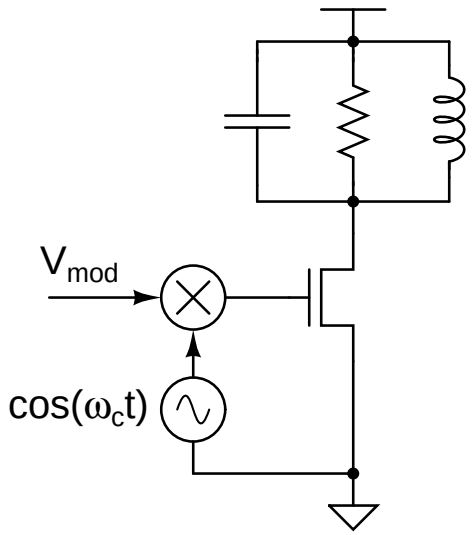


Figure 3-1: Class C power amplifier schematic.

linearity is not important. The solution is then to lower the DC operating point of the device so the zero input DC drain current is reduced to zero while keeping the input drive magnitude large enough to maintain the desired output swing. Lowering the DC operating point of the device such that it no longer conducts for 360° results in class AB operation for angles greater than 180° but less than 360°, class B operation for conduction angles exactly equal to 180°, and class C operation for conduction angles less than 180°.

In class C operation, the device is nominally off and only conducts for brief periods in each cycle when the input exceeds the transistor's threshold. When the transistor conducts, it acts as a current source, drawing spikes of current from the supply and passing it through the impedance of the LC tank. If the drive is periodic at the resonant frequency of the tank, the tank, which has maximum impedance at resonance, develops a large voltage across its terminals. At other frequencies away from resonance, the tank increasingly looks like a short circuit. The high frequency components of the drain current produce only small voltages at the output. The maximum possible efficiency of the class C power amplifier, η_{max}, as a function of conduction angle is calculated in [10] and shown in Equation 3.1. This equation is also valid for other classes of operation. It may be evaluated for any conduction

angle between 0 and 360°. Figure 3-2 shows a plot of the maximum power efficiency possible versus conduction angle.

$$\eta_{max} = \frac{\Phi + \sin\Phi}{4 \cdot \left(\sin\frac{\Phi}{2} - \frac{\Phi}{2}\cos\frac{\Phi}{2}\right)} \tag{3.1}$$

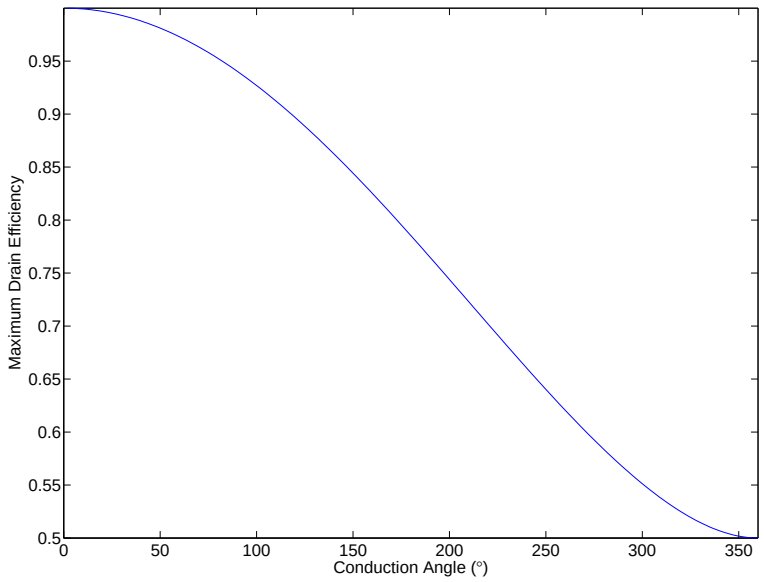


Figure 3-2: Maximum Power Amplifier Efficiency versus Conduction Angle.

This maximum efficiency curve makes several unrealistic assumptions. It assumes that the device behaves as a current source (remains saturated) even for $v_{DS} = 0$. It also assumes that the peak-to-peak output swings are a full $2V_{DD}$, and that there are no other losses in the circuit.

We have used this circuit successfully in lab to generate large peak-to-peak swings at the output, occasionally as large as 60V$_{pp}$. Although the circuit lacks the ability to modulate the output voltage swing in a linear manner, such accurate control is not needed for on/off AM. Simply switching the amplifier drive from full power to off is effective for transmitting on/off amplitude modulated data. However, one important drawback of this circuit is that it is not capable of oscillating autonomously. The drive must be at the transmit frequency of 13.5MHz, mixed by the modulation input V_{mod}. In a laboratory setting, high frequency generators are expensive and not often

available to dedicate to one application. Since the overall data rate is comparatively low (in the tens of kbps), a low cost, low frequency, programmable signal generator can supply just the modulation input to the amplifier. The amplifier may be a combination oscillator-transmitter described in the next section.

3.2 Oscillator-Transmitter (Transmixer)

A Colpitts oscillator solves the problem of the oscillator-transmitter. The circuit can be tuned to oscillate at 13.5MHz. Next, a data input can be used to turn the oscillator on and off. Hence, the oscillator is simultaneously a transmitter and a mixer.

The schematic of the oscillator, with the modulation input, is shown in Figure 3-3. Operation of the oscillator can be explained as follows: For now, ignore the transistor M_1 and the input V_{mod}. The remaining circuit is the Colpitts oscillator. Resistors R_1 and R_2 set a DC bias at the base of Q_1. Capacitor C_B is chosen large enough that the base of the transistor is at AC ground. The oscillator is essentially a common-base amplifier with positive feedback. Although drawn somewhat strangely, the combination of L, C_1 and C_2 is actually a resonant tank in the collector of Q_1. This configuration is an impedance transformer, sometimes referred to as a tapped-capacitor. Ignoring the "tap" connecting C_1 and C_2 to the emitter of Q_1, the configuration is more recognizable as a parallel tuned load with a resonance determined by L and the parallel combination of C_1 and C_2. Thus, the frequency of oscillation is given by:

$$f_0 = \frac{1}{2\pi\sqrt{L\frac{C_1C_2}{C_1+C_2}}} \tag{3.2}$$

In order to guarantee that oscillations will start up, we must determine an expression for the loop transmission, and choose component values such that the loop transmission is greater than one. If we view the circuit as a unidirectional common-base amplifier, we can see that any voltage that appears at the collector of Q_1 is fed back to the emitter via the "tap" on the capacitive voltage divider due to C_1 and C_2.

The tapped-capacitor structure also causes any impedance in the emitter to appear in parallel with the load, but transformed up by a factor of n^2, with:

$$n = \frac{C_1 + C_2}{C_1} \tag{3.3}$$

This impedance transformation only occurs near resonance. A more complete discussion of the tapped-capacitor is in [10].

The loop transmission at resonance is then given by:

$$(\omega_0) = g_m \cdot \left(R_L \parallel n^2 \cdot \left(R_E \parallel \frac{1}{g_m} \right) \right) \cdot \frac{1}{n} \tag{3.4}$$

where g_m is the small-signal transconductance of Q_1. It is acceptable in this case to use the small-signal g_m because we are simply testing for a loop transmission of greater than unity to ensure start-up. The signal amplitudes initially are small, which justifies this approach. The oscillation itself, of course, cannot be described using only small-signal analysis.

In this design, the impedance looking into the emitter of Q_1 is $\frac{1}{g_m}$, except at low frequencies. This is because at 13.5MHz, the base biasing resistors have been adequately bypassed by C_B, and therefore do not contribute to the emitter input impedance. At a standard bias current of 1mA at room temperature, this impedance is about 25Ω. Most practical values of R_E will be much larger than 25Ω. The total impedance at the emitter node is $\left(R_E \parallel \frac{1}{g_m} \right)$, but under these conditions it is approximately $\frac{1}{g_m}$. Likewise, most practical values of R_L will be quite large, even in comparison to $n^2 \cdot \frac{1}{g_m}$. An explicit R_L need not be used. It can be due entirely to the equivalent parallel impedance of the LC tank at resonance. For example, a 1μH inductor with a Q of 50 at 13.5MHz has a parallel equivalent impedance of 5kΩ. For values of $n < 4$, we can also ignore R_L in calculating the load impedance, since it is set by the lower impedance, $n^2 \cdot \frac{1}{g_m}$. In this situation, the loop transmission simplifies to:

$$L\left(\omega_0\right) \approx n \tag{3.5}$$

Table 3.1 shows a table of design values used in this Colpitts oscillator.

Table 3.1: Colpitts Oscillator Design Values.

Parameter	**Value**
Oscillation Frequency	13.5MHz
Q_1	2N2222
M_1	NDS351AN
R_1 and R_2	10kΩ potentiometer
C_b	1nF
R_L	not used
R_E	330Ω
L	700nH
C_1	220pF, tunable
C_2	680pF
n	4

3.3 Modulation

Unlike the Class C power amplifier, which we had to drive at the carrier frequency, the oscillator-transmitter is capable of generating its own carrier frequency. With the Class C amplifier, we could amplitude modulate the carrier simply by alternately driving and grounding the gate of the MOS transistor in Figure 3-1. How does one modulate the amplitude of the output of the colpitts oscillator? The most reliable, linear way of changing the output amplitude is to modulate the supply voltage. However, this is impractical and unnecessary because we are only interested in on/off modulation. The solution we propose to this problem is the addition of M_1 in Figure 3-3.

M_1 is operated as a switch. If its gate is grounded, then it behaves as an open circuit and is effectively removed from the circuit. If the gate is driven high, the switch turns on and creates a low impedance path to ground, discharging C_B. As the base of Q_1 is brought to ground, the oscillation quickly dies. When M_1 is opened, C_B begins to charge through R_1 and R_2. As Q_1's base returns to its normal operating point, the oscillation restarts.

This technique has been effective in laboratory tests at data rates exceeding

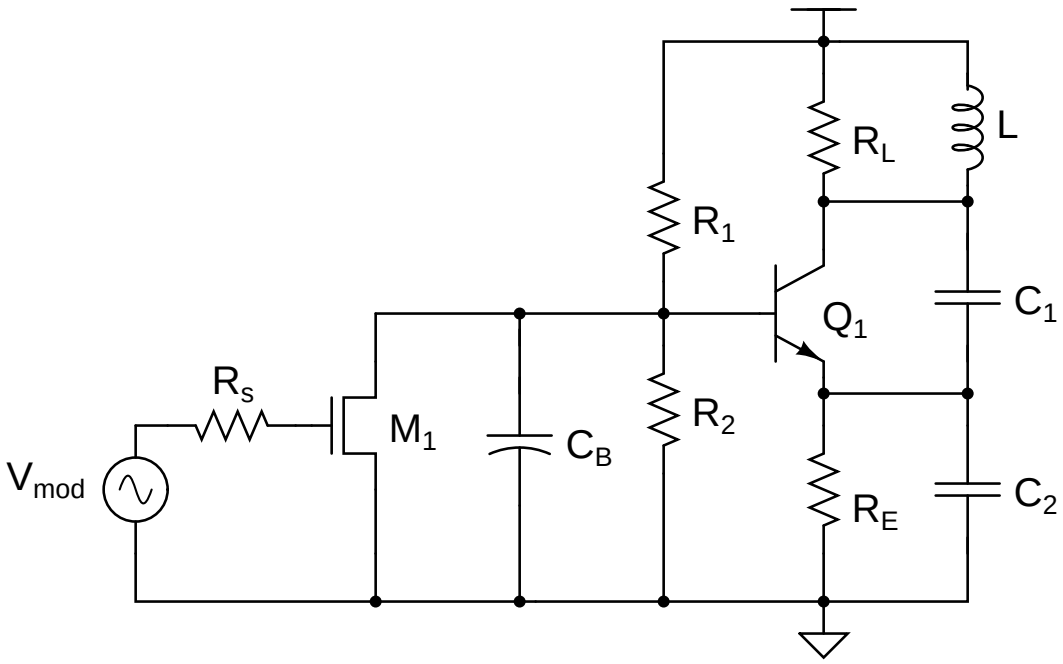


Figure 3-3: Colpitts Oscillator-Transmitter Schematic.

100kbps. However, pushing the data rate higher, which is not currently required by this application, would require some additional design modifications. The main limitation on the data rate is how quickly the oscillation can restart. The time-constant associated with C_B is $\tau = (R_1 \parallel R_2)\, C_B$. Reducing this time constant can increase the data rate. However, one must be careful not to raise the high-pass pole too high, since it is important to keep the base at AC ground. A similar modulation technique was reported in [21].

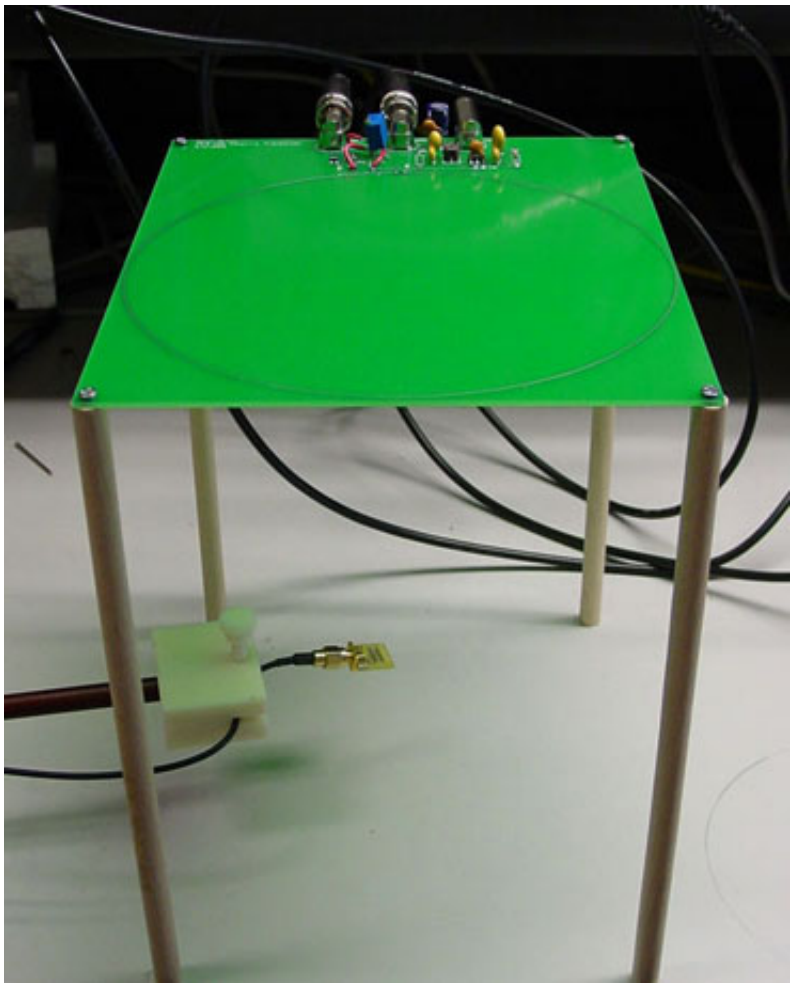


Figure 3-4: Photograph of the wireless transmitter. The transmitter coil is a single-turn PCB trace.

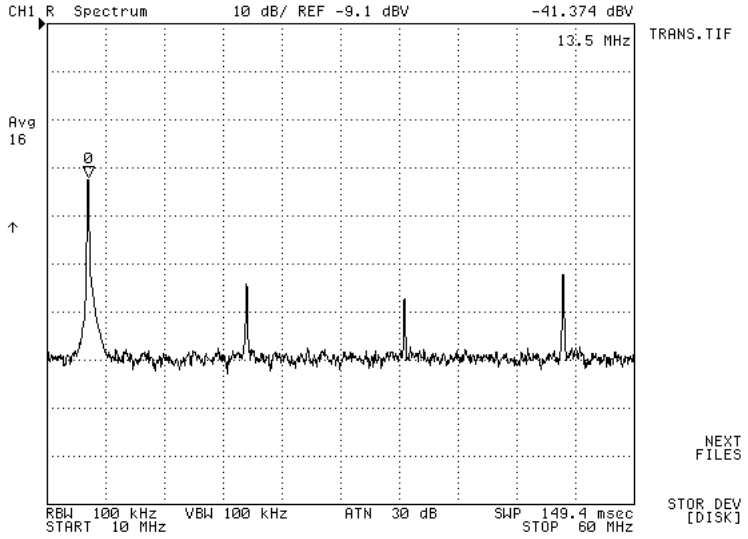


Figure 3-5: Transmitter power spectrum showing sidebands from data modulation. The first harmonic is about 23dB below the carrier.

Chapter 4

Electrode Modeling and Measurement

4.1 Introduction

The purpose of this chapter is to review some basic principles of electrodes and electrochemistry that are needed for understanding neural stimulation. Electrodes provide the interface between the electrical and chemical domains. In neural stimulation, the purpose of the electrode is to interface the stimulator circuitry with the neural tissue. Neural stimulation uses charge balanced, biphasic pulses of current to artificially induce neurons near the electrode to fire. In particular, we would like to know what impedance the electrode will present to the current drivers.

Electrochemical cells always consist of at least two electrodes. In experiments, the electrode under observation is termed the working electrode, and the secondary electrode for carrying the return current is called the counter electrode. Sometimes a third reference electrode is used as well. Implanted neural stimulation electrodes may be modeled by an electrochemical cell consisting of the stimulating (working) and counter electrodes, and tissue, which may be modeled as a 0.9% NaCl electrolyte (physiological saline). In Section 4.2, we will look at electrode modeling, and in Section 4.3, we will look at some electrode measurements. A more detailed treatment of electrodes as they apply to neural stimulation may be found in [9, 14], and a more

general reference on electrodes and electrochemistry may be found in [1].

4.2 Electrode Modeling

4.2.1 The Randles Model

The Randles model is a composition of electrical circuit elements that ideally behave identically to a real electrode in solution. However, this model is limited to small-signal excitations. Unfortunately, constant current drives for neural stimulation do not typically fall into the category of small-signal. Nevertheless, the Randles model is a good starting point for our investigation.

Figure 4-1 shows the general Randles model of the electrode. The model details two current paths, I_c and I_f. I_c is the current that charges the double-layer capacitance, C_{dl}. The other current path, I_f, represents the Faradaic current. Faradaic currents correspond to the oxidation or reduction of chemical species. The flow of Faradaic current is limited by the sum of two impedances, R_{ct} and Z_W. Physically, R_{ct} is the charge transfer resistance, or the impedance opposing an electron trying to transfer from the surface of the electrode into the electrolyte, and vice-versa. The Warburg impedance, Z_W, represents limitations on mass transport by diffusion inside the solution. The Warburg impedance is also frequency dependent. R_s is a general solution resistance, set by the conductivity of the electrolyte itself. Both the Faradaic and charging currents must pass through the solution resistance.

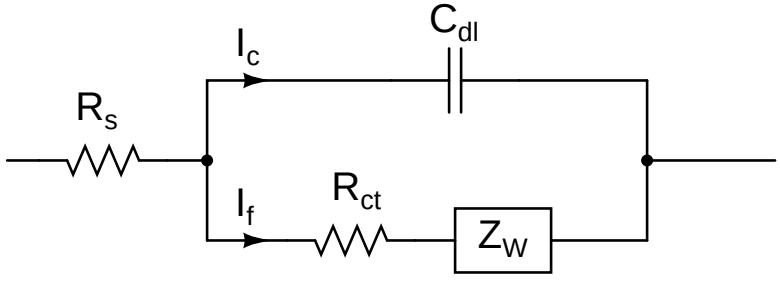


Figure 4-1: Randles electrode model.

4.2.2 The Double-Layer Capacitance

One of the most significant features of electrode impedance is the double-layer capacitance. To get a general idea of how the Faradaic and charging currents interplay, first consider the Randles model of the electrode, shown in Figure 4-1. If a sufficiently large DC step in potential is applied to the working electrode in an electrochemical cell, initially a large current, I_c, flows to charge the double-layer capacitance. The charging time constant is set by the product of the double-layer capacitance, C_{dl}, and the series solution resistance, R_s. In fact, the charging of this capacitance is what actually causes polarization of the electrode. Once the capacitance charges appreciably, the Faradaic current due to redox processes, I_f, begins to take over and eventually dominates the total electrode current. After a few RC time constants, the double-layer charging current falls below detectable levels.

Physically, the double-layer capacitance arises due to a diffuse charge region in the electrolyte near the surface of the electrode. The geometry is similar to that of a parallel plate capacitor, where the top plate is the electrode itself, and the bottom plate is formed from charges in solution piling up near the electrode. The separation between the "plates" is on the order of molecular scales; hence the double-layer capacitance can be quite large. The parallel plate view of the double-layer capacitor is known as the Helmholtz model.

Unfortunately, the Helmholtz model is not accurate because it predicts C_{dl} is a constant. It also does not account for the fact that real charge distributions in solution do not take the shape of a parallel plate, but rather are most heavily concentrated near the electrode, and then taper off. One can define the thickness of the diffuse charge region as where the strength of the electrostatic forces fall below that of thermal processes. The Gouy-Chapman capacitance model takes into account the average distance of the separation between charges on the surface of the electrode and charges in solution. This improved model also takes into account the fact that as the electrode becomes more heavily polarized, the double-layer capacitance actually increases. The diffuse charge region shrinks under strong polarization, which results

in an increased capacitance. However, the double-layer capacitance does not increase without bound as the applied potential increases because the charge separation cannot become infinitely small. The Stern model improves the Gouy-Chapman model to account for this limit.

For the purposes of estimating the double-layer capacitance of an electrode, we can simply apply the Gouy-Chapman formula in Equation 4.1 with the applied potential, ϕ_0, set to zero. C^* is the bulk concentration of the electrolyte in moles per liter. In this thesis, all experiments are performed in 0.9% NaCl solution, which has a molarity of 0.153. Therefore, we can calculate from Equation 4.1 the double-layer capacitance to be $89\mu F/\text{cm}^2$. For the 50μm diameter electrodes used in this thesis, we can expect $C_{dl} \approx 2\text{nF}$.

$$C_{dl} = 228zC^{*1/2}\cosh\left(19.5z\phi_0\right)\mu F/\text{cm}^2 \tag{4.1}$$

4.2.3 The Solution Resistance

The solution resistance arises because all the electrode current, both Faradaic and charging, must pass through the solution to the cell's counter electrode. One model for the solution resistance, the hemispherical model, is derived in [14]. The result is presented here in Equation 4.2:

$$R_{hemi} = \frac{\rho}{2\pi a_0} \tag{4.2}$$

where ρ is the resistivity of the solution and a_0 is the radius of the working electrode. It is assumed that the counter electrode has a significantly larger radius than the working electrode. A typical value for the resistivity of the physiological saline used in this thesis is $85\Omega\cdot$cm. For the 50μm diameter electrodes, the hemispherical model predicts about $5.5\text{k}\Omega$ of solution resistance. From this equation, we see that the solution resistance scales directly with electrolyte resistivity and inversely with electrode radius.

An alternative model of the solution resistance is the disk model. Instead of the

electrode current spreading from a hemisphere, it spreads from a disk. The disk model, given in Equation 4.3 has the same dependencies as the hemispherical model, but is a factor of $\pi/2$ greater. The disk model predicts a solution resistance of 8.6kΩ.

$$R_{disk} = \frac{\rho}{4a_0} \tag{4.3}$$

4.2.4 The Warburg Impedance

The Warburg impedance is one component of the impedance that sets the Faradaic current. Usually, it dominates the charge transfer resistance, i.e., most chemical reactions are mass transfer limited rather than charge transfer limited. Physically, the Warburg impedance arises from a diffusion current in the cell. As reactants in the cell are either oxidized or reduced, a concentration gradient arises in the cell due to a depletion of reactants near the electrode-electrolyte interface. This gradient is what allows for mass transport. More reactant will diffuse towards the interface to fuel the reaction. The Warburg impedance can be solved for by assuming a small, sinusoidally varying input to a semi-infinite RC transmission line. The Warburg impedance has a $1/\sqrt{\omega}$ dependence on frequency.

4.2.5 Low and High Frequency Limits

At low frequencies, the Warburg impedance and the double-layer capacitance compete to set the impedance of the cell. The impedance of C_{dl} rises linearly with decreasing frequency, but the Warburg impedance rises only with the square root of the decreasing frequency. Therefore, at extremely low frequencies, C_{dl} is effectively an open circuit, and Z_W in series with R_{ct} and R_s determine the total electrode impedance. Eventually, Z_W will also dominate R_{ct} and R_s. Strictly speaking, according to the Randles model, the Warburg impedance should have a constant phase angle of 45°. However, due to a number of simplifications in the Randles model, this is often not the case.

At high enough frequencies, the impedance of C_{dl} falls below the impedance of Z_W

and effectively shorts out both components of the Faradaic impedance. The electrode impedance then appears as the solution resistance R_s in series with the double-layer capacitance C_{dl}. Figure 4-2 shows the Randles high frequency model. Electrode drive under constant current usually involves biphasic pulses which are quick enough to fall into the high frequency category.

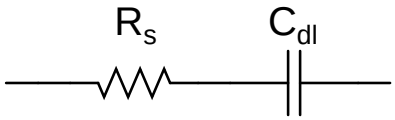


Figure 4-2: Randles high frequency electrode model.

4.3 Measurement Techniques for Electrochemical Systems

Any electrochemical system may be characterized in a number of different ways. However, all techniques involve either controlling the voltage in the cell and measuring the current, or controlling the current and measuring the voltage. Within each technique, it is possible to measure the cell step response to changes in current or potential, or measure steady state response to low frequency or AC sweeps of current or potential, or any combination of the previous. However, in any electrochemical measurement, two electrodes are required. An incomplete cell with only a single electrode cannot be measured. This presents a problem because in an electrochemical system, the two electrodes are always in series. Without being able to "probe" the middle node between the two electrodes, any measurement of cell potential is just the difference of the two electrode potentials, and cannot be attributed in part to either electrode. The potentiostat, described in the next section, is an instrument which solves this problem. Fortunately, having the electrodes in series also means that measuring the current in either electrode is sufficient because both electrodes must carry the same current.

4.3.1 Controlled Potential

The potentiostat is a common instrument for controlling the potential of the working electrode in an electrochemical cell. Figure 4-3 is a schematic of a simple potentiostat. In the electrochemical cell in this figure, the working electrode is depicted as a circle and the counter electrode is shown as a flat plate. The counter electrode is present because without it, the cell would be incomplete and no current could flow. The third electrode in the cell, represented with an arrow, is the reference electrode. The reference electrode is what allows us to effectively "probe" inside the cell between the two electrodes.

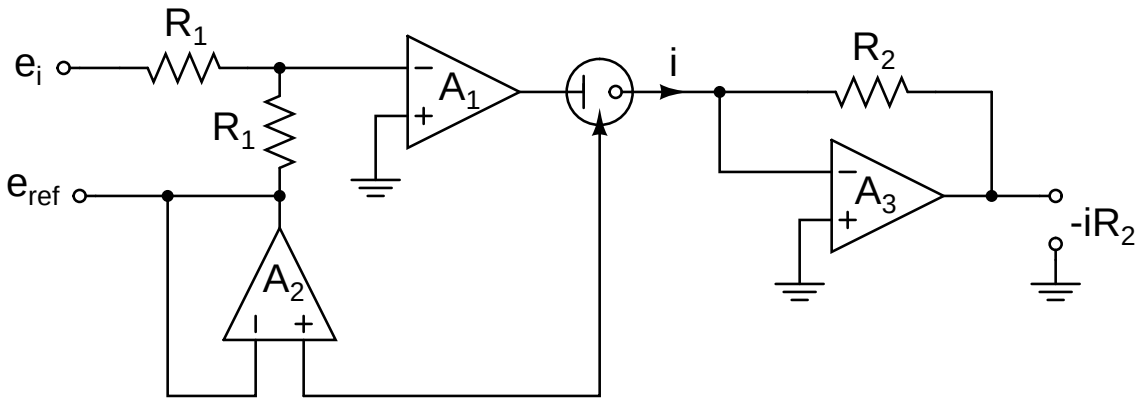


Figure 4-3: Schematic of a potentiostat.

The potentiostat controls the potential of the working electrode *with respect to the reference electrode.* The reference electrode allows us to adjust the potential of the working electrode with respect to the reference without the need to control the potential of the counter electrode directly. By setting a desired potential between the working electrode and the reference, the potentiostat will drive the counter electrode to whatever potential is necessary to support the current flowing through the working electrode. The current flowing through the working and counter electrodes are necessarily the same because the reference electrode is designed such that it does not participate in any chemical reaction and therefore carries no current. Additionally, its reference potential is independent of composition of the cell.

To control the potential of the working electrode, we apply the desired potential,

e_i, to the input. Since the inverting input to A_1 is a virtual ground, the circuit automatically drives the counter electrode in such a way that the voltage sensed by the reference electrode becomes $e_{ref} = -e_i$. The reference electrode should not carry any current since any reaction taking place at the reference electrode would change its potential. The buffer A_2 ensures that no current is drawn through the reference electrode. The op-amp A_3 actively holds the working electrode at virtual ground. In this regard, the potential of the working electrode is equal to e_i with respect to the reference electrode. The amplifier A_3 converts the cell current to a voltage so that it may be measured.

Controlled potential methods are especially useful for performing cyclic voltammetry. In cyclic voltammetry (CV), the potential of the working electrode is swept over some range while the cell current is monitored. Frequently, for low levels of polarization, negligible current is observed until a critical potential is reached, at which point an observable current corresponding to the oxidation or reduction of a chemical species begins to flow. Sweeping the potential through a complete cycle including the opposite direction can reveal information about the stability of products.

4.3.2 Electrode Impedance Spectroscopy

Unfortunately, CV is not directly applicable for characterizing the impedance of stimulation electrodes. Cyclic voltammetry is mainly useful in characterizing electrochemical systems that create products with DC currents, and where measurements of steady state currents versus potential are important. However, electrodes for neural stimulation are not used in this manner. In fact, a permanently polarized electrode can be extremely damaging to the electrode and the patient. In neural stimulation, electrodes are briefly charged by the stimulation current, then discharged in the opposite direction to remove all the transferred charge to ensure that there is no net DC current through the cell. A more effective technique than CV for determining cell impedance is with an AC technique called electrode impedance spectroscopy.

While the three electrode system is important for swept potential methods such as CV, where large voltages can appear even across the counter electrode, a two elec-

trode system is sufficient for electrode impedance spectroscopy. Electrode impedance spectroscopy is a small signal measurement that applies small sinusoidal excitations swept over a range of frequencies to the cell. By measuring the current drawn, one can determine the electrode impedance as a function of frequency.

Under small signal drive, the polarization of the electrode is minimal. Therefore, we can dispense with the reference electrode and apply the potential difference to the cell directly if we are absolutely sure that the all the applied potential is dropped across the working electrode. To meet this condition, one can design the counter electrode to be much lower impedance than the working electrode. This is usually accomplished by building the working electrode out of a relatively inert material, such as platinum, and making its surface area much larger than the surface area of the working electrode. Since the current through the counter electrode and working electrodes is the same, the counter electrode will have a much lower current density and hence potential drop than the working electrode.

4.3.3 Controlled Current

The final electrode measurement technique is controlled current. Since the implanted electrodes will be driven with biphasic pulses of current, the best way to characterize them in the laboratory is to drive them in the same manner that they will be used in the field. In constant current experiments, the working electrode is driven by a constant current, and the resulting back voltage that appears across the pair of electrodes is measured. Again, a large area counter electrode is needed to ensure that the measured potential difference is entirely attributable to the working electrode. The biphasic pulses of current are expected to last about 100μs, which means that the Randles high frequency model should be in effect.

Figure 4-4 shows example current and back voltage waveforms for electrodes under biphasic constant current drive. When the pulse of current initially starts, a voltage change of $-IR_s$ is observed, followed by a charging slope of $\frac{-I}{C_{dl}}\text{V}\cdot\text{s}^{-1}$. When the current switches direction, the back voltage moves by $2IR_s$, and the charges with a slope of $\frac{I}{C_{dl}}\text{V}\cdot\text{s}^{-1}$. When the current shuts off, if the pulse was precisely charge

balanced, the back voltage would return to 0.

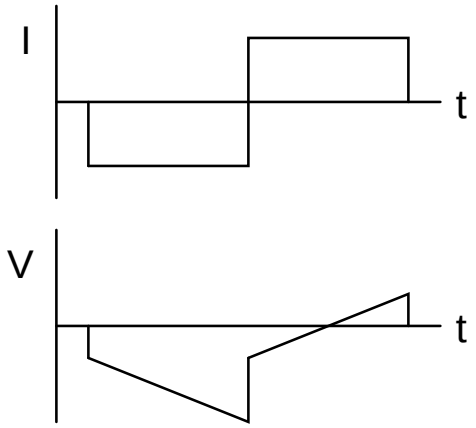


Figure 4-4: Example current and back voltage waveforms for electrodes under biphasic constant current drive.

4.4 Experimental Results

Next we took the experimental setup in Figure 4-5 into the laboratory and performed electrode impedance spectroscopy and biphasic amperometry. The working electrode is a piece of 50μm diameter stainless steel or platinum-iridium wire clipped at the end. The counter electrode is a section of platinum with a large surface area. The impedance of the working and counter electrodes will appear in series; however, the impedance will be dominated by the working electrode because the counter electrode, with its very large relative area, is much lower impedance.

An Agilent LCR meter was used to make impedance measurements from 20Hz to 1MHz for the platinum-iridium and stainless steel electrodes in 0.9% NaCl. The 0.9% NaCl is sometimes referred to as physiological saline because it approximates the composition of living tissue. Since the driving signal is small, the electrodes are never appreciably polarized. Hence most of the measured current can be attributed to charging the double-layer capacitance. Figure 4-6 shows the measured electrode impedance spectra. The four panels show the electrode impedance magnitude and angle as a function of frequency, as well as the decomposition of the impedance into an equivalent series RC circuit in an attempt to fit the data to the high frequency

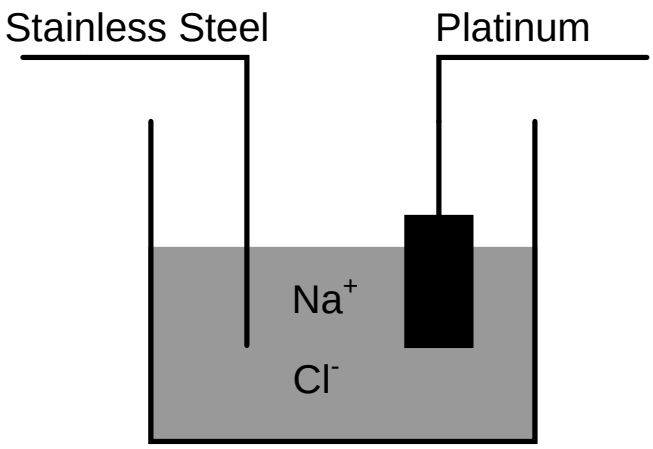


Figure 4-5: An example electrochemical cell.

Randles model. At a frequency of 10kHz, both electrodes show close to 1nF of capacitance (C_{dl}). The resistance varies, with both being about 20kΩ (R_s).

Next, the electrodes were measured under constant current drive using current driver circuitry that will be discussed in Chapter 5. Figure 4-7 shows measured back voltage waveforms of the working electrode under biphasic current drive. Note the similarity to the ideal waveform in Figure 4-4. The rounding of the corners is due to finite R_{ct} and Z_W. Each phase of current was 50μs in duration. Table 4.1 shows the extracted C_{dl} and R_s parameters from the stainless steel electrode measurements, and Table 4.2 shows the extracted parameters for the platinum-iridium electrodes. Note that R_s remains fairly constant as current level is increased, but C_{dl} actually increases. This is expected because the electrode becomes slightly polarized as the charge builds up, thus increasing the capacitance. The measured values of R_s and C_{dl} are also close to the predicted values of $R_s = 8.6$kΩ for the disk resistance model and of $C_{dl} = 2$nF nominally according to the Gouy-Chapman double-layer capacitance model.

Table 4.1: R_s and C_{dl} measurements for 50μ Stainless Steel Electrodes under constant current drive.

Level	R_s	C_{dl}
30μA	14.4kΩ	2.5nF
60μA	13.2kΩ	3.4nF
90μA	12.6kΩ	4.0nF

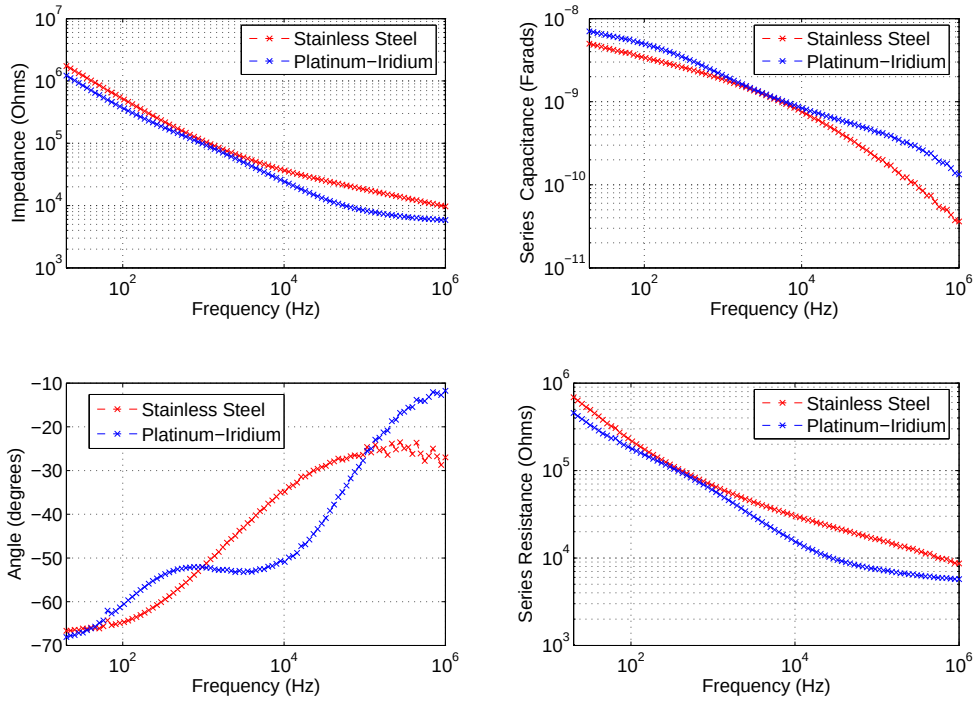


Figure 4-6: Pt-Ir and stainless steel electrode impedance spectra.

Table 4.2: R_s and C_{dl} measurements for 50μ Platinum-Iridium Electrodes under constant current drive.

Level	R_s	C_{dl}
10μA	6.4kΩ	3.6nF
50μA	5.8kΩ	4.4nF
100μA	4.8kΩ	7.3nF
150μA	6.0kΩ	9.0nF

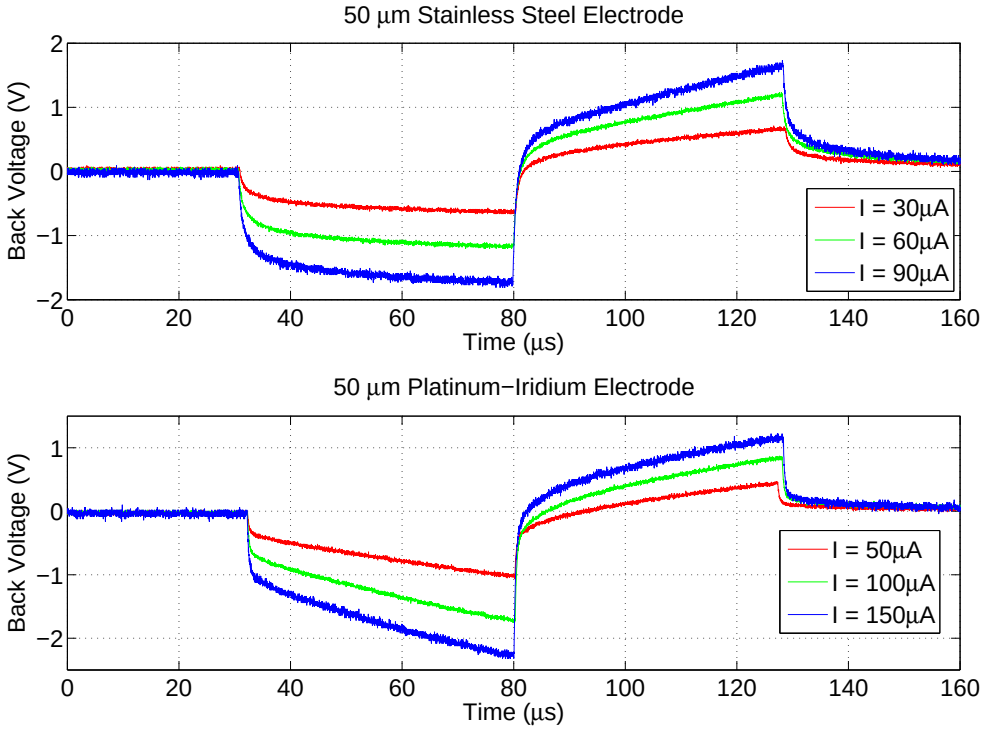


Figure 4-7: Biphasic pulses of electrodes under constant current drive.

Chapter 5

Electronic Integrated Circuits

5.1 Introduction

The core of this thesis project centers around the integrated circuits which I designed and were built in a standard 0.5μm N-well CMOS process. Integrated circuits are desirable because a single chip can implement a wide variety of functions at very low power. In particular, integrated circuits are necessary to meet the size and power requirements of the implantable system. Figure 5-1 shows a block diagram of the various parts of the implant. It consists of an off-chip receiver coil followed by a number of on-chip integrated circuits for amplifying the RF signal, demodulating data, decoding the data and applying the electrical stimulation to the electrodes implanted in the bird's brain. In this chapter, I will describe the design and experimental results of each circuit.

In all schematics, an N-well CMOS process is assumed, i.e., all NMOS bulk terminals are part of a common, grounded substrate.

5.2 High Speed Amplifier

As discussed in Chapter 3 on the wireless transmitter, data is modulated onto a 13.5MHz carrier prior to transmission. Figure 5-2, reprinted from Chapter 2 shows the transmitter and received coil arrangement. Unfortunately, the received signals

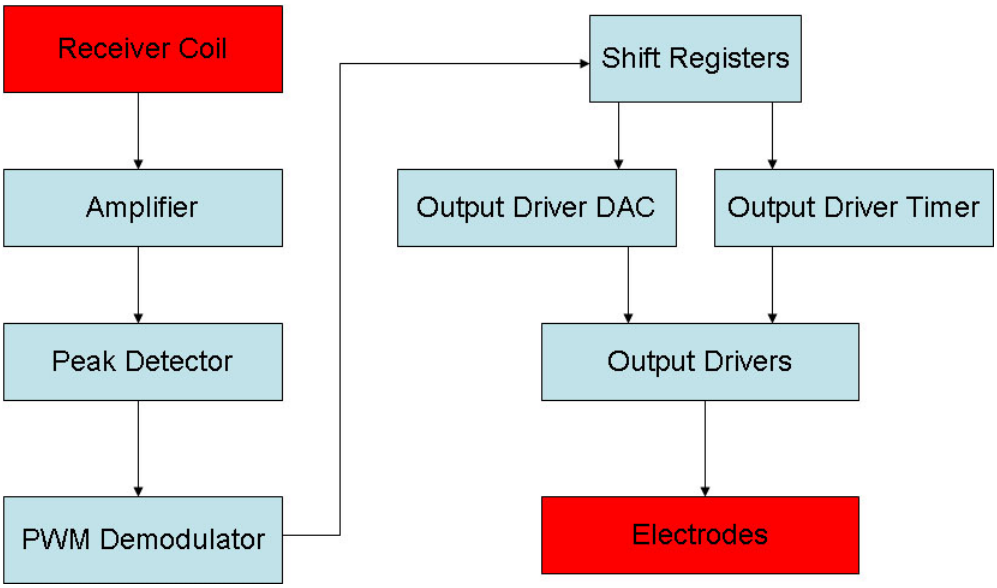


Figure 5-1: Block diagram of the implant. The red blocks are off-chip while blue blocks are implemented on-chip.

tend to be small – usually just tens of millivolts, even with the help from resonant amplification due to the passive LC network at the receiver input. Section 5.3 discusses the peak detector which will be used to recover the modulation input. However, due to a dead-zone nonlinearity of the peak detector, it is necessary to preamplify the received signal before it can be fed to the peak detector. The preamplifier should provide a gain of approximately 20dB to the received RF signal to ensure correct operation of the peak detector. The amplifier topology should be simple, low power, and have sufficient bandwidth to amplify at 13.5MHz.

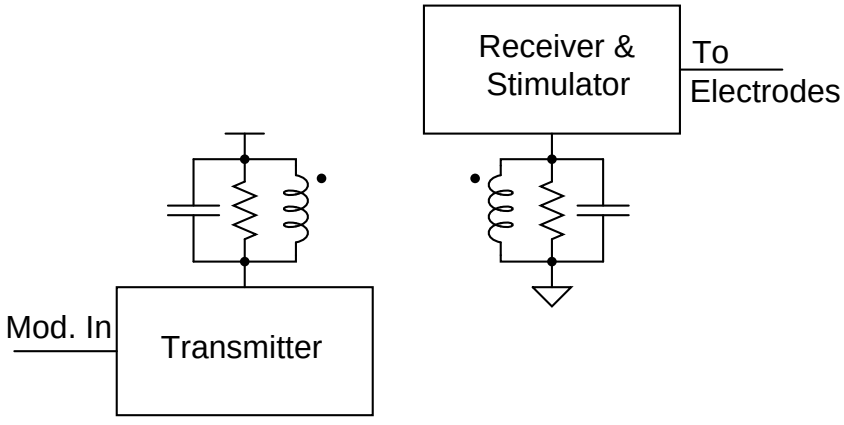
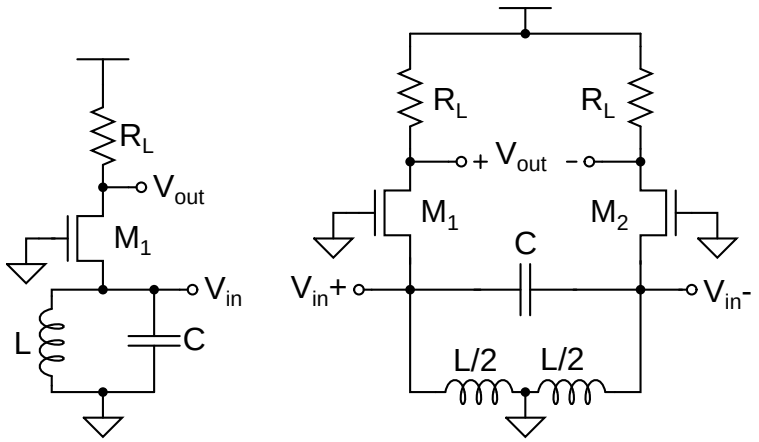


Figure 5-2: Overview of the primary and secondary coil functions.

Figure 5-3 shows single ended and fully differential versions of a common-gate amplifier. The advantage of the common-gate amplifier is its inherently high transconductance. Because the input is fed at the source and not the gate, both the gate and back-gate transconductances add to the total transconductance. An NMOS transistor is used because of its high mobility compared to PMOS transistors.



(a) Single Ended (b) Fully Differential

Figure 5-3: Common-gate Amplifiers

The design employed in this thesis uses the single ended version of the amplifier, shown in Figure 5-3(a). The main reason the differential version was not used is because it is difficult to create a completely symmetric inductor on a printed circuit board with access to the electrical centerpoint. Since one of the goals of this project is to minimize the size and weight of the implant, taking advantage of printed circuit board inductors is helpful towards achieving that goal. However, due to large PCB feature sizes, a compact differential inductor is not practical. Nevertheless, the single ended version is also adequate. The main advantage of the differential amplifier is that it permits full wave peak detection, which would naturally have lower ripple than a half wave version. Although differential signal processing provides other benefits such as excellent common mode rejection, this design would not be able to take advantage of that since the preamplifier simply serves to help drive the peak detector. Peak detectors inherently operate on single ended inputs.

Biasing of the amplifier in Figure 5-3(a) is provided by a current mirror to the gate of M_1. A bypass capacitor placed at the gate node ensures that the gate is at incremental ground. The external components L and C form a resonant low-pass filter to the node V_{in} at 13.5MHz. The inductor is also a short circuit at DC, setting the operating point of the source of M_1 at ground. This is important for the current mirror biasing the gate of M_1 to work correctly. The input impedance of the amplifier, which is approximately $1/g_s$, is large enough that it will not load the tank.[1] The gain of the amplifier is

$$A_v = (g_m + g_{mb}) \times R_L = g_s \times R_L \tag{5.1}$$

The bandwidth of the amplifier is set by the size of the resistive load, R_L, and the sum of all the capacitances at the output node. Those capacitances are the input capacitance to the next stage, C_i, and the small signal capacitances between the terminals of M_1, C_{gd} and C_{bd}. R_L in this design was selected to be 50kΩ, implying that any total capacitance less than approximately 200fF is acceptable. It should be noted that unlike actively loaded amplifiers, the bandwidth is fixed by the resistance and capacitance at that node. Increasing the amplifier bias current cannot improve the bandwidth, but it can increase the gain. The amplifier was designed for a nominal bias current of 15μA, giving a designed gain of roughly 20dB. The size of M_1 was 21.6μ/0.6μ. Figure 5-4 shows several amplifier gain curves under different bias conditions. Although the amplifier is DC coupled, measurements were AC coupled, and the filter cut-in around 100kHz can be observed. The amplifier exhibits approximately 80MHz of bandwidth. This indicates that the resistive load may be increased or M_1 may be widened in order to obtain the same gain at a lower bias current. The price is trading away some of the excess bandwidth. Table 5.1 is a summary of some parameters of the common-gate amplifier.

[1]The amplifier input impedance is on the order of a few thousand ohms while the tank output impedance is on the order of a few hundred ohms.

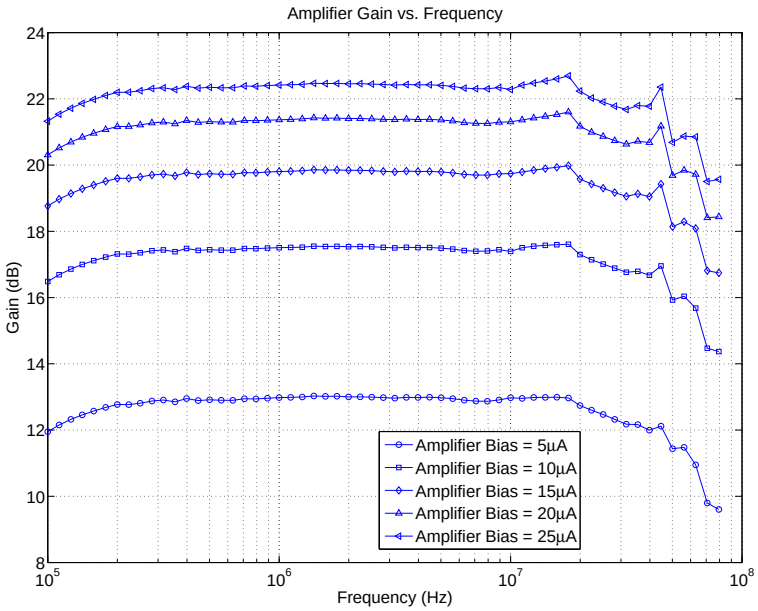


Figure 5-4: Measured amplifier gain curves under different bias conditions..

Table 5.1: Summary of Common-Gate Amplifier Parameters.

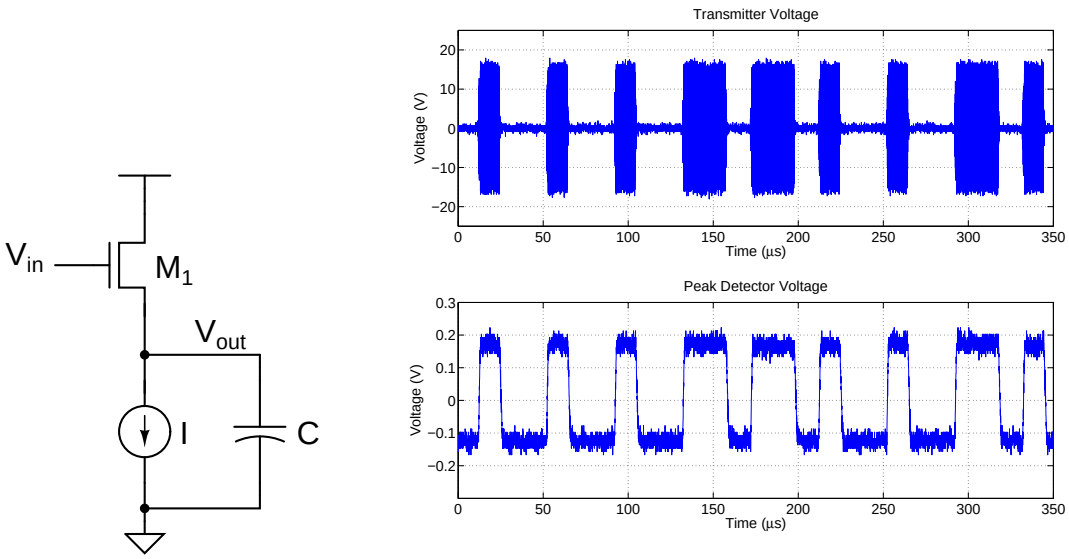
Parameter	**Value**
W/L of M	$21.6\mu/0.6\mu$
Bias Current	15μA
Gain	20dB
Bandwidth	80MHz

5.3 Peak Detector

The output of the RF amplifier is used to drive the input of the peak detector. The role of the peak detector is to distinguish the regions in time when the 13.5MHz carrier is on and when it is off, and output this information as a full scale digital signal. This effectively recovers the modulation input used at the transmitter. Figure 5-5(a) shows a simplified schematic of the peak detector circuit. The circuit operates as follows: when the input V_{in} rises, transistor M_1 turns on and charges the capacitor, C. When V_{in} falls, M turns off. The only discharge path for C is then through the current source, I. For the circuit to operate properly, the input amplitude of V_{in} should be on the order of 100mV or greater. This is necessary to drive M_1 into nonlinear operation. Otherwise, the circuit behaves as a simple source-follower. An alternative

way of thinking about this circuit is to assume that the amplitude of V_{in} is large enough that the current flow through M_1 is not symmetric with respect to each cycle. Therefore, the average current through M_1 will be greater than its bias current. The only way to balance this flow is for the average value of V_{out} to rise.

The choice of I and C depends on the frequency of the input V_{in}. The output of this peak detector is not constant. It will contain sawtooth ripples independent of the input amplitude due to the current source I discharging C at a constant rate. Increasing the W/L of M_1 increases the sensitivity of the peak detector since to charge C during positive cycles of V_{in}, M must overpower the discharge current, I. However, one must be careful not to design M too large because its input capacitance loads the output of the preamplifier. Table 5.2 summarizes some parameters of the peak detector design. Figure 5-5(b) shows an example of the peak detector in action.



(a) Simplified Schematic

(b) Observed Waveforms

Figure 5-5: Peak detector schematic and observed waveforms.

The peak detector may also be evaluated at DC for different input amplitudes. The curves in Figure 5-6 show the peak detector output voltage as a function of input voltage. It should be noted, however, that the input voltage is measured at the input to the preamplifier and not the peak detector itself. The preamplifier provides some

Table 5.2: Summary of Peak Detector Parameters.

Parameter	**Value**
W/L of M_1	$43.2\mu/0.6\mu$
I	1μA
C	5pF

gain to the input of the peak detector, which is needed to reduce the peak detector dead zone.

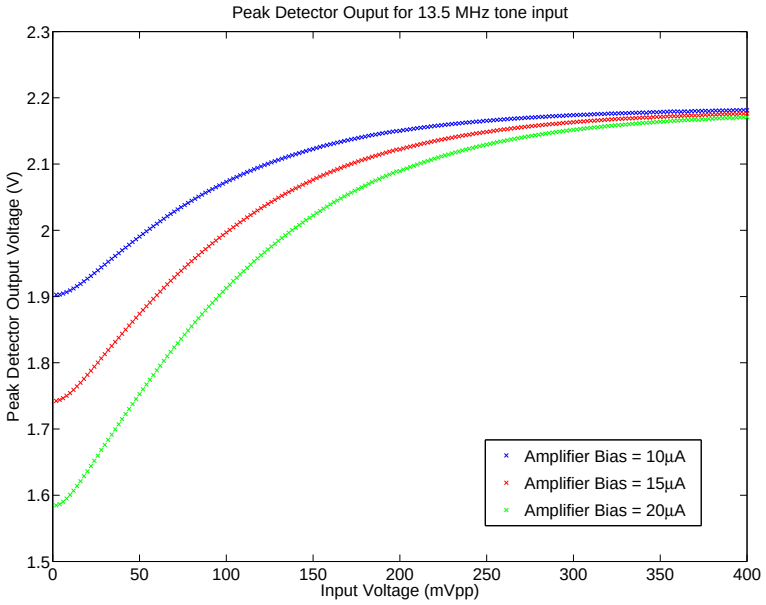


Figure 5-6: Measured peak detector characteristics under different amplifier bias conditions.

Although the peak detected voltage represents the transmitter modulation input, which is the desired signal, it is still not useful unless it can be amplified up to a full-scale digital signal. Conceptually, one would like to set a threshold voltage for comparison that is halfway between the minimum and maximum values of the detector output. A simple way to do this using feedforward is to low-pass filter the peak detector output and use that to set a comparison threshold. However, one problem with this approach is the large component values needed to get slow filter poles. Another problem occurs when the input is shut off for long periods of time. The peak detector output and low-pass filtered versions will converge on the same

value, resulting in rapid comparator triggering due to noise.

Figure 5-7(a) implements the low-pass filter concept. The time-constant of this filter is given by:

$$\tau = \frac{C}{g_m} \tag{5.2}$$

and the filter transfer function by:

$$\frac{V_{lp}}{V_{in}} = \frac{1}{\tau s + 1} \tag{5.3}$$

The problem of the comparator triggering due to noise can be removed by placing an intentional DC offset on either the transconductor of the low-pass filter or the comparator. Although this circuit is currently being used in this thesis, it could be improved. One serious problem is that the g_m-C low-pass filter is somewhat ineffective as a low-pass filter. To obtain a very slow filter pole of around a few hundred Hertz, the transconductor was biased with just 5nA of current. Still, this required a filter capacitor of 50pF. Peak detector output voltage swings routinely reach a few hundred millivolts, and this is easily enough to saturate the inputs to a subthreshold transconductor which only possesses a linear range of about 75mV. Hence, the internal nodes slew and the output never really becomes the average value of the input.

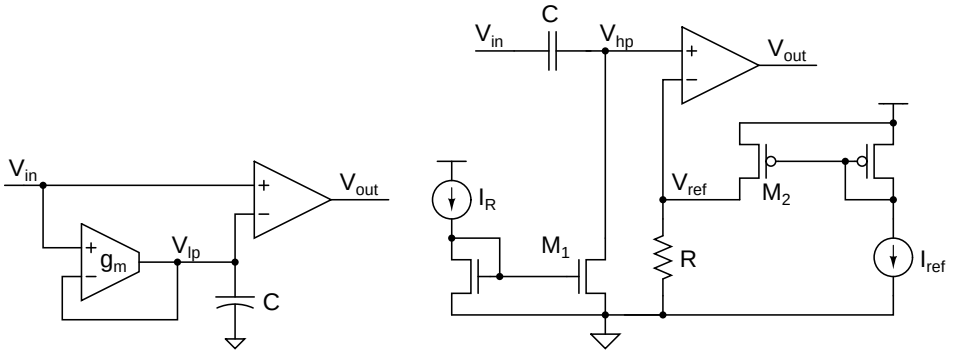
One possible solution is to replace the g_m-C filter with one made entirely of passive elements. However, the size of the required resistor and capacitor would occupy a very large area on chip. Figure 5-7(b) proposes an alternative method for amplifying the peak detector output to full scale. Instead of comparing the signal to its average value, Figure 5-7(b) proposes a a high-pass circuit that AC couples V_{in} to V_{hp} to be compared to a fixed, quiet reference voltage. M_1 acts as a resistance in this circuit to set the high-pass pole along with capacitor C. It is in the triode region because the drain node of M_1 is capacitive and therefore must be at the same potential for DC as the source node, which is grounded. The size of the resistance is set by the current I_R, and is therefore also tunable. Because V_{in} is not limited to just small

signal swings, M_1 may become saturated if its drain node swings 100mV or more above the source. However, M_1's output impedance transitions from a lower value in the linear region to a much larger value in the saturation region, thus lowering the high-pass pole cut-in frequency and actually improving the filtering, rather than worsening it. The high-pass time-constant is given by:

$$\tau = C \times r_1 \tag{5.4}$$

where $r_1 = 1/g_{ds}$ when M_1 is in the triode region and $r_1 = r_o$ when M_1 is in the saturation region.

Additionally, M_1 clamps V_{hp} to ground. The source of M_1 may be shifted above ground (and its mirror) to raise the DC input voltage to the comparator, if needed. The reference voltage is set by $I_{ref} \times R$ and is also tunable through M_2 and I_{ref}. The reference voltage may also be set to intentionally have a DC offset from the comparator's non-inverting input to prevent spurious triggerings when the input V_{in} is at 0.



(a) Level detector with feedforward adaptive threshold.

(b) Alternative level detector using AC coupling with adjustable offset.

Figure 5-7: Level Detectors

5.4 Pulse Width Demodulator

Stimulation data is encoded using pulse-width modulation (PWM). Although the peak detector downconverts the information encoded in the carrier to baseband, we still have not recovered any data. The resulting peak detected waveform must be passed through a pulse width demodulator to recover the desired data.

PWM is a return-to-zero strategy, meaning that every bit period is composed of a logic high phase and a logic low phase, in that order. Marks and spaces are distinguished by observing the width of the high period relative to the low period. In this design, a 50% duty cycle is defined as the boundary between a mark and a space. The demodulation circuit works by converting the duration of the high and low phases into voltages. At the end of each period, a comparator determines which voltage is greater. This is equivalent to determining whether the high or low phase was longer, and hence what the intended bit was. A bit sampling edge is provided for free by guaranteed transitions at the beginning and somewhere in the middle of each cycle of the modulated data. Once the data is decoded, on chip current drivers deliver the stimulation current.

The PWM demodulation circuit, implemented on chip, is shown in Figure 5-8. Operation of this circuit is as follows: Initially, all switches are open. The voltages on C_1 and C_2 are assumed to be zero. When the modulated data level goes high, switch ϕ_1 closes and charges C_1 with current I_1. When the data input changes from high to low, switch ϕ_1 opens and ϕ_2 closes, charging C_2 with I_2. In this design, $I_1 = I_2$ and $C_1 = C_2$ to set a crossover in the logic level of the demodulated data at 50% duty cycle. At the end of the bit period, switch ϕ_2 opens and ϕ_3 arrives to latch the output of the comparator and update V_{out}. After V_{out} has been latched, ϕ_4 briefly closes the two reset switches to return the voltages on C_1 and C_2 to zero.

Since the demodulation of each bit must take place in the span of one bit period, 4 clock phases must be generated from the 2 already available from the data. This is accomplished by creating two delayed versions of the data using the delay cell shown in Figure 5-9. The delay cell is a current starved inverter. When the input changes,

the inverter formed by transistors M_5 and M_6 switches slowly because it is current limited to I_{delay} by transistors M_2 and M_4. The output changes by charging C_{delay} with a constant current. The time it takes to switch is therefore set by the ratio I_{delay}/C_{delay} and the supply voltage. The delay period is set to be a small fraction of the bit period, i.e., 1μs out of a 20μs bit period. These delayed copies of the data input are used to create the charging signals ϕ_1 and ϕ_2 and the sampling and reset pulses ϕ_3 and ϕ_4. To accommodate these additional phases for sampling and reset, the long charging phases, ϕ_1 and ϕ_2, are equal to their respective high and low periods of the modulated data, less one delay period. Hence both the high and low charging phases are reduced by equal amounts, keeping the demodulation threshold at 50%. The full timing diagram for the PWM demodulator is shown in Figure 5-10. D_1 is the modulated data input, and D_2 and D_3 are delayed versions of the input created with the delay cell in Figure 5-9. The length of the delay period is exaggerated in the diagram. Table 5.3 shows design parameters for the PWM demodulator circuitry.

Table 5.3: Summary of PWM Demodulator Parameters.

Parameter	**Value**
I_1	50nA
I_2	50nA
C_1	1pF
C_2	1pF
I_{delay}	200nA
C_{delay}	100fF

Operation of this circuit in testing is shown in Figure 5-11. In Figure 5-11(a), the input voltage, the voltages on capacitors C_1 and C_2, and the demodulated output voltage can be seen for a 50kbps data rate. This strategy can also be used to demodulate PWM data at higher speeds. Figure 5-11(b) shows the same circuit re-biased to operate at 4Mbps. Due to the high data rate, it was not possible to observe the analog voltages on the charging capacitors. Instead, in addition to the modulated input and demodulated output, the clock signals ϕ_1 and ϕ_2 are shown. In this implementation, the charging switches ϕ_1 and ϕ_2 on chip are active low.

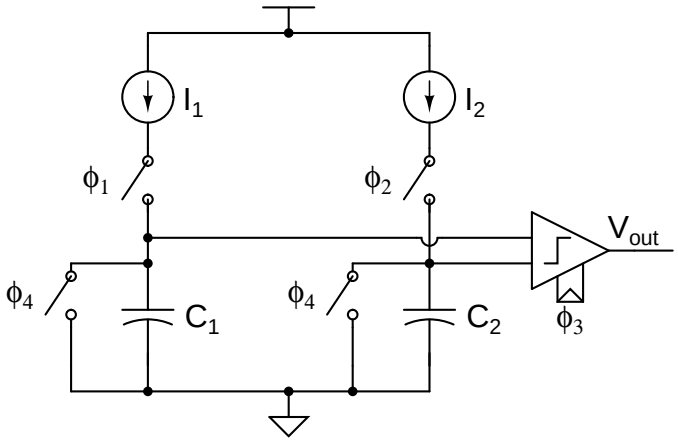


Figure 5-8: Simplified PWM Demodulator Schematic.

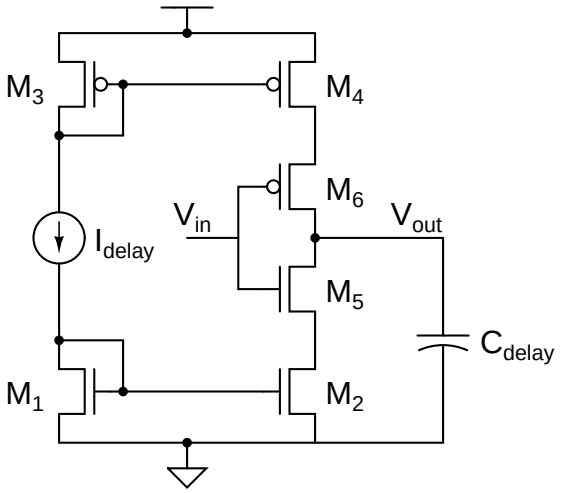


Figure 5-9: Analog Delay Cell.

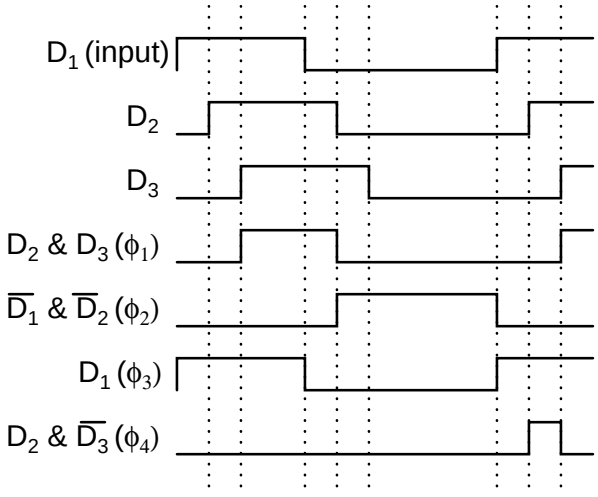
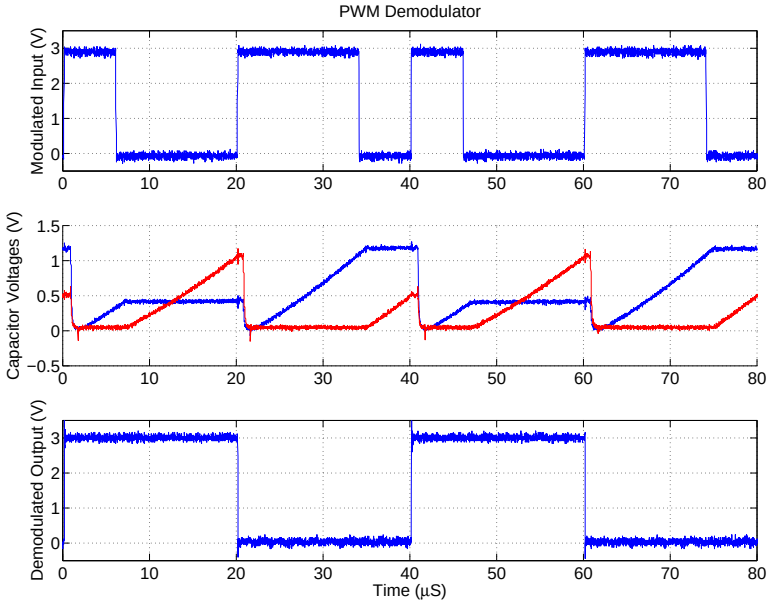
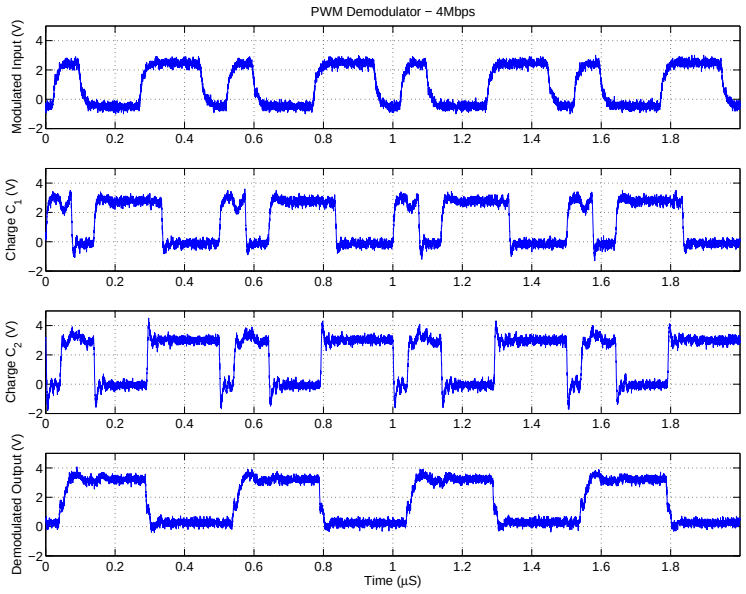


Figure 5-10: PWM timing diagram.



(a) PWM demodulator operating at 50kbps.



(b) PWM demodulator operating at 4Mbps

Figure 5-11: PWM demodulator circuit operating at two different data rates.

5.5 Output Driver Selection and Stimulation

The final step in the data processing chain is to take the demodulated data and decode it. The demodulated data is fed into a shift register using the falling edge of the modulated data as the clock (see Figure 5-11), to convert serial data into parallel words. These words configure a DAC which sets the stimulation current level for all the output drivers, selects a specific output driver, and triggers a timer circuit to initiate a biphasic current pulse on the desired electrode at the chosen current level.

Figure 5-12 shows a block diagram overview of the decoding and stimulation circuitry. The shift register contains 14 bits even though only 7 bits are actually used to select the output driver, phase and current level. The reason is to allow for a unique recognition sequence and additional spacer bits to prevent intended data from mimicking the recognition sequence. A recognition sequence is needed because the receiver and transmitter operate asynchronously. It is not known in advance when a stimulation command will arrive.

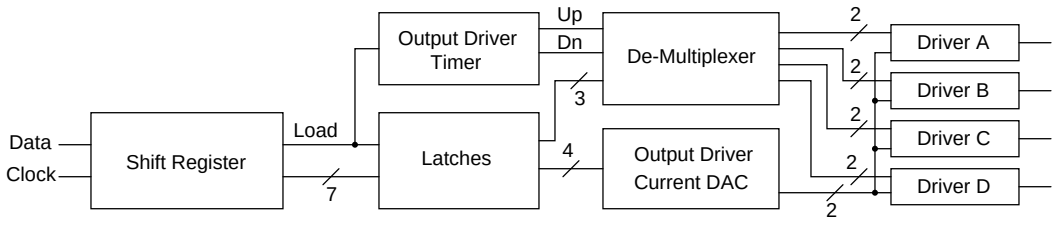


Figure 5-12: Complete output driving timing and selection circuitry.

The recognition sequence is a certain sequence of bits that are required to appear in order to indicate that a full set of stimulation parameters have been transmitted. Once the recognition sequence is observed, this triggers the *Load* signal to briefly become high. The load command moves a copy of the 7 desired data bits into the latches where they are stored until the next load command arrives. The latches enable the shift register to continue to operate and fill with the next set of parameters even while a stimulation is in progress. This is important because the latches must store stimulation parameters throughout the entire duration of the stimulation, which can last for many bit periods. A typical bit period may be approximately 40μs while a

stimulation may last 100μs per phase.

Figure 5-13 shows a general sequence of the 14 bit data packets. Bits D_0-D_2 form the recognition sequence, all 1's. D_3, D_6, D_9, and D_{12} are the spacer bits, all 0's. D_4 and D_5 are the data bits that select the stimulation site. D_7-D_8 and D_{10}-D_{11} select the current level, and D_{13} selects the phase ordering of the output pulse. Of the 7 actual data bits, 4 are used to configure a 16 level current DAC that sets the output current level. The remaining 3 are used in a de-multiplexer. Two of those bits are used to select 1 of 4 possible stimulation sites. The final bit is used to select the stimulation phase – either an anodic first or cathodic first pulse. The output driver timer times each phase and then shuts off the current. The timer circuit is essentially a cascade of two one-shots, of which the first is triggered by the load command to deliver the first pulse of current, and the second is triggered by the first one-shot finishing to deliver the second and opposite pulse of current.

D_{13}	D_{12}	D_{11}	D_{10}	D_9	D_8	D_7	D_6	D_5	D_4	D_3	D_2	D_1	D_0
P	0	C_4	C_3	0	C_2	C_1	0	S_2	S_1	0	1	1	1

Figure 5-13: General data packet sequence showing recognition sequence, spacer bits, and data bits.

5.5.1 Output Driver Cell

Figure 5-15 shows an output driver cell. There are four such cells on the chip – one for each electrode. The output of each cell is connected to an electrode. The cell is biased with the voltages V_P and V_N, set in common for all 4 cells across the chip by the DAC in Figure 5-14. The DAC output current, I_{out}, is set by I_{min} plus I_{ref} scaled by the 4-bit selection word, *Cur*. The output driver cell has switches ϕ_1 and ϕ_2 which can be closed one at a time to drive the output current I_{OUT} into the attached electrode. The timing and control of the switches is explained next in Section 5.5.2. Since the switches are not in the main current path, the on-resistance of the switches does not matter in determining the compliance of the output drivers. Only the $V_{DS,SAT}$ of the output transistors matter. A scaling factor of 1:10 in the output driver mirrors

also reduces current consumption in the biasing branch. The main output current ranges from 100μA to 1mA. Figure 5-16 shows measured biphasic current pulses demonstrating all 16 levels from 100μA to 1mA.

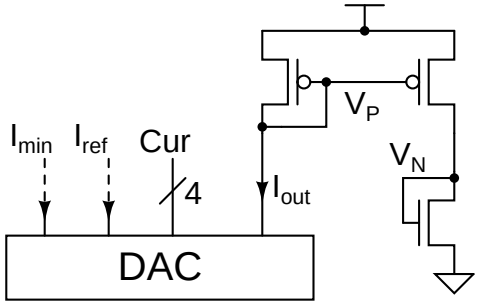


Figure 5-14: Output driver current DAC.

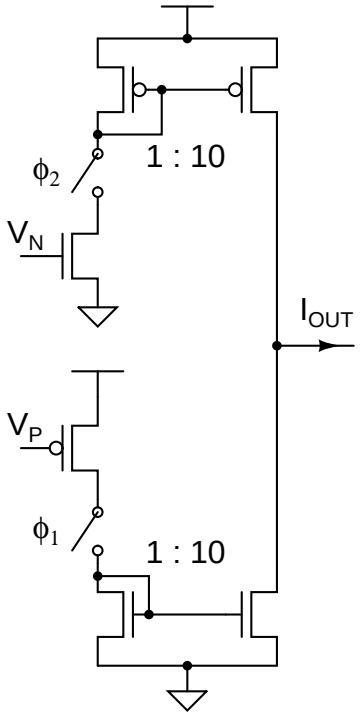


Figure 5-15: Output driver cell.

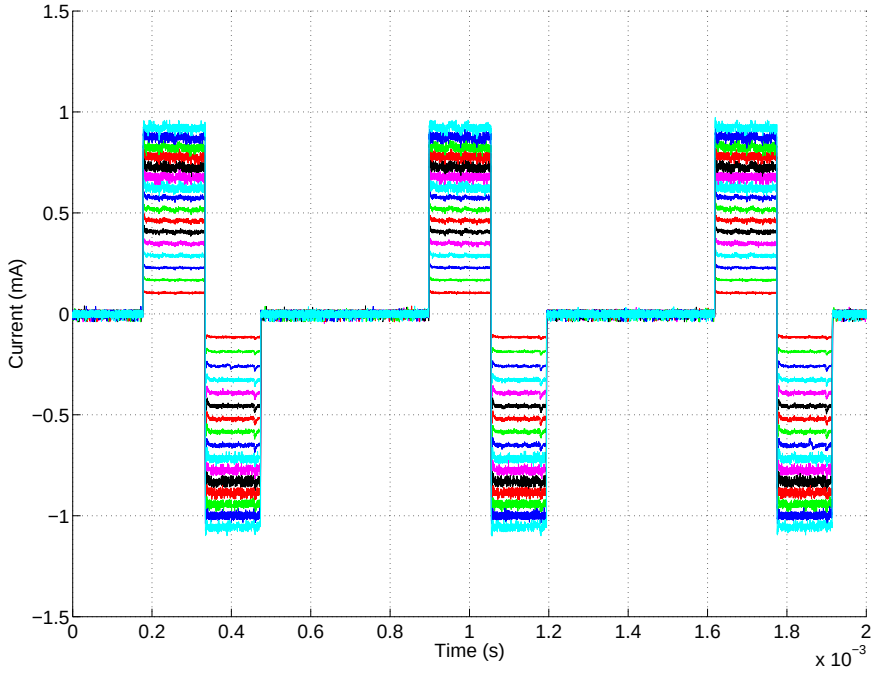


Figure 5-16: Measured biphasic current pulses demonstrating all 16 levels from 100μA to 1mA.

5.5.2 Output Driver Timer

Figure 5-17 shows a simplified schematic of the output driver timer. The timer circuit operates in a very similar manner to the pulse width demodulation circuit of Figure 5-8. The basic principle is that currents charging capacitors to a threshold can be used to produce pulses of a fixed width, T, set by:

$$T = \frac{C \cdot V_{ref}}{I} \tag{5.5}$$

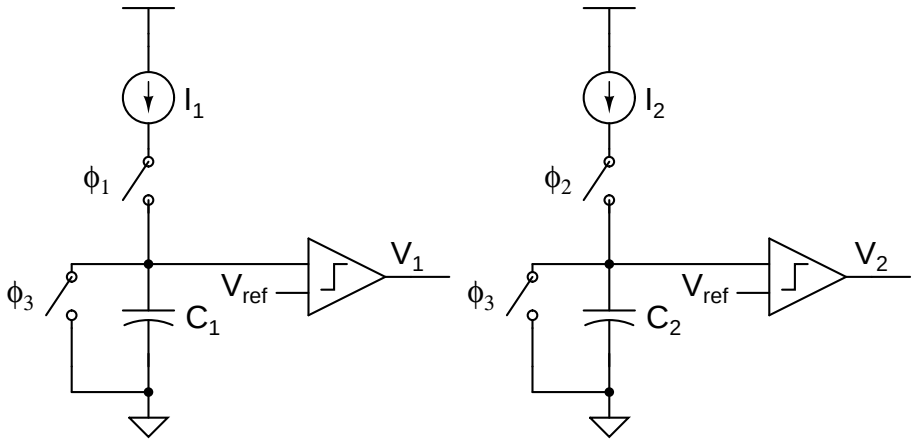


Figure 5-17: Output driver timer circuitry.

Figure 5-18 shows the logic used to control the charging and reset switches in Figure 5-17. The output driver timer may be described as a type of hybrid state machine which stores the stimulation state – either idle or stimulating – in FF1, and stimulation sub-states with the capacitors C_1 and C_2 which time the up and down pulses. The operation of the circuit may be described as follows:

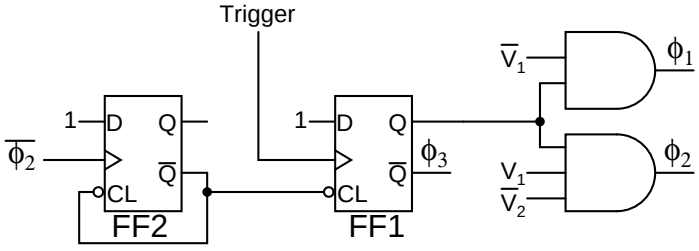


Figure 5-18: Output driver control logic.

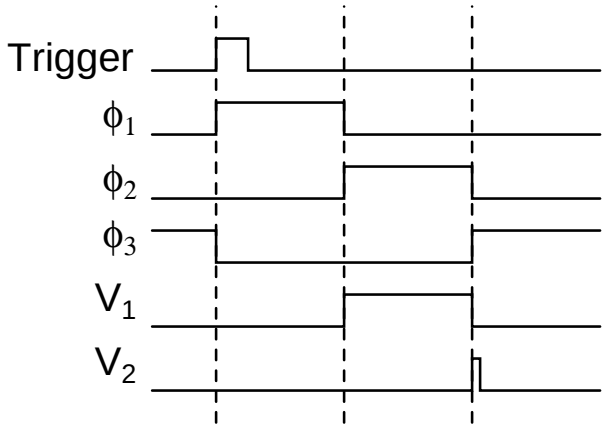


Figure 5-19: Timing diagram for output driver.

Initially, assume that capacitors C_1 and C_2 are both uncharged. Therefore the comparator outputs, V_1 and V_2, are both low. Additionally, assume that FF1 is in the reset state. Therefore, the switches ϕ_1 and ϕ_2 are also both open. When the recognition sequence is present, the clock input to FF1, marked $Trigger$ in Figure 5-18, is activated. This sets the output of FF1 high, and causes switch ϕ_1 to close. Simultaneously, switch ϕ_1 in Figure 5-15 also closes.[2] This connects the selected electrode to a NMOS current source. The switch remains closed while C_1 charges with a constant current, I_1. Once the capacitor voltage reaches the comparator threshold voltage set by V_{ref}, V_1 transitions from low to high. This causes ϕ_1 to open and discontinue charging C_1 while ϕ_2 closes and begins charging C_2 with the constant current, I_2. Like with the first phase, the control switches in the output driver cell each reverse phase and the current in the electrode reverses direction. Once C_2 charges to V_{ref}, V_2 changes state from low to high and opens ϕ_2. When ϕ_2 opens, this triggers FF2 to set itself. The $\overline{Q}$ output moves low, which clears both FF1 and FF2. When FF2 clears, ϕ_3 closes the two switches and discharges C_1 and C_2. The switches will remain closed until the next time FF1 is triggered by the recognition sequence. FF2 uses a feedback connection between its $\overline{Q}$ output and the $\overline{CL}$ input to generate a finite duration reset pulse. The duration of the pulse is set by the setup time of the

[2]Note that the phase ordering of the output current pulse is selectable. If the user selects the "up" phase to occur first instead of the "down" phase, switch ϕ_2 in the output driver cell would close first.

flip-flop. As soon as $\overline{Q}$ moves low, the $\overline{CL}$ input is activated, which attempts to move $\overline{Q}$ high, thus disabling $\overline{CL}$. However, in this time, Q of FF1 is reset low, C_1 and C_2 are discharged. The circuit is in the idle state, awaiting the next $Trigger$ pulse. The operation of the timing circuit may be summarized by the timing diagram shown in Figure 5-19.

5.6 Die Microphotograph

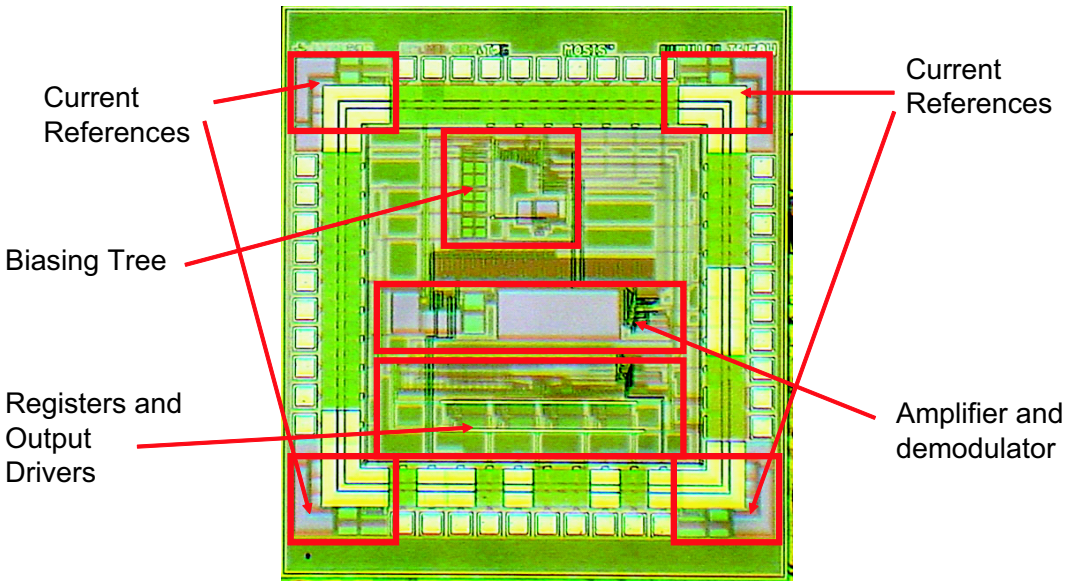


Figure 5-20: Die Microphotograph.

Chapter 6

On-Chip Supply Independent Biasing

This section contains excerpts from the paper "Fast Start-Up CMOS Current References" by S. Mandal, S. Arfin, and R. Sarpeshkar, which will appear in the 2006 IEEE International Symposium on Circuits and Systems [11].

6.1 Introduction

Analog circuits almost always require a well defined DC bias current in order to function. An on-chip current reference for setting DC biases throughout the chip is an important part of this wireless neural stimulator because the implantable device must be small and very light with a minimum number of off-chip components. In laboratory testing, DC supplies and potentiometers can be used to set biases, but during implanted operation, all biases must be generated by the chip itself. The chip's power supply comes from a battery whose voltage tends to droop with time. Supply independent biasing is therefore essential for battery powered operation without a voltage regulator.

In large systems, an effective biasing strategy is to implement a single current reference and scale it using current mirrors to generate all the analog bias currents on the chip [16]. These currents should be generated close to the current reference

and distributed across the chip. The alternative, distributing gate voltages which are then locally exponentiated (or squared) into bias currents, performs poorly because the high-impedance gate lines easily pick up noise.

Current references where the output is proportional to absolute temperature (PTAT) are useful in bipolar and subthreshold MOS design because they keep the transconductance $g_m \propto qI/(kT)$ constant with temperature. As a result, they are often known as constant-g_m references. Since the temperature coefficients of passive elements are small, circuits using constant-g_m references for biasing have temperature-independent time constants. For example, g_m-C filters with constant cut-off frequencies can be built.

The layout of this circuit is also rather interesting. Instead of laying out the circuit in the core, it was designed to fit in a corner of the pad ring. This layout technique has three advantages. The circuit can be instanced as a standard cell; it occupies no area of the core; it can be instanced in all 4 corners of the pad ring. The latter trick proved especially useful when earlier fabrication runs found that cell to cell variations in the reference current of 25% or more. The reference is nominally between 20 and 25nA, but can stray well outside those limits. Using 4 of them simultaneously and averaging all the reference currents together produces an average reference current reliably between 20 and 25nA. This has been verified over multiple fabrication runs.

6.2 Standard Current Reference

We first describe the basic self-biased supply-independent PTAT current reference. CMOS versions of this well-known bipolar circuit have been periodically reported in the literature [3, 13, 17]. A basic implementation is shown is Figure 6-1 (for now, ignore the startup circuits inside the box). The circuit works because the two NMOS transistors forming the current mirror have different W/L ratios. The ratio of their W/L's is designed to be M:1, with $M > 1$. Because the devices carry approximately the same current (this condition is imposed by the PMOS current mirror), the larger device has a smaller value of V_{GS} than the smaller device. The difference in V_{GS}

values is dropped across the resistor R. The value of R thus sets the value of the current flowing through the mirrors; this is mirrored out as the reference current I_{ref}. Assuming the devices operate in subthreshold, I_{ref} is given by

$$I_{ref} = \frac{\phi_T}{\eta R} \ln M \tag{6.1}$$

where $\phi_T = kT/q$ is the thermal voltage. If the bulk and source terminals of the NMOS transistors are tied together, $\eta = \kappa$, the subthreshold exponential constant. If the NMOS bulks are grounded, $\eta = 1$[1]. As expected, Equation (6.1) predicts that $I_{ref} \propto T$, the absolute temperature (i.e., is PTAT). The circuit shown in Figure 6-1 is nominally supply independent, since as long as the transistors are saturated, the value of I_{ref} does not depend on V_{DD}. In reality, the output impedance of the reference is set by the Early voltages of the transistors and is finite even if V_{DD} is high enough to guarantee saturation. Long devices are used in the circuit to improve the output impedance as much as possible. It can also be cascoded to further increase its output impedance if necessary. Unfortunately, cascoding increases the minimum value of V_{DD} required to guarantee that all transistors are saturated.

Capacitors C_P and C_N (normally implemented using MOS transistors to save layout area) are used to bypass the bias voltages V_P and V_N to V_{DD} and ground, respectively. As a result, high frequency noise on either V_{DD} or ground is shorted out and has no effect on the output current I_{ref}.

Finally, this current reference has two possible operating points: The desired state with the designed value of I_{ref} flowing, and an undesirable condition where there is no current flowing ($I_{ref} = 0$). It can be proved that the second state is metastable; the circuit eventually always starts up and goes into the normal operating state because of leakage currents and supply voltage transients. However, this process can take a long time. We have found that the startup time depends strongly on V_{DD}, the operating current I_{ref} and the sizes of C_P and C_N. If V_{DD} and I_{ref} are low and C_P and C_N are large, the circuit can take several seconds or minutes to start up. Since

[1]The body effect causes I_{ref} to decrease by a factor of $1/\kappa$, a significant amount since κ typically has a value of about 0.7.

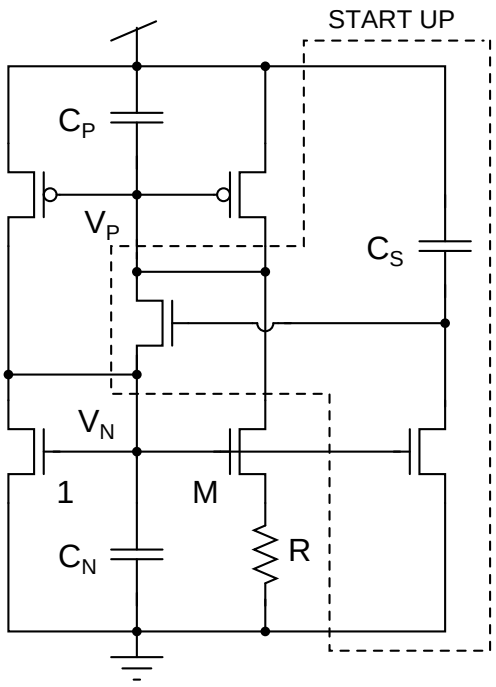


Figure 6-1: CMOS current reference with capacitively-coupled startup circuit.

this is rarely acceptable, we have thus designed an innovative startup circuit for the reference (shown inside the dashed lines in Figure 6-1). Most startup circuits for constant-g_m references use a string of diode-connected devices to generate a leakage current. This is not an elegant approach – the leakage current is V_{DD}-dependent and consumes extra static power. In contrast, our startup circuit consumes no static power and works over a wide range of power supply voltages.

6.3 Capacitive Startup Circuit

Consider the initial state of the system, with the supply voltage at zero and all capacitors uncharged. When V_{DD} is switched on, V_P goes to V_{DD}, and V_N stays at ground. Thus no current flows. However, C_S is also uncharged, and thus the gate of the NMOS switch connected between V_P and V_N is at V_{DD}. The switch thus turns on, shunting charge from V_P to V_N. This shunting lowers V_P and raises V_N, so current starts flowing and the circuit starts up. In addition, a copy of the output current placed in series with C_S discharges the bottom plate of C_S to ground. Thus the gate

voltage of the startup switch eventually goes to zero; the startup circuit, its job done, shuts off.[2] Startup time can be reduced by increasing the sizes of the startup switches. We have experimentally verified the operation of this startup circuit in several CMOS process technologies.

6.4 Improved Current Reference

The current reference used on the chip in this thesis, shown in Figure 6-2, is derived from the topology in Figure 6-1. The mirrors were replaced with cascoded mirrors. However, high transistor threshold voltages (particularly PMOS) in this process and the absence of wide-swing cascodes limit the minimum V_{DD} needed for saturation to 2.2V. The layout area of the reference was also reduced by replacing the large resistor, R, that sets the value of the current by a string of NMOS transistors operating in the linear (triode) region [3]. There are $N_1 + N_2$ identical devices connected in series as shown to form the current-defining resistor. The reference current is given by:

$$I_{ref} = K\mu_n(T)\phi_T^2 C_{ox} S \left(\ln M\right)^2 \tag{6.2}$$

where K is a technology-independent constant that depends on N_1, N_2 and the current-mirror ratio $R = I_{triode}/I_{ref}$ (where I_{triode} is the current through the triode transistors). Analytical formulas have been found for K but are too complicated to provide insight for actual design. Also, the triode devices have $W/L = S$, gate oxide capacitance density C_{ox} and electron mobility μ_n. In order to reduce the effect of threshold voltage variations while minimizing layout area, S was set to a value just low enough to place these transistors in moderate inversion at the operating current. Finally, I_{ref} varies with temperature as $\mu_n(T)T^2$, and so is not *guaranteed* to be PTAT. In our design, $R = 1$, $N_1 = 6$, $N_2 = 6$, and the devices were sized to yield a current of $I_{ref} = 25$nA with $M = 10$. All the important gate voltages in the circuit (V_P, V_{CP}, V_{CN} and V_N) were bypassed to V_{DD} or ground. The startup capacitor C_S

[2]Power-on reset circuits for digital logic tend to operate in a similar way. The current source is typically replaced by a resistor.

was set to 3pF. The total static current consumption of the reference was $3I_{ref}$.

This reference was designed to produce 50nA. However, measured values of I_{ref} were about 25nA. Inaccurate transistor models, particularly in the triode region, are suspected. The circuit was designed to be a standard analog cell; it was built into a corner pad in order to occupy no layout area in the core. Each chip contained four references (one in each corner). Deviations in output current across chips can be significantly reduced by adding together the currents produced by the four references on each chip. Chip-to-chip deviations were reduced to within ±10% of the mean value when this procedure was carried out.

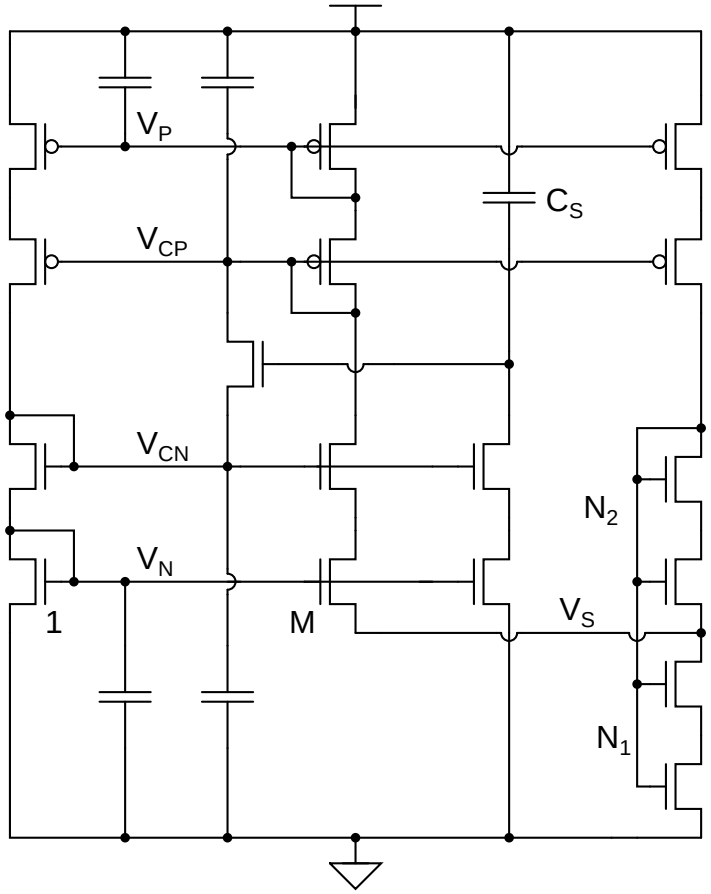


Figure 6-2: CMOS current reference used in this thesis.

6.5 Startup Time and Temperature Dependence

We define startup time as the time taken by the reference current to reach 90% of its steady state value. Startup times are strongly dependent on the supply voltage. Figure 6-3 shows startup time versus V_{DD}, measured by switching on V_{DD} while simultaneously measuring the current flowing through the circuit. The measured startup times spanned nearly four orders of magnitude as V_{DD} was varied from 1.8V to 5V.

We note three distinct startup regimes, visible in Figure 6-3 as different slopes in the startup time curve. For $V_{DD} < 3$V, a dead zone, when almost no current flows in the reference, appears as soon as the supply is switched on. Eventually, however, a rapid rise in the output current is observed, and the circuit is said to have started up. The shape of this rise is similar to that of a first-order RC circuit. The length of the dead zone decreases exponentially with the supply voltage until about 3V, when it becomes too small to be observable. In this second regime (3V$< V_{DD} <$4V) switching on the supply causes the current to rise immediately while following a first-order response. The rise time decreases weakly with V_{DD}. Once $V_{DD} > 4V$, the rise time begins to decrease rapidly. Overshoot and ringing in the step response can be observed when $V_{DD} > 5$V. In this regime startup is so fast that the dynamics of parasitic poles in the circuit start affecting it. Figure 6-4 shows measured startup waveforms at $V_{DD} = 2.5V$. A dead zone about 17ms long followed by a first-order rise to steady state is visible in I_{ref}.

The temperature dependence of this reference was measured from $-10°$C to 90°C. It was found to be very close to PTAT over the entire range, again indicating that $\mu_n \propto 1/T$ (see Figure 6-5). The behavior of hole mobility μ_p versus temperature can be studied by building a PMOS version of this circuit.

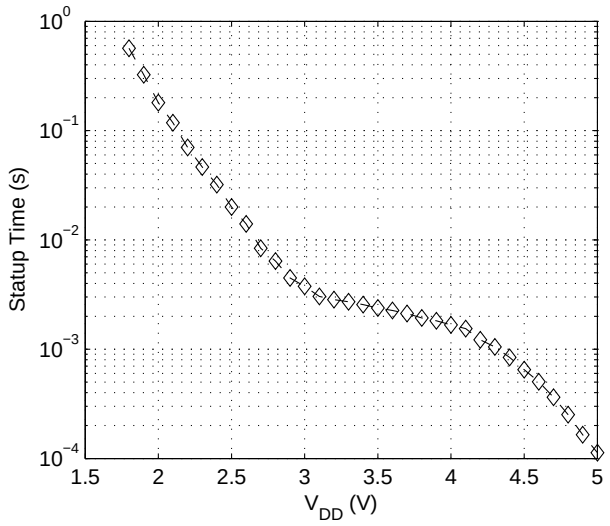


Figure 6-3: Experimentally observed 0.5μm current reference startup time versus supply voltage.

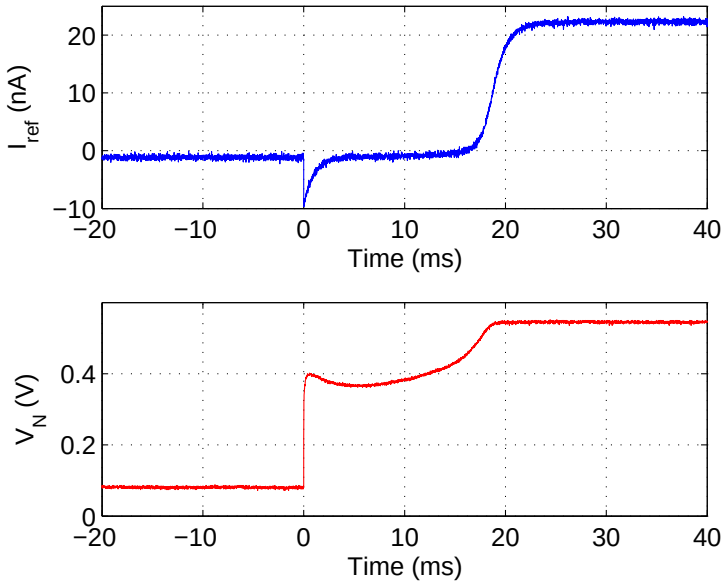


Figure 6-4: Measured output current (top) and NMOS bias voltage (bottom) of the 0.5μm current reference. V_{DD} was stepped from 0 to 2.5V at time=0.

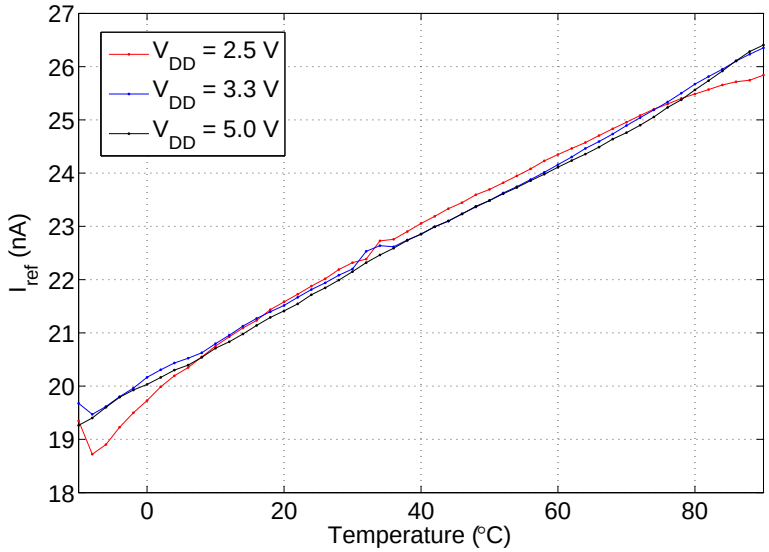


Figure 6-5: Current reference temperature data.

Chapter 7

The Implantable System

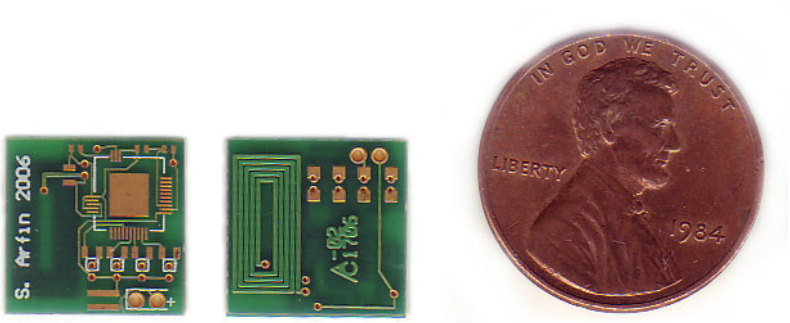
7.1 Printed Circuit Board Design

The implant is a printed circuit board designed to mount directly on the bird's cranium, with implanted wire electrodes extending from the underside of the printed circuit board. Figure 7-1 shows the top and bottom sides of the printed circuit, enlarged to show detail, next to a penny coin for scale. The actual dimensions of the circuit board are 1.2mm $\times$ 1.2mm $\times$ 31 mils. The board is finished with 30 micro-inches of soft, wire bondable gold.

The top side of the board contains a gold paddle for mounting the chip with epoxy to the board, 28 gold fingers for wire bonding the die to the board, and several pads for surface-mount capacitors and resistors. The lower-right corner of the board also contains two drill holes for inserting the PCB mount legs of the miniature battery. The chip will be encapsulated in an epoxy to protect it.

The bottom side of the board contains a printed coil of wire for receiving the RF signals. Also visible are 8 pads – 4 pairs of 2 – for soldering the electrodes which will be implanted into the bird. Each pair of electrodes contains a stimulating electrode and a ground electrode. The electrodes are also AC coupled to prevent any DC current from flowing into the bird.

Figure 7-2 shows the final assembled board with the bare die attached and wire-bonded to the board, and all components mounted.



(a) Top (b) Bottom (c) Coin for Scale

Figure 7-1: Printed circuit board design, shown next to a coin for scale.

Figure 7-2: Populated chip-on-board photograph.

7.2 Weight Estimates

The implant is very light. The battery weighs just 0.3g. The printed circuit board, made of FR4, is estimated to weigh 0.215g. The chip itself weighs approximately 0.065g. Data on the weights of the surface mount components were not available. However, they are not expected to contribute much to the overall weight of the implant. Other materials include copper+gold finish on the PCB, wirebonds, die epoxy, die encapsulation, and solder for mounting the surface-mount components. It is very reasonable to expect that the populated implant weighs under 1.5g.

7.3 Control System

The entire wireless system is currently controllable through a Matlab graphical user interface. Figure 7-3 shows a screenshot of the graphical user interface. The interface, designed primarily for testing the system, allows the user to select any of the 4 electrodes, the phase order of the pulse, and the current level from a 16-item drop down list. In an experimental setting with the zebra finch, pre-programmed sequences might replace the user interface. Instead, a microphone would listen for bird sounds and send out neural stimulation signals completely on its own. Figure 1-1, first shown in the Introduction at the beginning of this thesis, depicts how the system is likely to work with an actual subject.

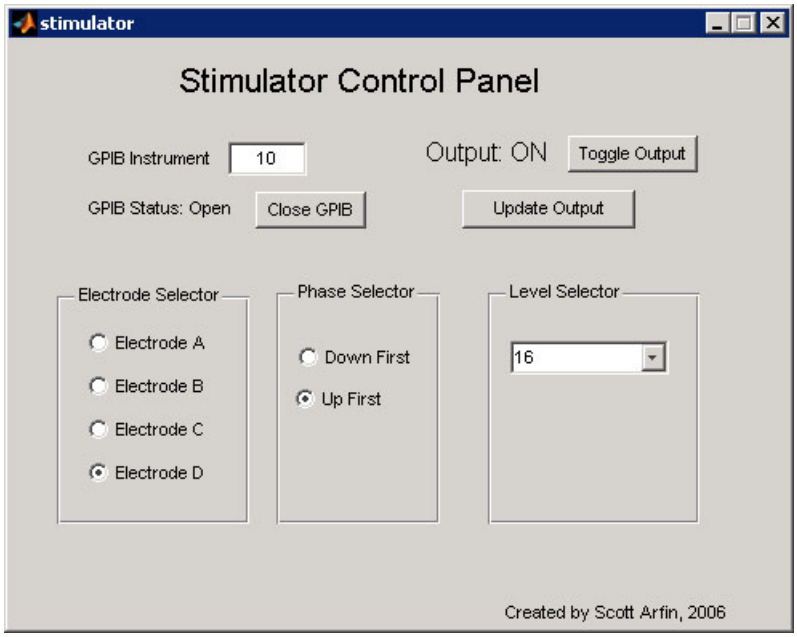


Figure 7-3: Screenshot of the graphical user interface.

Chapter 8

Future Work and Conclusions

Although the system has been tested extensively in lab with bench equipment and beakers full of saline to model avian tissue, it has not yet been tested inside a live animal. I would like to begin testing the device in birds as soon as possible. A longer term goal may be to adapt the design into a miniature, implantable cortical visual prosthesis for the blind. In the shorter term, some circuit improvements are always possible, including a power management system to extend the life of the implant.

8.1 Power Management System

One area where the implantable system can be improved is its power consumption. The chip currently draws a steady current of approximately 60μA at idle. This current is supplying the DC biasing to all circuits throughout the chip. Current consumption of course increases when data is transmitted and the output drivers are activated. However, due to the nature of the bird experiments, most of the time is spent idle with only occasional neural stimulations. On a 5mAh battery, this translates to approximately 3 days of continuous operation. To reduce standby power consumption, a power management system could power down all the circuits on the implant except RF amplifier and peak detector when not being used. This could increase the life of the implant to approximately 20 days.

Figure 8-1 shows a block diagram of the proposed standby system, to be imple-

mented on a future revision of the neural stimulator chip. The circuit uses a standby capacitor to store the state of the system. If no data has been transmitted to the receiver coil for a long time, the standby capacitor discharges to ground through a steady subthreshold leak current. The voltage on the capacitor opens a switch which removes all the biasing from the all the circuits on the chip other than the amplifier and peak detector. If the bird starts singing, the transmitter may begin sending data to the implant. The peak detector will pick up the data transitions, and for each transition will trigger a one-shot circuit that increases the voltage on the standby capacitor by a finite, moderate amount. If several transitions occur consecutively, the capacitor will eventually charge up to V_{DD}. That capacitor voltage will close the switch which will bias up the rest of the circuits on chip, including the PWM demodulator and the output drivers. As long as data continues to arrive, the standby capacitor will remain fully charged.

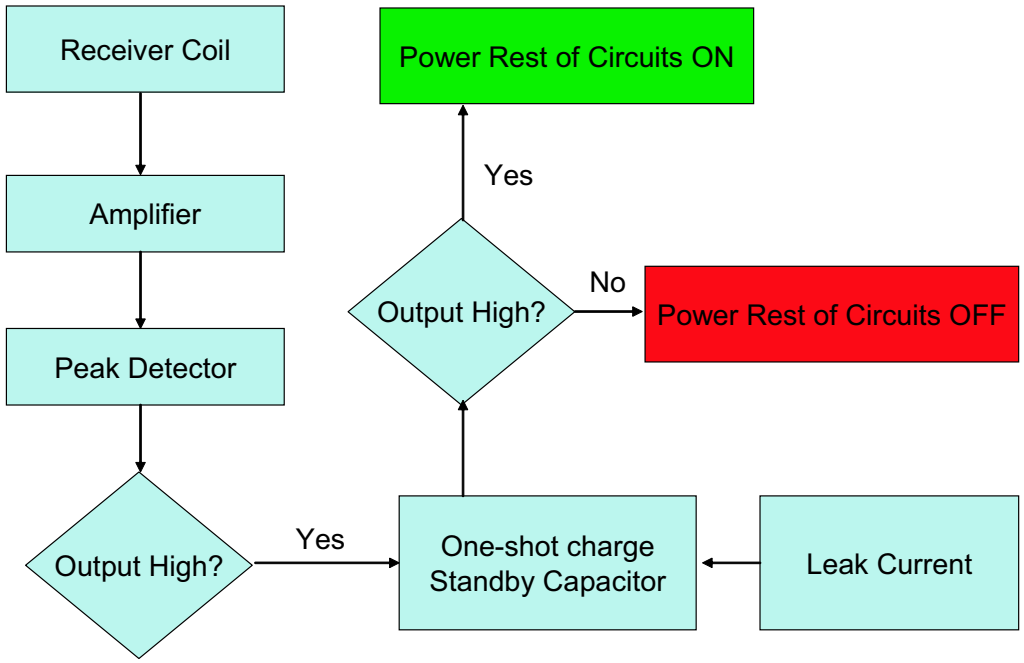


Figure 8-1: Proposed standby circuit block diagram.

8.2 Conclusions

In this thesis, I presented the design of a wireless neural stimulation system. The design of a discrete transmitter as well as an integrated circuit receiver were also presented. The implantable portion of the system was designed with attention to the size and weight, achieving a board area of 1.5cm^2 and a weight of less than 1.5g. The chip was fabricated in a standard 0.5μm CMOS process. It is capable of delivering biphasic current pulses to 4 addressable electrode sites at 16 selectable current levels ranging from 100μA to 1mA. The phase order of the pulses is also selectable. This thesis also briefly explored electrode models and presented some impedance measurements conducted in physiological saline solution.

Appendix A

Battery Properties

The battery used on the implant was a Panasonic ML621S. It is a Lithium-Manganese rechargeable battery. Although it is designed for ultra-low power applications, such as memory backup batteries for portable electronics, it is sufficient for the wireless implant. Its small size, light weight, high capacity, and high open circuit voltage make it an excellent choice for this application. The battery also comes with PCB mount legs, which make it even more desirable.

Table A.1 show some battery specifications as provided by the manufacturer. The battery is rated for a standard load current of 10μA, but we expect the quiescent load current to be higher than that – approximately 50μA. In order to verify that the battery would be effect at the higher load current, I laboratory tested the battery under simulated 50μA load conditions. The battery discharge curve is shown in Figure A-1. It is evident from the plot that this battery possesses a 5mAh capacity, as claimed by the manufacturer even when discharged at 5 times the standard load current.

Table A.1: ML621S Battery Details

Parameter	**Value**
Nominal Voltage	3V
Nominal Capacity	5mAh
Continuous Standard Load	10μA
Weight	0.30g
Diameter	6.8mm
Height	2.15mm

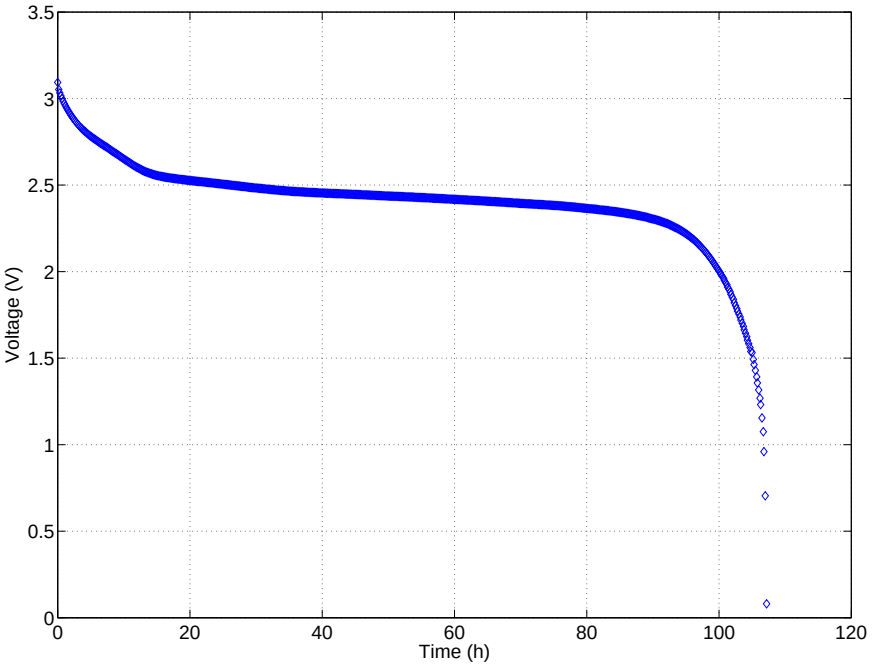


Figure A-1: Discharge curve for battery with 50μA load current.

Bibliography

[1] A. Bard and L. Faulkner. *Electrochemical Methods: Fundamentals and Applications*. John Wiley and Sons, Hoboken, NJ, Second edition, 2001.

[2] G. Brindley and W. Lewin. The Sensations Produced by Electrical Stimulation of the Visual Cortex. *J. Physiol.*, 196(2):479–493, May 1968.

[3] E. M. Camacho-Galeano, C. Galup-Montoro, and M. C. Schneider. A 2-nW 1.1-V Self-Biased Current Reference in CMOS Technology. *IEEE Transactions on Ciruits and Systems-II*, 52(2):61–65, February 2005.

[4] A. B. Carlson. *Circuits*. Brooks/Cole, Pacific Grove, CA, 2000.

[5] M. S. Fee. Neural Mechanisms of Sequence Generation and Learning. website.

[6] M. S. Fee and A. Leonardo. Miniature Motorized Microdrive and Commutator System for Chronic Nerual Recodring in Small Animals. *Journal of Neuroscience Methods*, 112(2):83–94, December 2001.

[7] K. Finkenzeller. *RFID Handbook*. John Wiley & Sons, West Sussex, England, Second edition, 2003.

[8] O. Foerster. Beitrage zur Pathophysiologie der Sehbahn und der Sehsphare. *J. Psychol. Neurol., Lpz*, 39:463–485, 1929.

[9] S. K. Kelly. *A System for Efficient Neural Stimulation with Energy Recovery*. Ph.D. Thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, October 2003.

[10] T. H. Lee. *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge Uinversity Press, Cambridge, UK, Second edition, 2004.

[11] S. Mandal, S. Arfin, and R. Sarpeshkar. Fast Startup CMOS Current References. *IEEE International Symposium on Circuits and Systems*, May 2006.

[12] S. S. Mohan, M. del Mar Hershenson, S. P. Boyd, and T. H. Lee. Simple Accurate Expressions for Planar Spiral Inductances. *IEEE Journal of Solid-State Circuits*, 34(10):1419–1424, October 1999.

[13] H. J. Oguey and D. Aebischer. CMOS Current Reference Without Resistance. *IEEE Journal of Solid-State Circuits*, 32(7):1132–1135, July 1997.

[14] K. L. Roach. Electrochemical Models for Electrode Behavior in Retinal Prostheses. M. Eng. Thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, May 2003.

[15] K. A. Sandquist. Design of a Wireless Power and Data Link for a Cranially-Implanted Neural Prosthesis. M. S. Thesis, Kansas State University, Department of Electrical and Computer Engineering, 2004.

[16] R. Sarpeshkar, C. Salthouse, J.-J. Sit, M. W. Baker, S. M. Zhak, T. K.-T. Lu, L. Turicchia, and S. Balster. An Ultra-Low-Power Programmable Analog Bionic Ear Processor. *IEEE Transactions on Biomedical Engineering*, 52(4):711–727, April 2005.

[17] F. Serra-Graells and J. L. Huertas. Sub-1-V CMOS Proportional-to-Absolute-Temperature References. *IEEE Journal of Solid-State Circuits*, 38(1):84–88, January 2003.

[18] G. Wang, W. Liu, et al. Design and Analysis of an Adaptive Transcutaneous Power Telemetry System for Biomedical Implants. *IEEE Transactions on Circuts and Systems—I*, 52(10):2109–2117, October 2005.

[19] G. Xu, Y. Chen, S. Yang, M. Wang, and W. Yan. The Optimal Design of Magnetic Coil in Transcranial Magnetic Stimulation. *Proceedings of the 2005 IEEE Engineering and Medicine and Biology 27th Annual Conference*, pages 6221–6224, September 2005.

[20] D. C. Yates, A. S. Holmes, and A. J. Burdett. Optimal Transmission Frequency for Ultralow-Power Short-Range Radio Links. *IEEE Transactions on Circuits and Systems—I*, 51(7):1405–1413, July 2004.

[21] B. Ziaie, K. Najafi, and D. J. Anderson. A Low-Power Miniature Transmitter Using a Low-Loss Silicon Platform for Biotelemetry. *Proc. IEEE 19th Intl. Conf. Engineering in Medicine and Biology Society*, 5:2221–2224, 1997.