WIRELESS BI-DIRECTIONAL BIO-POTENTIAL
RECORDING AND STIMULATION SYSTEM

by

CHENG-KUAN LU

Submitted in partial fulfillment of the requirements
For the degree of Doctor of Philosophy

Dissertation Advisor: Dr. Steven L. Garverick

Department of Electrical Engineering and Computer Science
CASE WESTERN RESERVE UNIVERSITY
Cleveland, Ohio

August, 2014
We hereby approve the dissertation of

CHENG-KUAN LU

candidate for the degree of DOCTOR OF PHILOSOPHY*

Committee Chair
Kenneth A. Loparo

Committee Member
Hillet J. Chiel

Committee Member
Swarup K. Bhunia

Committee Member
Francis L. Merat

Date of Defense
April 4th, 2014

*We also certify that written approval has been obtained for any proprietary material contained therein.
I would like to thank Dr. Hillet Chiel, Dr. swarup Bhunia, and Dr. Francis Merat for being part of my thesis committee. Their valuable help and suggestion on my thesis led to this final text. I am specially thankful to Dr. Kenneth Loparo for being head of committee.

I am highly grateful to my advisor Dr. Steven Garverick for his guidance and patience throughout the duration of this work. Without his help and advice in every point, this work would not go through. I will miss you, Dr. Garverick.

Many thanks are due to all members within my research group, MISC lab. I am specially thankful to Steve Majerus for lots of help with this project. I am also thankful to Paras Samsukha, Ming-Hsuan Hsu, Feng-Hsu Lin, Christopher Dorr, and Yingying Wang who participated and contributed to this project before. Thanks are also due to past member of lab, Dan Howe, Ting He, David Tian, and Anand Narayanaswamy for providing a wonderful atmosphere.

I would also like to thank Dr. Mohseni and the members of his group for constant support and warm friendship. Also many thanks to Dr. Chiel and his group members for their invaluable suggestions and help.

I am thankful to all my friends in Cleveland during the past ten years. Specially friends in EECS, Chia-Hung Kuo, Chia-Wei Soong, Chia-Hua Lin, Ting-Jung Chen, Kanokwan Limnuson, and many others, the time with you would be part of the best memory in my life.

Lastly, I owe too many thanks to my family. My parents, Kun-Lu Lu and Chia-Hsin Huang, my brother, Cheng-Hsin Lu, two kids, Emma and Brian, and my lovely wife, Sing Chau Ng, without your support non of this would have happened. Now that this is complete, we can finally move on to next journey together.
Contents

List of Tables iii

List of Figures v

Abstract vi

1 Introduction 1
  1.1 Motivation ................................ 1
  1.2 Challenges and Constraints ...................... 2
  1.3 Thesis Outline ................................ 4

2 Compression Algorithm Verification 5
  2.1 Introduction .................................... 5
  2.2 Spike Detection Algorithm .......................... 9
  2.3 Verification ........................................ 11
  2.4 Feature Preserving Spike Compression Algorithm .......... 18
  2.5 Reconstruction ..................................... 19
  2.6 Conclusion ........................................ 21

3 Four Channel Recorder/Stimulator Integrated Circuit 23
  3.1 Introduction ......................................... 23
  3.2 Overview .......................................... 27
  3.3 ASIC Design ........................................ 28
    3.3.1 Pre-Amplifier ................................... 28
3.3.2 Row Amplifier .............................................. 33
3.3.3 ADC Pre-Amplifier ........................................... 37
3.3.4 Timing ......................................................... 40
3.3.5 SAR ADC ..................................................... 41
3.3.6 Stimulator ..................................................... 42
3.3.7 SPI Interface .................................................. 44

3.4 Testing Result .................................................. 44
3.4.1 Reference Voltage ........................................... 45
3.4.2 Amplifier ....................................................... 45
3.4.3 Stimulator ....................................................... 49

3.5 Summary ......................................................... 49

4 Conclusion and Future Work 50
4.1 Conclusion ....................................................... 50
4.2 Future Work ...................................................... 51

Bibliography 52
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Size of main amplifier transistors</td>
<td>31</td>
</tr>
<tr>
<td>3.2</td>
<td>Size of feedback amplifier transistors</td>
<td>33</td>
</tr>
<tr>
<td>3.3</td>
<td>Size of the ADC pre-amplifier transistors</td>
<td>40</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Experimental setup for wireless neural recording with Aplysia . . . 2

2.1 The NEO and thresholding circuit diagram . . . . . . . . . . . . . . 6
2.2 Block diagram of Peng’s adaptive spike detector . . . . . . . . . . . 7
2.3 Block diagram of Rodríguez-Pérez’s neural recording ASIC . . . . . 8
2.4 Block diagram of the adaptive threshold spike detection algorithm . 10
2.5 Simulation of spike detection with different values of Q . . . . . . . 11
2.6 Illustration of B4/B5 and BN3 signal recording setup . . . . . . . . 12
2.7 Simulation of cross correlation with different window sizes . . . . . 13
2.8 Simulation of cross correlation with different time delay . . . . . . . 14
2.9 Simulation of normalized cross correlation coefficient with varying Q 15
2.10 Simulation of cross correlation with different start points . . . . . . 16
2.11 Simulation result in different correlation regions . . . . . . . . . . . 17
2.12 Illustration of the data compression algorithm . . . . . . . . . . . 18
2.13 Illustration of compressed neural data package . . . . . . . . . . . 19
2.14 Measured and average shapes of normalized waveform . . . . . . . 20
2.15 A comparison between reconstruction and original waveform . . . . 20
2.16 original and reconstructed waveforms comparison . . . . . . . . . 21
2.17 The correlation between original and reconstructed waveform . . . 22

3.1 Block diagram of the neural interface system by Harrison et. al . . . 24
3.2 Block diagram of the neural stimulation system by Brandonn et. al. 25
3.3 Block diagram of ASIC front end by Borna and Najafi . . . . . . . 26
3.4 Block diagram of the 4-channel ASIC . . . . . . . . . . . . . . . 27
Wireless Bi-Directional Bio-Potential Recording and Stimulation System

Abstract

by

Cheng-Kuan Lu

In this research, a system is design to achieve the bi-directional neural activity recording and stimulation system for a sea slug. The system includes 3 different modules, a data compression algorithm, and a signal recording/stimulation ASIC.

The neural spike detection compression algorithm developed by Ming-Husan Hsu is able to run in real time on a simple μ-controller and achieves an average compression ration of 20:1. The spike detection algorithm is verified by the correlation between pre-recorded signals from two buccal nerves (BN3 and B4/B5). In active period, the data showed high correlation value. The compression and re-construction algorithm can also achieve correlation value as high as 0.9 between original waveform and re-construction waveform.

The 4 channel neural recording and stimulation ASIC is designed and fabricated in a CMOS 0.5 μm process. The 3 serial amplifiers provide signal gain more than 10000 with an input refer noise of 2 μV_{rms}. After the signal digitized by a 8-bit SAR ADC, it can be further processed by on board μ-controller or sent directly by on chip FSK transmitter. The stimulator can provide a variety of strength from 50 μA to 300 μA. The command feeds through the SPI port to control the ASIC function.

The ASIC is designed capable of expanding input array size for large channel number. A implant unit integrated with ASIC and a compression algorithm loaded micro-controller, and a external tranceiver unit would complete the bi-directional system.
Chapter 1

Introduction

1.1 Motivation

Connections between the brain, neural activity and body motion is an interesting topic and with a deeper understanding of the mechanisms relating these, researchers are able to create new applications or devices to improve human life, for example: improved prosthesis. The in-vitro investigation of neural activity can provide a more detailed view of neural functions that enables neural-scientists to correlate behaviors with recorded neural activities. The development of devices that record and stimulate multiple channels of neurons simultaneously is greatly improved by the availability of advanced semiconductor processes.

Although animals have less complex neural structure, they neural systems are sufficiently similar to that of humans to be useful in research. Seashulls, *Aplysia Californica*, have a nervous system with ~20,000 neurons making them suitable for this kind of research. The large identifiable neurons enable the attachment of recording/stimulation probes and allow for high specificity recording/stimulation. Morton and Chiel have successfully identified the patterns of neural activity asso-
ciated with ingestion or egestion behavior \[7\][8].

A multi-channel neural recording and stimulation device that is small enough to be implanted into a seaslug and with wireless capability that allows the animal to move freely is required to advance research with more reliable data. Moreover, such a device can further be used in other studies that involve the study of neural activities, leading to improvements in human health.

1.2 Challenges and Constraints

Figure 1.1 illustrates the system setup that includes two main components: an implant unit and an external unit. The implanted device is immersed inside *Aplysia* in a 6 inch diameter saltwater container. Electrodes from the implanted device are attached to the ganglion for recording and stimulating purposes. The antennas transmit the neural activity data and receive setup commands from the external unit wirelessly through the saltwater. The implant unit is powered by a battery that is completely sealed inside the unit. The external unit is connected to a PC to log the data and send commands to the implant unit.

![Figure 1.1: Experimental setup for wireless neural recording with *Aplysia* \[9\].](image_url)
The battery powered system is expected to operate at least one hour a day and able to last for more than one week. The implant unit size limits the battery size and capacity. A compact size 80 mAh lithium ion battery is chosen to meet the size requirement, with the implant unit current consumption limited to be no more than 10 mA. The ASIC power consumption could be further reduced to insure battery performance during its lifetime or to permit using a smaller size battery.

The minimum detectable signal is related to the amplifier power consumption. Many present microsystems require signals with a minimum amplitude of 50 μV, with a noise floor of 10 μV. The expected minimum signal from intra cellular neural recording in Aplythia amplifier noise, but this would use more power. Reducing the number of amplifiers can reduce power consumption, and this can be accomplished by multiplexing the second and third stage amplifiers and integrating precise high cut-off frequency pre-amplifiers.

The recorded data is transmitted through the saltwater environment of the Aplysia. Chestek showed that signals are highly attenuated once the carrier frequency exceeds 50 MHz [10]. The highest Industrial, Scientific and Medical (ISM) band below 50 MHz is at 27.1 MHz. This frequency is chosen for this work to achieve a balance between limiting the attenuation of data transmission through saltwater and the bandwidth of the data bandwidth.

A 200 kHz bandwidth in the 27.1 MHz ISM band [11] is the bottleneck while the data rate increases with the number of channels or higher sampling rates. Traditional spike sorting only provides the temporal position of the neural spikes, but some researchers require more information about the neural signal waveform [9]. A data compression system specific for neural data is thus required for this application. The recorded neural data should be compressed before it is transmitted. The compression algorithm must preserve information about neural spike amplitude, duration and shape, and also has to be implemented in a PIC μ-controller.

The unit should also include the ability to stimulate neurons for neurodynamics
studies. To control the recording or stimulation function and channel switching, a in unit receiver is also required to close the loop. To avoid the interference between recording and data transmission and control command receiving, the carrier frequency for the command function should use a much lower ISM band.

1.3 Thesis Outline

This chapter introduces the motivation of building an implantable bidirectional neural recording and stimulation system. The system diagram is illustrated and the design challenges are addressed.

Chapter 2 details the compression algorithm. The literature review gives a broad view of related work done by other researchers. The background section provides the initial idea and design constraints. The full digital spike detection and waveform compression algorithm developed by Ming-Hsuan Hsu [12] is presented in detail and evaluated.

Chapter 3 focuses on the implantable ASIC design and testing. Four channels of a 3 stage neural amplifier are connected to an 8 bit SAR ADC. The system was initially developed by Paras Samsukha [13] and improved on in this work to increase the system performance and yield.

Finally, Chapter 4 provides the conclusion of this work and discusses the future complete system and ASIC design improvement.
Chapter 2

Compression Algorithm

Verification

2.1 Introduction

The overall data rate from an implanted neural sensor increases in proportion to the number of recording channels [14, 15]. To transmit all data, a wide bandwidth transmitter is required. Such a transmitter would consume considerable power and require a high carrier frequency. The transmitter for implantation and neural recording and stimulation in *Aplysia* should operate with low carrier frequency and bandwidth to reduce the absorption by the salt water. In this application the 27.1 MHz ISM band is chosen based on the previous constraints. To preserve enough waveform features while reducing data rate to fit in a limited transmission bandwidth can require very complex digital signal processing [16, 17].

The action potentials acquired from extracellular recordings are of interest to neural researchers[18]. By detecting and transmitting only the information conveyed by action potentials, the output data rate can be greatly reduced. Various
spike detection and compression algorithms have been developed and reported in the literature. Some detection algorithms were too complex to be implemented in an implantable device and require off-line computing \[19, 20\]. While other spike detection methods use an ASIC to minimize power consumption and achieve real-time compression, similar to our work.

For example, Holleman et. al. presented an analog spike detection ASIC \[5\] that includes two parts. Figure 2.1(a) shows the nonlinear energy operator (NEO) that provides a differential output current as the neural activity indicator. The other part is an adaptive thresholding circuit, shown in Figure 2.1(b) that uses the feedback loop formed by A1 and MN3 with the low-pass corner frequency set by MN3 and MN4 to filter out background noise and minimize false detections.

![Figure 2.1: The NEO output (a) connects to the thresholding circuit (b) \[5\]](image-url)
There are several other research works that also use NEO to sort spikes in real-time [5, 16, 21]. The NEO is integrated with the analog front end circuits and thus requires a strong input signal. When the signal-to-noise ratio is low, the NEO is confused by the high noise peaks, and the quality of spike detection is reduce. Other algorithms detect spikes after the neural signals are digitized [22, 23, 24, 16, 17], and can be used in real-time applications.

Adaptive spike detector increasing the functionality of device by sufficiently filter the noise to more accurately preserve spike information. Peng et. al. presented an adaptive spike detection algorithm shown in Figure 2.2. The first step is to average a certain period of the recorded signal to estimate the baseline level of the noise, and the a Finite State Machine (FSM) continuously adjusts the unit step size and window history size to adjust the noise level. A digital comparator is used to sort the spikes in the incoming neural signal.

An action potential also provides enough information about the timing of neural firing and in some applications researchers were able to connect body reaction with neural activity using both timing and the neural activity waveforms [9].

Rodríguez-Pérez et. al. used an ASIC to realize spike detection and compression in real-time, and the block diagram is shown in Figure 2.3(a). A neural signal is first digitized and processed to obtain a first order Piece-Wise-Linear (PWL)
A spike is represented as several positive or negative peaks, and each peak includes the peak position ($\Delta_1$), the overall peak duration ($\Delta_2$), and the peak amplitude ($V_p$). In between the positive and negative peaks, a separation duration ($\Delta_3$) is also measured. Figure 2.3(b) shows an example of the PWL spike approximation. All information is coded into 8-bit vectors, and the total data for each spike is 56 bits.

Figure 2.3: (a) Block diagram of the neural channel architecture. (b) Piece-Wise Linear spike approximation [17]
Digital signal processing approaches are very flexible, and the current algorithms reported in the literate are realized using either a high-speed FPGA or ASIC. Additional components are generally required to interface with front-end circuits and these will likely increase the size and weight of the implantable recording unit.

A compression algorithm requires spike detection and feature preservation, with small size compatible with a 8 bit microprocessor for future implant system integration. A recreated waveform from the reconstruction algorithm provides important features of the original neural signals that support the objectives of the research study.

The compression and reconstruction algorithm developed by Hsu is light suitable for a 8 bit microprocessor and preserves important features of the original neural signals [12]. The contributions of this work are the verification of the Hsu algorithm.

2.2 Spike Detection Algorithm

The spike detection algorithm uses an adaptive threshold and a non-linear filter to improve noise estimation [12]. As illustrated in Fig. 2.4, the noise power of the input signal is calculated using a low-pass filter (LPF) that includes “spike masking”, i.e. when the power of the input signal exceeds a defined threshold, the comparator output detects a spike and the multiplexer recycles the current output of the LPF to its input, ignoring the current input signal. Thus, the noise calculation is not affected by high-amplitude input signals, which enables the detection of a low-amplitude spike that follows a high-amplitude spike.

The threshold for spike detection is calculated by multiplying the noise power by an integer constant Q. The DC offset of the input signal is assumed to be negligible, but could be easily removed using a digital high-pass filter prior to the LPF, if necessary.
The low-pass filter is first-order and recursive (IIR), as given by Equation (2.1):

$$H_{LP}(Z) = \frac{(1 - \alpha)(1 + z^{-1})}{2 \times (1 - \alpha \times z^{-1})}$$  \hspace{1cm} (2.1)

The constant $\alpha$ is slightly less than 1, resulting in a very low cut-off frequency. The numerator includes a zero to reject high-frequency noise. The parameter $\alpha = 1 - 2^{-N}$ and $\frac{1 - \alpha}{2}$ is included to achieve a pass-band gain of 1.

The cut-off frequency of the low-pass filter is given in Equation (2.2). It is chosen to provide adequate attenuation of the noise while providing a reasonable response time to a change in noise power.

$$f_c = \frac{\omega_c f_s}{2\pi} \cos^{-1}(\frac{2\alpha}{1 + \alpha^2}) \times \frac{f_s}{(2\pi)} \approx 2^{-N} \frac{f_s}{(2\pi)}$$  \hspace{1cm} (2.2)

where $f_s = 3.3$ ksp s is the sampling frequency. Simulations have shown that $N = 8$, i.e., $f_c \approx 2$ Hz, provides a relatively stable noise level and the best spike detection results. A fixed-point format with rounding was used and found to have very similar performance to floating point simulation initially performed using MATLAB.

The threshold was computed by multiplying the LPF output by a constant $Q$, which can be changed to vary the sensitivity of spike detection. $Q$ is chosen to be an integer to minimize computation. Spikes are detected when the squared input signal exceeds the threshold. As $Q$ is reduced, high-amplitude noise might be
detected as spikes, and as Q is increased it is possible that low-amplitude spikes may be missed.

Fig. 2.5 shows an example of the performance of the spike detection algorithm when Q varies from 1 to 4. By visual inspection, Q = 2 or 3 appears to be good choices. A detailed numerical analysis presented in the next section confirms this observation.

Figure 2.5: Simulation of spike detection with different values of Q. The upper trace is the original data, and the following traces are detected spikes for Q = 1, 2, 3, and 4.

2.3 Verification

The performance of the algorithm was determined by comparing the results of spike detection from two different recordings that are expected to be highly correlated, with one having much better signal-to-noise-ratio (SNR) than the other.
The two channels of data were recorded from *Aplysia californica* as shown in Fig. 2.6. One signal was recorded using an extracellular needle electrode placed near the B4/B5 neuron, and the other was recorded using a suction electrode on the buccal nerve 3 (BN3). Signals generated at B4/B5 transmit through an axon to BN3 resulting in a time delay between the two recordings. The BN3 recording has high SNR and is considered to be a reference for the signal generated at B4/B5. The correlation between these two signals should be high, but not perfect since other neurons can contaminate the neural recordings from these two locations.

The normalized cross correlation is computed as given by Equation 2.3:

\[
\delta(d) = \frac{\sum_i [(x(i) - x_{avg}) \ast (y(i - d) - y_{avg})]}{\sqrt{\sum_i (x(i) - x_{avg})^2} \cdot \sqrt{\sum_i (y(i - d) - y_{avg})^2}}.
\]  

(2.3)

where \(x_{avg}\) and \(y_{avg}\) are the average values of the signals \(x\) and \(y\), respectively, and \(d\) is the delay between the signals being correlated. A window size of \(W\) samples is used for the computation. The correlation coefficient \(\delta(d)\) ranges from +1 (positive linear dependence) to 0 (uncorrelated) to −1 (negative linear dependence).

The size of the correlation window must be chosen properly to obtain mean-
meaningful results. MATLAB simulations were performed to determine the optimum window size using an active starting point in the data set and a fixed delay $d$, which had been estimated as 10 samples. As shown in Fig. 2.7, the best window size to test this pre-recorded data is about 1000 samples, i.e. about 0.3 sec at 3.3 kHA sampling frequency. A very small window size yielded a slightly stronger, but less meaningful, cross correlation.

Figure 2.7: Simulation of cross correlation with different window sizes.

The next simulation was used to quantify the optimum time delay $d$ between the two neural signals. The window size was fixed at 1000 samples and the starting sample was chosen as above. The delay time was varied, and the correlation coefficient should peak at the optimum delay, and decrease quickly when misaligned. The data of Fig. 2.8 reveals that the time delay between the two recordings is 11 samples, or approximately 3.3 msec.
Figure 2.8: Simulation of cross correlation with different time delays between B4/B5 and BN3 recordings.

Once the optimum window size and time delay are determined, the normalized cross correlation between B4/B5 and BN3 for varying Q was calculated using the pre-recorded data, as shown in Fig. 2.9. The peak occurs for Q = 2 or 3 for starting points in the pre-recorded data, in agreement with the visual inspection.
Figure 2.9: Simulation of normalized cross correlation coefficient with varying Q.

Fig. 2.10 shows the normalized cross correlation between B4/B5 and BN3 across the entire pre-recorded data set with the window size and delay chosen as above and Q = 3. In this case, the cross correlation of the original signals is compared to the cross correlation of the detected spikes. Most of the time, the detected spikes are more highly correlated than the original signals, suggesting that noise in the signals is filtered by the spike detection algorithm. There are periods of “0” correlation, that are represented “undefined”, because there was no activity in one or both recordings.
Figure 2.10: Simulation of cross correlation with varying starting points in the data set

Fig. 2.11 shows regions of high (0.75) and low (0.32) correlation. It is clear from visual inspection that the correlation is low when the B4/B5 recording has spikes that are only slightly above the background noise. Considering the relatively strong BN3 recording to be the reference, the false positive and false negative percentages are 3.1% and 4.5%, respectively in the region of high correlation.
Figure 2.11: Simulation result in high- and low- correlation regions. (a) high; (b) low.
2.4 Feature Preserving Spike Compression Algorithm

The proposed neural signal compression algorithm preserves only the amplitude and duration of the spikes, and the duration of an inactive period between spikes, or “run”. A spike begins when the input signal exceeds the threshold and ends when the signal drops below the threshold or changes polarity. When a spike is detected, a “run” is terminated and a run packet that includes the duration of the inactive period is transmitted. When a spike ends, the maximum amplitude and duration are encoded into a spike packet that is transmitted, and the next run begins. An example of this encoding scheme is illustrated in Fig. 2.12.

![Figure 2.12: Illustration of the data compression algorithm.](image)

<table>
<thead>
<tr>
<th>Run/Spike</th>
<th>Duration</th>
<th>Amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>34</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>-72</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>48</td>
</tr>
<tr>
<td>0</td>
<td>13</td>
<td></td>
</tr>
</tbody>
</table>

Each spike/run packet is allocated a 12-bit slot within a 61-bit frame, as illustrated in Fig. 2.13. The spike slot uses 1 bit to signify the spike, 3 bits to represent duration, and 8 bits for the two’s complement amplitude. The run slot
uses 1 bit to signify the run and 11 bits for the duration or amplitude. Exceptions are handled for spikes and runs whose duration exceed the maximum encoded values. For example, a run terminates when the duration reaches 2047, and a new run begins. Every 5 slots are combined with 1 synchronization bit to form the frame alignment word “100011011100”, which has been adopted from the T1 communication standard [25]. Synchronization overhead is therefore just $1/61$, less than 2%.

![Diagram of compressed neural data package](image.png)

**Figure 2.13**: Illustration of compressed neural data package.

### 2.5 Reconstruction

A reconstruction algorithm is used to create a noise-free visualization of the compressed neural data. Each run is reconstructed with zero amplitude for the given duration. Each spike is reconstructed using an average shape corresponding to the spike duration. The library of spike shapes was created by averaging the raw data for each detected spike in the pre-recorded data sets, after normalization to unity amplitude. A subset of this library is shown in Fig. 2.14. A comparison of one neural spike to its compressed reconstruction is shown in Fig. 2.15. This simple
algorithm does surprisingly well in creating a visual reconstruction of the neural signal.

Figure 2.14: Measured (faint) and average (dark) shapes of normalized pre-recorded waveform with different duration.

Figure 2.15: A comparison between reconstruction waveform (solid line) from the compressed data in Fig. 2.12 and the original waveform (dash line).

An example showing the original waveform, the detected spikes, and the visual
reconstruction from compressed data is shown in Fig. 2.16. The reconstructed waveform is qualitatively similar to the original waveform, except for the reduction in noise.

Figure 2.16: Original waveform (top), detected spikes (middle), and the reconstructed waveform (bottom).

Finally, Fig. 2.17 shows the normalized cross correlation between the original and reconstructed waveforms. Correlation is as high as 0.9 in regions of high neural activity. As before, correlation is “0” (undefined) when no spikes are detected within the 3.3 msec window.

2.6 Conclusion

A novel algorithm for neural spike detection, 20:1 data compression and visual reconstruction is presented and verified. High correlation between spike identification and noise filtering shows the algorithm has high accuracy. Using the reconstruction algorithm we obtain a waveform that preserves most of the visual features of the original signal.
Figure 2.17: The normalized cross correlation between the original and reconstructed waveform.
Chapter 3

Four Channel Recorder/Stimulator Integrated Circuit

This chapter presents detailed design of the Application Specific Integrated Circuit (ASIC) and its bench testing results. The emphasis is to meet the gain, noise and power requirements. The ASIC is fabricated in the 0.5 $\mu$m process from ON semiconductor.

3.1 Introduction

A notable multi-channel neural recording device was developed by researchers in the University of Utah. Harrison et. al [4] developed a 100-channel neural recording ASIC. Their system has a $10 \times 10$ amplifier array. A 10-bit ADC with 15kSamples/s is shared by all amplifiers. The data including spike detection, a total of 250 kb every second, is transmitted though a 433 MHz Frequency Shift Key (FSK) transmitter. Commands are Amplitude Modulated (AM) using a
2.64 MHz carrier. The cascade two-stage amplifiers boost the extracellular neural signal by 60 dB. Each channel has a comparator to detect the spike by comparing the neural signal and a reference voltage from a 7-bit DAC. The neural interface system consumes 13.5 mW of power and operates at minimum voltage of 3.55 V. Figure 3.1 shows the block diagram of the system.

![Block diagram of the neural interface system by Harrison et. al.](image)

The same team also presented another system with stimulation capability. Thurgood et. al. [14] designed an ASIC for 100 channel of charge balanced neural stimulation. Each stimulator in the $10 \times 10$ has a $R-2R$ DAC and an active charge recovery circuit to maintain the charge balance. The command to adjust stimulation current amplitude, duration, inter-phasic delay and repetition rate is delivered over a 2.765 MHz Amplitude Shift Key (ASK) inductive link. The current resolution is 8-bit, and maximum pulse width is $370 \mu s$. Figure 3.2 shows the block diagram of the stimulation system.

Borna and Najafi presented an ASIC with 8 channel recording chip integrated with wireless transmitter [26]. One of the front end channel can be assigned as input reference voltage. The active low frequency suppression core (ALFS) and the miller integrator sharpen the high pass filter and the cut-off frequency is at 300 Hz. This can eliminate the noise from 60 Hz and muscle movement artifacts.
The front end OTA gain is 40 dB. All bias current mirror transistors operate in strong inversion regime to minimize the effect of process variations. The input pair transistors operate in sub-threshold regime to increase the transconductance under fixed current. Sizing the transistors is done to gain the balance between low noise and low power. The front end block diagram is shown in Figure 3.1 and the circuit diagram of the front end amplifier is shown in Figure 3.1.
Aplysia neural signals are obtained using stainless steel hook electrodes [7]. The animal is tethered by electrode leads which would affect its behavior. The leads also become a significant source of noise burying an already weak neural signal. A implantable device combining low noise recording, stimulation, and wireless transmission through saltwater is required for more complicated Aplysia neural
patterns with specific behaviors experiments.

### 3.2 Overview

Figure 3.4 shows the architecture of the 4-channel ASIC. The ASIC includes 4 neural amplifier/stimulators, an analog-to-digital converter (ADC), an FSK data transmitter, and an OOK command receiver. This architecture is proposed by Paras Samsukha [13]

![Block diagram of the 4-channel ASIC](image)

The neural signal is amplified in three stages. The first stage pre-amplifier has a gain of 50, the second stage row-amplifier provides gain of 16, and the third stage ADC pre-amplifier has a variable gain from 2 to 16. The total gain is between 1600 to 12800. The shared ADC is an 8-bit Successive Approximation Register (SAR) type. The operating channel can be selected through row and column select controls. This approach is scalable to larger channel counts. The FSK transmitter sends recorded data to an external receiver at 27.12 MHz. The Manchester encoded command is received by the OOK command receiver at 125 kHz.
3.3 ASIC Design

3.3.1 Pre-Ampifier

The preamp architecture is shown in Figure 3.5. The ground symbol represents the reference voltage. The low-noise, low-power amplifier is ac-coupled to avoid offset propagation. The feedback amplifier and capacitors limit the bandwidth. The ratio of input capacitor and integrating capacitor, $-\frac{C_1}{C_2}$, sets the closed loop gain to 100. The feed-forward gain transfer function is:

$$FF(s) = -\frac{C_1}{C_2}(s + \frac{1}{RC_2})(s + \frac{1}{R_{out}C_3})$$  \hspace{1cm} (3.1)$$

The amplifier output resistance, $R_{out}$, together with $C_3$ sets the high cutoff frequency. In the absence of the feedback amplifier, the low cutoff frequency is set by $R$ and $C_2$, where $R$ is the pseudo resistor created by two sub-threshold MOS transistors on top of $C_2$. The value of this resistor can have large variation from chip to chip and is far too large to implement a desirable low cutoff frequency. Instead, the feedback amplifier is used to set the low cutoff frequency at a de-
sirable value between 60 Hz and 100 Hz. A current generated by the feedback amplifier cancels the signal current driven through $C_1$ at the frequency within the loop bandwidth of the overall pre-amplifier. The closed-loop bandwidth is set by the unity gain frequency, which is a $\frac{g_m}{C}$ ratio. The transfer function is:

$$LT(s) = -\frac{g_mA}{(C_1 + A'C_2)s + \frac{A'}{R}}$$  \hspace{1cm} (3.2)

$A$ is the open loop gain of the main amplifier and $A' = A + 1$. $g_m$ is the transconductance of the feedback amplifier. Assuming the $A$ and $A'$ are about the same while open loop gain is large, and $R$ is very high. If the closed loop gain is much smaller than the open loop gain, $A'C_2 > C_1$, the bandwidth is approximately $\frac{g_m}{C_2}$. The overall pre-amplifier response is:

$$H(s) = -\frac{C_1}{C_2}(\frac{s}{s + \frac{g_m}{C_2}})(s + R_{out}C_3)$$ \hspace{1cm} (3.3)

$\frac{g_m}{C_2}$ sets the low cutoff frequency. In this design $C_2$ is selected as 500 fF. $g_m$ would be a very low value in order to achieve a desirable low cutoff frequency range.

The detailed design for each circuit component is presented in the following subsections.

**Main Amplifier Design**

Figure 3.6 shows the pre-amplifier schematic. The design is based on [4]. The input-referred thermal noise power density is:

$$v^2_{\text{thermal}} = \frac{16kT}{3g_{m1}}[1 + 2\frac{g_{m3}}{g_{m1}} + \frac{g_{m7}}{g_{m1}}]$$ \hspace{1cm} (3.4)

To minimize the noise, a large value of $g_{m1}$, $g_{m1} \gg g_{m3}, g_{m7}$ is needed, which can reduce the contribution from transistors M₃ - M₇. This means we need to make $(W/L)_1 \gg (W/L)_3, (W/L)_7$. 

29
The requirement of input-referred noise is $2\mu V_{rms}$ in 4 kHz bandwidth or $31 nV/\sqrt{Hz}$. Assuming 75% of the total noise is given by $\frac{16kT}{3g_m}$ contributed from the input pair, the $g_m$ of the input pair has to be greater than 122.6 $\mu$-mho. Biasing the input pair transistors in weak inversion region can maximize the $g_m$ ratio, and minimize the current consumption. Weak inversion biased MOSFET transconductance is given by

$$g_m = \frac{I}{\eta V_{th}}$$

(3.5)

where $I$ is the bias current, $\eta$ is the sub threshold factor and $V_{th}$ is the thermal voltage. $\eta V_{th}$ for this process is around 40 mV estimated from previous run [13]. The current for the input pair is 4.9 $\mu$A under the $g_m$ requirement. An inversion coefficient [27], $\frac{L}{T_{sp}} = 0.25$, is used to design the input pair, where $I_{sp}$ is surface current and is given by

$$I_{sp} = 2\mu C_{ox}\eta V_{th}^2 \frac{W}{L}$$

(3.6)

With $\mu C_{ox} = 38\mu A/V^2$ in this process, the input pair shape factor is $\frac{W}{L} > 260$. In this design, $\frac{W}{L} = 300$ in order to ensure reaching the noise requirement. Large size of the PMOS input pair also reduces the flicker noise and helps pushing the noise corner frequency lower. The corner frequency is around 1 Hz in this design.
which is well below the desired low-cutoff frequency.

A high open-loop gain is required to minimize the voltage offset and reduce the bandwidth limited capacitor, $C_3$. Therefore, the cascode output stage is used. To reducing the noise contribution, the $\left(\frac{W}{L}\right)_3$ and $\left(\frac{W}{L}\right)_7$ are decreased which means higher $V_{Dsat}$ for $M_3$ and $M_7$ under the same amount of current. To maximize the output swing, the $V_{Dsat}$ of $M_{11}$, and $M_{12}$ need to minimize while the transistors still operating under saturation region. The $V_{Dsat}$ of $M_7$, $M_8$ and $M_{11}$, $M_{12}$ are 0.4 V and 0.2 V, respectively. $M_{11}$ and $M_{12}$ bias voltage are generated locally from a stable current source.

The sizes of each transistor in the main amplifier are listed in Table 3.1.

<table>
<thead>
<tr>
<th>Transistor</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$, $M_2$</td>
<td>$540 \mu m/1.8 \mu m$</td>
</tr>
<tr>
<td>$M_3$, $M_4$, $M_5$, $M_6$</td>
<td>$1.8 \mu m/5.4 \mu m$</td>
</tr>
<tr>
<td>$M_7$, $M_8$</td>
<td>$3 \mu m/1.8 \mu m$</td>
</tr>
<tr>
<td>$M_9$, $M_{14}$</td>
<td>$2.1 \mu m/1.8 \mu m$</td>
</tr>
<tr>
<td>$M_{10}$</td>
<td>$8.4 \mu m/1.8 \mu m$</td>
</tr>
<tr>
<td>$M_{11}$, $M_{12}$</td>
<td>$45 \mu m/1.8 \mu m$</td>
</tr>
<tr>
<td>$M_{13}$, $M_{15}$</td>
<td>$1.8 \mu m/4.5 \mu m$</td>
</tr>
<tr>
<td>$M_{16}$</td>
<td>$1.8 \mu m/1.8 \mu m$</td>
</tr>
</tbody>
</table>

Table 3.1: Size of main amplifier transistors
Figure 3.7: Schematic of the feedback amplifier

Figure 3.7 shows the schematic of the feedback amplifier, M₁ ∼ M₆ and the bias generating circuit, M₇ ∼ M₁₄. To set the low-cutoff frequency at 80 Hz, the required transconductance of the feedback amplifier is 250 p-mho with C₂ = 0.5 pF. The $g_m$ of sub-threshold biased transistor is based on equation 3.5. Such a very low value of $g_m$ requires only 10 pA current passing through the differential pair. The bias circuit generates an ultra-low current through M₇, and a current division technique [28] is used to split the current resulting in the required value.

The ultra-low bias current generator is a self-biased circuit [29]. M₁₁ M₁₄ are biased in the weak inversion region. The circuit generates a voltage of M₁₁ drain given by $V_{th}ln\left(\frac{S_{12}S_{11}}{S_{14}S_{13}}\right)$ like bandgap voltage reference, where $S_{11}$ $S_{14}$ are the $\frac{W}{L}$ ration of transistors M₁₁ M₁₄. The $\left(\frac{S_{12}S_{11}}{S_{14}S_{13}}\right)$ is set to 10 for stable voltage.

The current division technique uses M₃ and M₄, which are 9X bigger than M₁ and M₂, to reduce the effective tail current by 10X. Thus, the differential pair of feedback amplifier has accurate low bias current for the ultra low $g_m$ requirement. The ASIC has two different $\frac{W}{L}$ ratios of M₇ which can be switched by SPI command.
to adjust the low cutoff frequency closer to the desired corner.

The sizes of each transistor in the feedback amplifier are listed in Table 3.2.

<table>
<thead>
<tr>
<th>Transistor</th>
<th>Size</th>
<th>Multipiler</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1, M_2$</td>
<td>18(\mu m/9\mu m)</td>
<td></td>
</tr>
<tr>
<td>$M_3, M_4$</td>
<td>18(\mu m/9\mu m) m=19</td>
<td></td>
</tr>
<tr>
<td>$M_5, M_6$</td>
<td>3.6(\mu m/7.2\mu m)</td>
<td></td>
</tr>
<tr>
<td>$M_7$</td>
<td>18(\mu m/2000\mu m)</td>
<td></td>
</tr>
<tr>
<td>$M_8, M_{12}, M_{14}$</td>
<td>36(\mu m/600\mu m)</td>
<td></td>
</tr>
<tr>
<td>$M_9$</td>
<td>21(\mu m/810\mu m)</td>
<td></td>
</tr>
<tr>
<td>$M_{10}$</td>
<td>21(\mu m/1620\mu m)</td>
<td></td>
</tr>
<tr>
<td>$M_{11}$</td>
<td>18(\mu m/72\mu m) m=10</td>
<td></td>
</tr>
<tr>
<td>$M_{13}$</td>
<td>18(\mu m/72\mu m)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Size of feedback amplifier transistors

### 3.3.2 Row Amplifier

The pre-amplifier cell outputs from the same row are multiplexed to the row amplifier. The row amplifier is a differential input single-ended output with a gain of 16. The overall gain at this stage is 800. The output signal swing requirement is 100 mV which is under the design limit of the pre-amplifier and the bandwidth is still the same as the pre-amplifier. The row amplifier transistors are simply copy of the corresponding ones in the pre-amplifier.

**Switch Capacitor Circuitry**

The switch capacitor (SC) circuitry is used in both the row amplifier and the ADC pre-amplifier. A SC topology was chosen to minimize the crosstalk between each channel and is suitable for ADC sampling. The wide band topology \([30]\) with a offset storage capacitor is selected. This can reduce the power consumption by reducing the voltage change at capacitors between phases, and minimize the input offset between previous stage output and reference voltage.

This SC circuit has five capacitors as shown in Figure 3.8. The ground symbol represents the reference voltage. $C_1 = C_3$ and $C_2 = C_4$ form the gain stage. $C_I$ is the offset storage capacitor. The SC circuit has two phases, sample ($\phi_1$) and hold.
(ϕ₂). In the sample phase, the input signal goes through Capacitor $C_1$ and $C_2$ to the output. Both ends of $C_3$ and $C_4$ connect to $V_{ref}$ clearing the charge. In the hold phase, the input signal moves to $C_3$ and $C_4$ with output disconnecting. $C_1$ and $C_2$ reset and $C_I$ stores the offset between $V_{ref}$ and the gain signal path replica. The normal output gain is $V_{out} = -(C_1/C_2)V_{in}$, and the gain error is given by:

$$\hat{A}_v = \frac{1 + \frac{C_1}{C_2}}{A^2}$$  \hspace{1cm} (3.7)

where $A$ is the open loop gain of the amplifier. All switches are implemented using CMOS transmission gates with NMOS and PMOS ($W/L$) = $(5\mu m/0.6\mu m)$ and $(1.5\mu m/0.6\mu m)$, respectively. The ON resistance of the switches is 3.2 $k\Omega$s. For a 5 pF load, the settling bandwidth is much higher than the circuit requirement.
Figure 3.8: Schematic of row amplifier and pre-ADC amplifier switch capacitor circuit.

There are four non overlapping clock phases for controlling the switches in the SC amplifiers, as shown in Figure 3.9. φ₁ and φ₁d are high during the sample phase. In evaluation, φ₂ and φ₂d are active. φ₁d and φ₂d are the delayed phase of φ₁ and φ₂, respectively. φ₁ and φ₂ flip the switches, S₅ and S₆, before other part of circuit changing phases, and φ₁d, φ₂d control the rest of switches. This can ensure the input node of amplifier will not be disturbed between phase transitions.
Noise Analysis

There are two main noise sources for this amplifier: the switch capacitor thermal noise \( v_n^2 = \frac{kT}{C} \) and the amplifier thermal noise.

The input referred noise due to sampling charge is given by:

\[
\overline{v_{n,s}^2} = \frac{2}{C_1^2}(kT C_1 + kT C_2) = \frac{2kT}{C_1}(1 + \frac{C_2}{C_1}) \tag{3.8}
\]

Noise charge \( kT C_1 + kT C_2 \) is stored on the capacitor in the sampling phase and remains inside through the holding phase. Noise of \( C_3 \) and \( C_4 \) is not included because they are not present in the final signal path.

The amplifier output referred thermal noise is shaped by the voltage gain and bandwidth, and is given as:

\[
\overline{v_{n,\text{out}}^2} = \overline{v_{n,\text{open}}^2} (1 + \frac{C_1}{C_2})^2 \times f_B \times \frac{\pi}{2} \tag{3.9}
\]

where \( \overline{v_{n,\text{open}}^2} \) is the open loop input referred noise of amplifier, \( f_B \) is the bandwidth, and \( \frac{\pi}{2} \) is from the 1st order filter. The amplifier thermal noise as designed is dominated by the input pair, \( \overline{v_{n,\text{open}}^2} = \frac{16kT}{3g_m} \). The closed loop amplifier bandwidth
is:
\[ f_B = \frac{C_2}{C_1 + C_2} \times \frac{g_m}{2\pi(C_L + C_2)} \]  
(3.10)

By replacing \( v_{n,open}^2 \) and \( f_B \) in equation 3.9 and dividing the amplifier gain, \( \left( \frac{C_1}{C_2} \right)^2 \), the closed loop input referred thermal noise of the amplifier is:

\[ v_{n,input}^2 = \frac{4kT}{3C_1} \times \left( \frac{C_1 + C_2}{C_1} \times \frac{C_2}{C_L + C + 2} \right) \]  
(3.11)

Adding input referred noise in equation 3.8 to the referred thermal noise in equation 3.11 provides the total input referred noise of the row amplifier. There is no \( g_m \) dependent, only the capacitance ratio of \( C_1, C_2 \) and \( C_L \). Since the ADC pre-amplifier input capacitor is relatively higher than \( C_2 \), the noise is primarily capacitor sampling noise. In this design, \( C_1 \) is 2 pF and \( C_2 \) is 0.1 pF. The input referred noise is calculated as 66 \( \mu \)V, which is well below the desired noise level of 100 \( \mu \)V.

### 3.3.3 ADC Pre-Amplifier

There is another amplification stage to achieve the desired gain of 12800 before digitizing signals. The bandwidth linearly increases along with the channel number count. The output of this amplifier needs to swing between ADC reference high and low, 1.9 V and 0.5 V, respectively. ADC input capacitor array, 3 pF, and the integration capacitors in switch capacitor network, 1 pF total, are loaded at the output of ADC Pre-Amplifier. The ADC-preamplifier is shared by all 4 channels. To achieve sufficient amount of slew rate for this stage, class-AB amplification topology is chosen.

Figure 3.10 shows the class-AB amplifier topology, which is called super class-AB. This amplifier has a better efficiency due to adaptive input biasing and local common mode feedback (LCMFB).

The resistors \( R_1 \) and \( R_2 \) provide the local feedback loop of the differential, and a balance between high slew rate and gain bandwidth of the amplifier. The output
Figure 3.10: Schematic of the ADC pre-amplifier.

current is given by:

\[
I_{out} \approx \pm \frac{\beta_{11,13}}{2} \left( \sqrt{\frac{2I_{cm}}{\beta_{12,14}}} + \frac{R|I_d|}{2} \right)^2
\]

(3.12)

where \( R \) is the value of \( R_{1,2} \), and \( I_{cm} \) and \( I_d \) are the common mode current and differential currents, respectively [31]. For typical operation, \( I_d \) is proportional to \( V_{id}^2 \) instead of \( V_{id} \) in the traditional class AB for \( V_{id} \) larger than \( M_1 V_{dsat} \).
Figure 3.11: The input stage with adaptive bias circuit.

Figure 3.11 shows the adaptive bias circuit. The flipped voltage follower (FVF) output is a shifted input voltage, \( V_{in} - V_T - V_{dsat} \). The source of \( M_1 \) or \( M_2 \) is connected to the cross-coupled level shifted input voltage. Therefore, the transconductance of the input pair is twice that of a conventional differential pair. The ac small signal differential current of the input pair is:

\[
    i_d = i_1 - i_2 \approx 2g_{m1}v_{id}
\]  

(3.13)

The large signal current flowing through \( M_1 \) is:

\[
    I_1 = \frac{\beta}{2}(\sqrt{\frac{2I_B}{\beta} + V_{id}})^2
\]  

(3.14)

where \( V_{id} \) is differential input voltage and \( I_B \) is the bias current of FVF. As \( V_{id} \) increasing, \( M_2 \) is not in the saturation region and the differential current is still \( I_1 \). The current is strongly depend on \( V_{id} \) and not limited by \( I_B \).

The gain bandwidth equation of the amplifier is given by:
\[ GBW = \frac{g_{m,\text{eff}}}{2\pi C_L} = \frac{2 \times 2 \times g_{m1,2}(g_{5,8}R)}{2\pi C_L} \]  

(3.15)

where \(2 \times 2\) is due to cross-coupling and the size ratio between \(M_{3,4}\) and \(M_{5,6}\). The reference bias current is \(1.25 \mu A\) and \(g_{m5,8}R \approx 2\). With about 10 pF load from ADC, the gain bandwidth is \(> 2\) MHz which is sufficient for our application.

The gain variation prevents saturating the ADC in case there is a strong incoming signal. This is done by implementing an array of capacitor as variable capacitors, \(C_2\) and \(C_4\), whose capacitance can be selected by programming the SPI port. By keeping the same input capacitors, \(C_1\) and \(C_3\), as constant, the row amplifiers’ output loading and the frequency response would stay the same while the gain changes. The output of ADC pre-amplifier drives a larger capacitor when the gain is smaller. This keeps the first pole around a same frequency. The gain can be selected from 2x to 16x by sending a gain control command through SPI interface.

The design parameters of each transistor are shown in the following table.

<table>
<thead>
<tr>
<th>Transistor</th>
<th>Size</th>
<th>Multipiler</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M_1, M_2)</td>
<td>9(\mu m/7.5\mu m)</td>
<td></td>
</tr>
<tr>
<td>(M_3, M_4)</td>
<td>4.5(\mu m/1.5\mu m)</td>
<td></td>
</tr>
<tr>
<td>(M_{11}, M_{13})</td>
<td>4.5(\mu m/1.5\mu m)</td>
<td>(m=2)</td>
</tr>
<tr>
<td>(M_5, M_6)</td>
<td>18(\mu m/3\mu m)</td>
<td></td>
</tr>
<tr>
<td>(M_7, M_8)</td>
<td>7.5(\mu m/7.5\mu m)</td>
<td></td>
</tr>
<tr>
<td>(M_9, M_{10})</td>
<td>9(\mu m/3\mu m)</td>
<td></td>
</tr>
<tr>
<td>(M_{12}, M_{14})</td>
<td>1.5(\mu m/1.5\mu m)</td>
<td>(m=2)</td>
</tr>
<tr>
<td>(M_{15})</td>
<td>18(\mu m/3\mu m)</td>
<td></td>
</tr>
<tr>
<td>(R_1, R_2)</td>
<td>200 k(\Omega)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.3: Size of the ADC pre-amplifier transistors

### 3.3.4 Timing

The Timing diagram is shown in Figure 3.12. All clock phases are non-overlapping. The SAR ADC uses a total of nine clock cycles for 8-bit conversion. The ADC pre-amplifier hold phase completes after the ADC sampling. The amplifier sampling
phase is slightly shorter than hold phase to ensure the next stage completing the sampling phase. Row amplifier hold phase completes one clock later than ADC pre-amplifier sampling phase to provide a stable signal.

![Timing diagram of the row amplifier, ADC pre-amplifier and SAR ADC clock phases.](image)

Figure 3.12: Timing diagram of the row amplifier, ADC pre-amplifier and SAR ADC clock phases.

### 3.3.5 SAR ADC

The 8-bit SAR ADC [32] digitizes the input signal. At the sampling slot, the output bit is replaced by a bit from the pseudo random number generator (PRNG) for wireless synchronization [33]. The PRNG outputs a string of 31 bits, which allows the external receiver to lock in the receiving data by tracking sequence. Figure 3.13 shows the schematic of PRNG and ADC setup.

The ADC module output does not include the channel information. The \( \mu \)-controller will reset the PRNG every time while sending a command to switch the recording channel. The external receiver will detect the PRNG sequence change and add the channel information to the receiving data. This can reduce the amount of data transmitted by an implant unit and power consumption.
3.3.6 Stimulator

The stimulator injects or sinks current into or from neurons, and the amplitude of the current determines the stimulation strength. Based on [10], the circuit diagram for the stimulator is shown in Figure 3.14. The common mode level is floated by using a pair of current sources to equally source and sink. The current is generated from a 3-bit DAC. The control bits stored in SPI port set the enable, strength, and polarity. The advantage of this design is that there is no charge remaining in the tissue causing damage.
The 3-bit DAC has a range from 50 to 350 $\mu$A. The reference current is generated and magnified from a bandgap voltage reference. Although the large magnification reduces the accuracy of the absolute current value, the relative current strength difference is sufficient to provide the required stimulation waveform. In order to lower the power consumption, the DAC is shared by all 4 channels which means only one channel can be stimulated at a given time.

The cross-bar switches, $S_1$ $S_4$, change the polarity of current flow through the electrodes. The tissue resistance, usually a few k$\Omega$, is in series with the switches ON-resistance which might limit the output swing if the switch transistors are not wide enough. The pair electrodes at each channel are shared by recording amplifiers and stimulator. While stimulation, the input of recording amplifier requires isolated from the electrode. The amplifier inputs at the stimulating channel are tied to reference voltage when activating stimulator. This not only avoids saturating the amplifier but also aids a fast recovery when recording is again turned
3.3.7 SPI Interface

The control command programs through the SPI port and the decoder to set a function of the ASIC. Figure 3.15 shows the timing diagram of this 8-bits SPI interface. The first bit (b0) select the pre-amp high pass filter to change the low cutoff frequency. The gain control bits b[1 : 3] defines the gain of ADC pre-amp or the strength of stimulation current depending on the stimulation select bit (b7). B[4 : 5] are column select and row select. The recording or stimulation channel is setup by both bits. The polarity select bit (b6) define the polarity of stimulation current. The stimulation select bit (b7) controls the chip status in either recording mode or stimulation mode.

![SPI Interface Diagram](image)

Figure 3.15: A timing diagram of SPI interface.

After receiving a command, the SPI module would switch the function at the falling edge of ADC sample clock. This can ensure the ADC has valid last output while changing recording channel or switching between recording and stimulation.

3.4 Testing Result

The ASIC was fabricated in the ON-semiconductor 0.5μm, 2-poly and 3-metal process. The die photo is shown in Figure 3.16. The IC size is 3 mm × 1.5 mm, and is packaged in a TQFP 64 pins open cavity plastic package. The testing bench
includes a micro-controller as master clock generator and create the SPI control command to switch in between different chip functions.

Figure 3.16: Die micro-photograph of the ASIC.

3.4.1 Reference Voltage

The bandgap reference voltage is measured across different chips as 1.25V, shown in Figure 3.17(a). Figure 3.17(b) shows the voltage regulator output with different supply voltage to simulate the lithium ion battery operating range. A 500 Ω resistor is added at the output to stimulate the circuit loading. The voltage stays constant at $\approx 2.6V \pm 0.05$ V as the battery voltage higher than 3V.

3.4.2 Amplifier

Figure 3.18 shows the frequency response of a neural pre-amplifier. The high pass filter cut off frequencies are measured as 30 Hz and 50 Hz with different cut off frequency setting bit. The desired gain is 50, and the measurement is 33.9 dB. The high cut off frequency is measured as 2 kHz. This is because the measurement pad is connecting to the amplifier output directly without any buffer. The pre-amplifier total loading capacitor from parasitic capacitor from bond wire, bond pad and package plus the loading capacitor of scope probe is estimated at least 15
Figure 3.17: Measurement results of (a) the bandgap voltage reference output, and (b) regulator output with supply voltage swing.
pF. The extra loading cap lower down the high cut off frequency. The real corner frequency of pre-amplifier should be very close to the design value which is 8 kHz.

![Figure 3.18: Measured frequency response of the pre amplifier](image)

The row amplifier has flat gain of 23.5 dB from low frequency to high cut off frequency at 5.66 kHz. The ADC pre-amplifier has gain of 23.6 dB with highest gain setting and 5.97 dB with smallest gain setting. The measured frequency response plots are shown in Figure 3.19(a) and 3.19(b).
Figure 3.19: Measured frequency response of (a) the row amplifier and (b) the ADC pre amplifier
3.4.3 Stimulator

A 3 kΩ resistor is placed across the input and reference electro pads to simulate the slug body resistance. The currents are measured under different amplitude setting and for different channels. The measurement results are shown in Figure 3.20. Due to the large gain from the bias circuit reference circuit, the absolute output current is sensitive to fabrication process and varied from chip to chip. The linearity is considered adequate for this application.

![Figure 3.20: Measured current with different amplitude setting at each channel.](image)

3.5 Summary

The design and test of ASIC signal path blocks are presented in details in this section. The gain of each stage of the amplifier was tested. Due to some layout errors, the SAR ADC requires a reset after each sample.
Chapter 4

Conclusion and Future Work

4.1 Conclusion

A system capable of wirelessly recording and stimulating neural signals in a salt-water environment was developed. A fully digital compression algorithm that can be implemented in an 8 bit microprocessor is evaluated, and the algorithm showed average 20:1 compression ratio with pre-recorded neural signals. The reconstructed waveform has reduced noise, was highly correlated to the original signal, and preserved important features for visual identification by researchers. The neural recording and stimulating ASIC was designed to have a specific gain for each amplifier stage and the ability to switch between recording and stimulation. A programmable low cutoff frequency for the neural pre amplifier was designed and tested. A on chip voltage regulator provided the ability for consistent circuit performance even with varying lithium battery voltage through its life cycle.
4.2 Future Work

The pseudo differential setup of the first stage amplifier was chosen to reduce the area of the input capacitors for future channel expansion. However, this also created an offset between each amplification stage, and reduced the dynamic range of the signal. To implement a fully differential signal path, with minimum area, high density MOSCAP can be considered [34]. Although the capacitance linearity is moderate (< 3.5%), with careful layout and ratio only architecture, the MOSCAPs should be sufficient.

The neural pre amplifier and row amplifier should be changed from single ended output to fully differential. Testing showed that the amplifier offset voltage is large enough to saturate the ADC pre amplifier even with high DC gain design. The noise measurement also showed the original architecture could not meet the requirement.

To achieve sensing of small input signal, eliminate channel to channel cross talk, and minimize the noise created by long signal traces from different gain stages, the signal path should be fully differential. Although fully differential implementation requires more silicon area, the low noise advantage is crucial.

For long term in-vivo testing, the battery life has to be further extended. One solution is to use a re-chargeable lithium ion battery instead. The small size of the implant unit can maintained without sacrificing the operating time. The battery recharging circuit developed by Feng-Hsu Lin [35] has been ported to the same ASIC and requires further testing.
Bibliography


[25] “North america t1 framing information.”


