A CLOSED-LOOP BIDIRECTIONAL BRAIN-MACHINE INTERFACE SYSTEM FOR FREELY BEHAVING ANIMALS

Xilin Liu

A DISSERTATION

in

Electrical and Systems Engineering

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

_____

Jan Van der Spiegel
Professor of Electrical and Systems Engineering

Graduate Group Chairperson

_____

Alejandro Ribeiro, Rosenbluth Associate Professor of Electrical and Systems Engineering

Dissertation Committee
Nader Engheta, H. Nedwill Ramsey Professor of Electrical and Systems Engineering
Firooz Aflatouni, Skirkanich Assistant Professor of Electrical and Systems Engineering
Timothy H. Lucas, Assistant Professor of Neurosurgery
Naveen Verma, Associate Professor of Electrical Engineering

A CLOSED-LOOP BIDIRECTIONAL BRAIN-MACHINE

INTERFACE SYSTEM FOR FREELY BEHAVING ANIMALS

Acknowledgements

Benjamin Franklin, the founder of this amazing university, used to say: "either write something worth reading, or do something worth writing." I find myself motivated and encouraged by this saying in both my research and my life. Penn is an incredible place, where I had the privilege to work with the brightest people. The immense guidance, support, encouragement, and friendship I received not only helped me accomplished this thesis, but also had a profound impact in my life. So, let me start by thanking them.

First and foremost, I must express my gratitude to my advisor and my role model Prof. Jan Van der Spiegel. Jan is the most brilliant, generous, and humble person I know. Professor, thank you for always guiding and encouraging me. Whenever I had problems in my research, I knew you would always be there to support me. I have so much to learn from you, and I only hope I would be able to carry a small portion of your wisdom.

I am grateful to my thesis committee members, Prof. Nader Engheta, Prof. Timothy H. Lucas, Prof. Firooz Aflatouni, and Prof. Naveen Verma. Having the opportunity to discuss my research with such outstanding scientists is my great honor. Thank you for all your feedback, inspiration, and support. I could only wish to have more of your fantastic advices.

I am especially thankful for Dr. Milin Zhang and Dr. Andrew G. Richardson for all their help during my research. I can't imagine accomplishing this thesis without them. Milin, thank you for bringing me into the world of integrated circuit design. You are my mentor in both research and my life, and I always trust and lean on you

as my big sister. Drew, thank you for making all of the amazing research happen! I'm indebted for your ideas and feedbacks all these years. You taught me how to do fantastic scientific research by example.

I am also thankful to my collaborators and friends. I would like to thank Dr. Xiaotie Wu and Dr. Matt Hongjie Zhu for their help and insightful discussion in my circuit designs. Basheer Subei, Rohit Dureja, Jacob Sacks offered great help in the embedded system programming and system integration. Hanfei Sun, Tian Qiu, Wanqing Xin helped in the digital implementation and circuit layout. The students and fellows in the Translational Neuromodulation Laboratory at Penn carried most of the animal experiments presented in this thesis, they are Yohannes Ghenbot, Sam DeLuccia, Srihari Y. Sritharan, Pauline K. Weigand, Ivette Planell-Mendez, Solymar Torres Maldonado, Gregory Boyek, and many others. Thank you all for the help, support and friendship!

I would also like to thank Dr. Dengteng Ge and Prof. Shu Yang for the help in the device development, and Dr. Tao Xiong, Prof. Peter S. Chin, Prof. Trac D. Tran, and Prof. Ralph Etienne-Cummings for their collaboration in the work of compressed sensing recording front-end.

Finally, Im blessed to have the endless love from my family. Mom and Dad, saying thank you is too inconsequential for everything you give me. I only hope I could spend more time with you all these years. I want to tell you that I feel your unconditional love with me all the time. I'm proud to be your son and I know you are proud of me too. Another person behind this accomplishment undoubtedly is my girlfriend, Shurui Zhou. Shuishui, thank you so much for being in my life. In all the happy, difficult, joy, and frustrating moments, you are there for me. I want you to

know that means everything to me, and I wish I could do the same for you, from my heart.

The past five years of my life was full of passion and imagination. To be able to work with the most intelligent people in a world-class university is an unparalleled experience, which changed my vision, and my life. In the end, I want to thank the days and nights of hard work, which pushed myself to be a better person. To all aforementioned, thank you, for giving the courage, confidence, humility, knowledge, and ability in pursuing my future.

ABSTRACT

A CLOSED-LOOP BIDIRECTIONAL BRAIN-MACHINE INTERFACE SYSTEM FOR FREELY BEHAVING ANIMALS

Xilin Liu

Jan Van der Spiegel

A brain-machine interface (BMI) creates an artificial pathway between the brain and the external world. The research and applications of BMI have received enormous attention among the scientific community as well as the public in the past decade. However, most research of BMI relies on experiments with tethered or sedated animals, using rack-mount equipment, which significantly restricts the experimental methods and paradigms. Moreover, most research to date has focused on neural signal recording or decoding in an open-loop method. Although the use of a closed-loop, wireless BMI is critical to the success of an extensive range of neuroscience research, it is an approach yet to be widely used, with the electronics design being one of the major bottlenecks. The key goal of this research is to address the design challenges of a closed-loop, bidirectional BMI by providing innovative solutions from the neuron-electronics interface up to the system level.

Circuit design innovations have been proposed in the neural recording front-end, the neural feature extraction module, and the neural stimulator. Practical design

issues of the bidirectional neural interface, the closed-loop controller and the overall system integration have been carefully studied and discussed. To the best of our knowledge, this work presents the first reported portable system to provide all required hardware for a closed-loop sensorimotor neural interface, the first wireless sensory encoding experiment conducted in freely swimming animals, and the first bidirectional study of the hippocampal field potentials in freely behaving animals from sedation to sleep.

This thesis gives a comprehensive survey of bidirectional BMI designs, reviews the key design trade-offs in neural recorders and stimulators, and summarizes neural features and mechanisms for a successful closed-loop operation. The circuit and system design details are presented with bench testing and animal experimental results. The methods, circuit techniques, system topology, and experimental paradigms proposed in this work can be used in a wide range of relevant neurophysiology research and neuroprosthetic development, especially in experiments using freely behaving animals.

Contents

Acknowledgements **iii**

List of Tables **xiii**

List of Figures **xiv**

1 Introduction **1**
1.1 Background and Motivation 1
1.2 Review of Prior Work 6
1.3 Overview of Bidirectional Closed-loop Brain Machine Interface 18
1.4 Thesis Outline and Contributions 25

2 Neural Recording Front-end Design **29**
2.1 Introduction 29
2.1.1 Signal Characteristics 31
2.1.2 Design Specifications 32
2.2 Design of a Low-Noise Neural Amplifier 36
2.2.1 Review of Prior Work 36
2.2.2 Circuit Implementation 41
2.2.3 Measurement Results 48
2.3 A Pre-whitening Neural Amplifier 53
2.3.1 Introduction 53
2.3.2 Analysis of the Pre-whitening Neural Amplifier Design 55
2.3.3 Circuit Implementation 66
2.3.4 Measurement Results 67
2.4 Design of a Low-Power Analog-to-Digital Converter 73
2.4.1 Introduction 73
2.4.2 Circuit Implementation 74

2.4.3 Measurement Results 78
2.5 A Compressed Sensing Neural Signal Acquisition System 83
2.5.1 Introduction 83
2.5.2 A Brief Background of Compressed Sensing 85
2.5.3 System Overview 87
2.5.4 Circuit Implementation 89
2.5.4.1 Energy Efficient Analog Front-end 89
2.5.4.2 On-chip Wireless Power and Data Link 93
2.5.4.3 External Wireless Relay Board 94
2.5.5 Measurement Results 96

3 Neural Feature Extraction 103
3.1 Introduction 103
3.2 Natural Logarithmic Domain Neural Energy Extraction 106
3.2.1 Introduction 106
3.2.2 System and Circuits Implementation 107
3.2.2.1 Design of the GmC Filter 109
3.2.2.2 Biasing Current Generation 111
3.2.2.3 Design of the Biquad filter 114
3.2.2.4 Multiplier and Integrator 116
3.2.3 Measurement Results 117
3.3 Action Potential Detection 123
3.3.1 Introduction 123
3.3.1.1 Integrate and Fire Model 123
3.3.1.2 Review of Action Potential Detection Methods . . . 125
3.3.2 Circuit Implementation 127
3.3.3 Experimental Results 132
3.4 Matched Filter for Neural Feature Extraction 137
3.4.1 Introduction 137
3.4.2 Matched Filter and Pre-whitening for Optimum Correlation Detection 138
3.4.3 Methodologies 141
3.4.4 Experimental Results 144
3.4.4.1 Detection of Synthesized Signal 144
3.4.4.2 Detection of Recorded Neural Signal 146
3.4.4.3 Application of Compressed Sampling 148

4 Neural Stimulator Design 151

4.1 Introduction . . . 151
4.1.1 Background of Neurostimulation . . . 152
4.1.2 Electrode and Electrolyte Interface . . . 155
4.2 Overview of Electrical Stimulator Design . . . 158
4.2.1 Methods of Stimuli Generation . . . 159
4.2.2 Stimulation Waveform and Electrodes Configuration . . . 161
4.2.3 Methods for achieving Charge Balance . . . 164
4.2.3.1 Matching and Calibration . . . 165
4.2.3.2 Passive and Active Discharge . . . 166
4.3 Design of A General-Purpose Stimulator . . . 168
4.3.1 Architecture of the Stimulator . . . 168
4.3.2 Circuit Implementation . . . 169
4.3.3 Measurement Results . . . 172
4.4 An Energy Efficient Net-Zero Charge Neural Stimulator . . . 175
4.4.1 Introduction . . . 175
4.4.2 Motivation and Innovation . . . 177
4.4.2.1 Net-zero Charges Stimulation . . . 177
4.4.2.2 Adaptive Driving Voltage . . . 180
4.4.2.3 Arbitrary Channel Configuration . . . 182
4.4.3 Circuit Implementation . . . 183
4.4.3.1 System Architecture . . . 183
4.4.3.2 Output Stage with Dynamic Element Matching . . . 185
4.4.3.3 Feed-forward Error Compensation Comparator . . . 188
4.4.4 Experimental Results . . . 192
4.4.5 Conclusion and Future Work . . . 198
4.5 Conclusion . . . 198

5 Bi-directional Neural Interface and Closed-loop Control 200
5.1 Introduction . . . 200
5.2 Stimulation Artifacts in the Bi-directional Neural Interface . . . 202
5.2.1 Introduction . . . 202
5.2.2 Review of Prior Work . . . 203
5.2.3 Analysis of Stimulation Artifacts . . . 205
5.2.3.1 Origins of Stimulation Artifacts . . . 205
5.2.3.2 Configuration of the Interface Circuits . . . 209
5.2.3.3 Practical Issues in the Electrical Circuits . . . 211
5.2.4 Methods . . . 214
5.2.5 Experimental Results . . . 216

5.2.5.1 *In-Vitro* Experiment 216
5.2.5.2 *In-Vivo* Experiment 222
5.2.6 Conclusion 224
5.3 Closed-loop Neural Interface System 228
5.3.1 Introduction 228
5.3.2 Mechanism of Closed-loop Neural Interface System 230
5.3.3 Design of a Closed-loop Neural Interface with a PID controller 233
5.3.3.1 Introduction 233
5.3.3.2 System and Circuit Implementation 234
5.3.3.3 Experimental Results 237

6 System Integration and Animal Experiments 243
6.1 Introduction 243
6.2 The PennBMBI: A General Purpose Experimental Platform 246
6.2.1 Introduction 246
6.2.2 System Overview 247
6.2.3 Hardware Implementation 249
6.2.3.1 Neural Signal Analyzer 249
6.2.3.2 Neural Stimulators 252
6.2.3.3 Body Area Sensors 254
6.2.3.4 Computer Interface 255
6.2.4 Experimental results 257
6.2.4.1 Bench Testing 257
6.2.4.2 *In-Vivo* Testing 261
6.3 The Watermaze 264
6.3.1 Introduction and Background 264
6.3.2 System Overview 266
6.3.3 Hardware Implementation 267
6.3.3.1 Design of the Watermaze Stimulator 267
6.3.3.2 Electrode and electrode connector 275
6.3.3.3 Image Sensor and Computer Interface 276
6.3.4 Software Implementation 277
6.3.4.1 Communication Protocol 277
6.3.4.2 Animal Tracking 282
6.3.4.3 User Interface 282
6.3.5 Experimental Results 285
6.4 Bidirectional Neural Interface for Freely Behaving Macaque 291
6.4.1 Introduction and Background 291

6.4.2 System Integration 292
6.4.3 Experimental Results 300

7 Conclusion and Future Works **315**
7.1 Summary of the Work 315
7.2 Future Work 319

A Acronyms **321**

B Propositions **323**

List of Tables

1.1 Survey of Bidirectional Neural Interface Designs (Part I) 9
1.2 Survey of Bidirectional Neural Interface Designs (Part II) 10
1.3 Key Features of the Proposed Bidirectional BMI 21

2.1 Summary of Specifications for Neural Recording [1] 33
2.2 Comparison of capacitive and resistive gain elements [2] 37
2.3 Summary of commonly used low-noise OTA [3, 4] 40
2.4 The Transistors' Parameters for the OTA Design 43
2.5 The Neural Front-end Specifications Summary 51
2.6 Comparison with Prior Works 52
2.7 Specification Summary of the 10-bit SAR ADC 81
2.8 Chip Specifications Summary 101
2.9 Comparison with State-of-the-art Works 102

3.1 The Measured AP Detection Accuracy of Signal With Different SNR 133

4.1 Chip Specification Summary 197

5.1 *In-vitro* Experiments for Stimulation Artifact Study 217
5.2 Stimulation Artifacts with Common Ground 227
5.3 Effects of PID Parameters [5] 234

6.1 Power consumption of the sensor node 255
6.2 Specifications of the PennBMBI system 260
6.3 Comparison of MCUs used in this work 294
6.4 Key Specifications of the Bi-directional Neural Interface SoC 304
6.5 Comparison with the State-of-the-Art Bidirectional Neural Interface Designs 314

List of Figures

1.1 Envisioned bidirectional clinical hand neuroprosthesis. Motor function is restored through brain-controlled electrical stimulation of hand muscles, and somatosensation is restored through sensor controlled electrical stimulation of the brain. The motor pathway replacement has been extensively studied, and this work is focused on sensation restoration. 4

1.2 The historical trend for publications of BMIs over the past 15 years. Specific categories of BMIs including invasive approach, closed-loop study, and bidirectional interface are plotted for comparison. 6

1.3 The trend for the percentage of closed-loop and bidirectional designs among all BMI publications. Second order polynomial fitting curves are plotted for showing the trend. 7

1.4 The block diagram and signal routing of the Neurochip-2 system developed by S. Zanos and Eberhard E. Fetz *et al.* from the University of Washington in 2011 [6]. Reused with permission, copyright 2011, IEEE. 12

1.5 Electrical device architecture and system partition for the prototype developed by A. G. Rouse and T. J. Denison *et al.* from Washington University, St. Louis and Medtronic Inc. in 2011 [7, 8]. Reused with permission, copyright 2011, IEEE. 13

1.6 The system architecture of the activity-dependent cortical microstimulation (ICMS) chip proposed by M. Azin and Pedram Mohseni *et al.* from the Case Western Reserve University in 2011 [9]. Reused with permission, copyright 2011, IEEE. 14

1.7 The system architecture of the SoC proposed by Wei-Ming Chen *et al.* from Nation Chiao Tung University in 2014 [10]. Reused with permission, copyright 2014, IEEE. 16

1.8 The overall system block diagram of the SoC proposed by Hyo-Gyuen Rhew and Michael P. Flynn *et al.* from the University of Michigan, Ann Arbor in 2014 [11]. Reused with permission, copyright 2014, IEEE. 17
1.9 The architecture of (a)the wireless neural recording device developed by Ming Yin, and Arto V. Nurmikko *et al.* from the Brown University [12], and (b) the chronic wireless recording system developed by David A. Schwarz and Miguel A. L. Nicolelis *et al.* from the Duke University [13]. Both systems have been validated in animal behaving experiments. 18
1.10 The block diagram of the proposed generalized closed-loop, bidirectional BMI system 22
1.11 Various closed-loop configurations of the proposed BMI system. . . . 23
1.12 The building blocks of a typical bidirectional closed-loop BMI system and the organization of this thesis. 25

2.1 The sources of the neural signals and their locations relative to the brain, modified from [14]. 31
2.2 The amplitude and frequency characteristics of the neural signal, in comparison with main noise sources. This figured is modified from [15]. 32
2.3 The block diagram of the typical neural amplifiers with (a) capacitive gain element [16] and (b) resistive gain element [17]. The x1 symbol is for a unity gain buffer. 36
2.4 The circuit schematic of commonly used low noise OTAs: (a) the current mirror OTA, (b) the Miller OTA, (c) the folded cascode OTA, and (d) the telescopic OTA. 39
2.5 (a) The block diagram of a chopping amplifier. (b) Illustration of the signal and noise spectrum before and after the chopping. 41
2.6 The block diagram of the neural recording front-end. 42
2.7 The circuit schematic of the MOS pseudo-resistor. 42
2.8 Simulation of the resistance of the MOS pseudo-resistor. 43
2.9 The circuit schematic of the fully differential low-noise OTA with complementary input stage. 44
2.10 Simulation of the input impedance boosting. 47
2.11 The microphotography and layout of one channel of the neural recording front-end. The major building blocks are highlighted. 49

2.12 The measured response of the neural amplifier with a 1kHz sinusoidal signal. A resistor divider of 2kΩ and 1Ω is applied at the output of the function generator, and the amplifier is configured with the maximum gain of 1900. The gain error is 0.37%. 49
2.13 The measured differential-model and common-mode frequency response of the low-noise neural amplifier. 50
2.14 The measured input-referred voltage noise spectrum. An integration under this curve from 1Hz to 7kHz yields a rms noise of 2.55μV. . . . 50
2.15 Illustration of the prewhitening filter. (a) The neural signal displays a $1/f^n$ power characteristic, while the recording front-end has a $1/f$ noise power characteristic. (b) The pre-whitening filter shapes the frequency response of the recording front-end to reduce the overall dynamic range requirement, while still preserves a sufficient SNR. 53
2.16 A 1st order RC highpass filter with noise source. 55
2.17 Noise simulation of RC highpass filters with frequency corners at 10Hz and 100Hz. The capacitor value is set to be 20pF. 56
2.18 The block diagram of a pre-whitening filter after a wideband low-noise amplifier. 57
2.19 The circuit schematic of a pre-whitening filter after a wideband low-noise amplifier. Noise source from the resistor is shown. 57
2.20 An implementation of an active highpass filter. Noise sources from the resistor and the 2nd stage amplifier are shown. 58
2.21 Capacitor coupled neural amplifier. Noise sources are marked in the figure. 61
2.22 Capacitor coupled instrumentation amplifier with a DC servo loop. . 62
2.23 Capacitor coupled instrumentation amplifier with DC servo loop implemented by switched capacitor circuits. 65
2.24 A capacitor coupled chopping amplifier with DC servo loop and input impedance boosting. 65
2.25 The circuit schematic of the designed pre-whitening amplifier. The Pseudo resistors R_A and R_B used in the 1st and 2nd stage is shown in subplot (a) and (b), respectively. 66
2.26 The microphotography and layout of one channel of the pre-whitening amplifier. 67
2.27 The measured frequency response of the pre-whitening amplifier in comparison with the simulation result. 68

2.28 Comparison of the PSD of the pre-whitening amplifier and the original signal. 69
2.29 Comparison of (a) the wideband signal, (b) the measured output of the pre-whitening amplifier, and (c) the reconstructed signal from the pre-whitening amplifier's recording. 70
2.30 Comparison of the spectrum of (a) the original signal and (b) the reconstructed signal from the pre-whitening amplifiers recording. 71
2.31 The architecture of the 10-bit voltage-mode SAR ADC. 75
2.32 A comparison of the waveforms of (a) the conventional switching to V_{CM} procedure, and (b) the monotonic switching procedure. Modified from [18]. 75
2.33 The circuit schematic for the SAR timing generation module. *clk* is the input clock, *clks* is the signal for the sampling switch, *clkc* is the clock for the comparator, and *clk*[x] is for the bit[x] of the DAC. . . 76
2.34 The circuit schematic of the bootstrapped switch. 77
2.35 The circuit schematic of the comparator. (a) A pre-amplifier, and (b) a dynamic latch. 77
2.36 The circuit schematic of a single-to-differential converter. 78
2.37 The layout of the 10-bit SAR ADC with major building blocks highlighted. 79
2.38 Measured DNL and INL of the 10-bit SAR ADC. The worst DNL is -0.49/+0.56LSB, and the worst INL is -0.82/+0.77LSB. 79
2.39 Measured FFT spectrum at 1MS/s with an input tone of 3kHz. The SFDR is 76.54dB and the SNDR is 56dB. The ENOB at 3kHz is 9.01. 80
2.40 Measured FFT spectrum at 1MS/s with an input tone of 493kHz. The SFDR is 71.6dB and the SNDR is 54.6dB. The ENOB at 493kHz is 8.77. 80
2.41 Historical trend for publications using compressed sensing technique in biomedical signal acquisition in the past decade. Data retrieved from Web of Science. 84
2.42 (a) Illustration of the hypothetical chronic neural signal recording system using the fully integrated compressed sensing chip, and (b) the architecture of the chip. 88

2.43 Analog front-end of the proposed system (Part I). The signal chain includes the signal amplification, filtering, voltage-to-current conversion and multiplexed to a shared ADC. The circuit schematic of the (a) low noise amplifier, (b) OTA with extended linear range, and (c) OTA with programmable transconductance. 90
2.44 Analog front-end of the proposed system (Part II). A current-to-voltage conversion with programmable gain and a 10-bit SAR ADC is used to digitize the signal. The boxed windows show the circuit schematic of the (a) comparator and the (b) SAR ADC. 92
2.45 The circuit schematic of the clock recovery and clock division module. 94
2.46 The circuit schematic of the CMOS Schmitt trigger [19]. 94
2.47 Inductive power management module, including active rectifier and LDOs for analog and digital power supplies. (a) circuit schematic of the comparator, (b) bandgap reference, and (c) LDO (start-up circuits are not shown). 95
2.48 The micrograph of the fabricated fully integrated compressed sensing neural recording front-end chip. 96
2.49 The measured frequency response of the low noise amplifier (without filtering stages). 97
2.50 The measured input-referred voltage noise spectrum. 97
2.51 Comparison between the uncompressed sampling results and data restored from different compression ratio (CR). 98
2.52 Comparison of the spectrograms of (a) the uncompressed sampling results and (b) the data restored with a CR of 8. 99
2.53 Photography of an assembled demonstration system. (a) Power and data transmission testing setup across 5mm plastic cap, (b) external transceiver board, (c) implantable device. 99
2.54 A 24-hour continuous recording in the hippocampus of a rhesus macaque during free behavior. 100

3.1 Brain oscillation bands are shown together with the traditional frequency tuning bins in linear steps, and the proposed tuning bins in the natural logarithmic domain. A total of 32 steps in a frequency range from 1Hz to 200Hz is shown for illustration. 107
3.2 The processing flow of the LFP energy extraction. 108
3.3 The diagram of the 16-channel energy extraction module. 108

3.4 The circuit schematic of the transconductance (Gm) block with source degeneration. 110
3.5 The circuit schematic of the transconductance (Gm) block with extended linear range. The biasing current is used to tune the transconductance. 111
3.6 The circuit schematic of the 6-bit current-mode DAC for generating the biasing current for the Gm block. Thick oxide transistors are used to reduce the leakage current. 112
3.7 The circuit schematic of the biasing current generation module. a 64-step natural exponentially spaced biasing current can be generated. . 113
3.8 The circuit schematic of the designed biquad filter. The center frequency and the quality factor of the filter can be tuned independently. 115
3.9 The circuit schematic of the Gilbert multiplier and the integrator. . . 117
3.10 The microphotography and layout of the LFP energy extraction channel. The major building blocks are highlighted in the layout. 118
3.11 The measured frequency response of the biquad filter tuning in the proposed natural logarithmic steps. A total of 16 steps were measured. 118
3.12 The measured the frequency response of the biquad filter with different quality factors. The center frequency is configured at 10Hz. 119
3.13 The measured response of one biquad filter with a synthetic sine wave with frequency sweeping from 0.1Hz to 1kHz. 120
3.14 The measured output of the Gilbert multiplier with an amplitude modulated 10Hz signal. The measurement result is plotted in comparison with the simulation (after a gain calibration). 120
3.15 The measured outputs of the multiplier and the LFP energy integrator (phase shift have been corrected). 121
3.16 The spectrum of a 6-hour continuous recording using the prototype device. The animal was from awake (high-frequency oscillation more active) to sleep (low-frequency oscillation more active). 121
3.17 In-vivo recording in a Rhesus macaque using the designed chip. The extracted energy in four brain oscillation bands (Theta, Beta, Gamma, and Fast) compared with the theoretical computations (dashed lines). 122
3.18 Matlab simulation of the implemented HHM model. 125
3.19 The block diagram of the action potential detection module. 128
3.20 Illustration of the window discriminator for action potential detection principle. 129

3.21 The block diagram of the programmable current mode spike detection unit with integrated programmable amplitude-window discriminator. The filtering state is not shown in this diagram. 129
3.22 The circuit schematic of the current-mode DAC. 130
3.23 The circuit schematic of the current-mode comparator. 131
3.24 The micrograph and layout of the designed analog action potential detection module. 132
3.25 The action potential signal used for testing the designed detection module. The real action potentials are marked by triangle markers. (a) The original signal with an SNR of 25.7dB. (b) Testing signal adding white noise and artifacts with an SNR of 20dB. (c) Testing signal adding white noise and artifacts with an SNR of 15dB. 134
3.26 A cluster analysis of the action potentials from two neurons. (a) Normalized maximum and minimum amplitudes are calculated and used as two features for the analysis. (b) The action potentials are labeled with different colors according to the classification results. 136
3.27 The block diagram of the matched filter. 138
3.28 The block diagram of the matched filter in combination with the pre-whitening filter for the correlation optimization. 140
3.29 (a) A 5-sec recording from an anesthetized rat. (b) Power spectrum density of the recording. (c) Phase-amplitude coupling analysis of the signal. 141
3.30 (a) The neural feature waveform (template). The waveform has more time in the "up-state" than the "down-state", so it is not an ideal sinusoid wave. (b) Frequency analysis of the neural feature. 142
3.31 A comparison of the frequency response of different filters for extracting the slow oscillation. 143
3.32 Examples of the synthesized test signals with SNR ranging from 0.2 to 6. The last one is the template. 145
3.33 Detection accuracy of different filters for signals with different SNR (0.2 to 7 in a step of 0.2). The detection accuracy for each SNR step was an average of 100 trials with random pink noise. A total of 3,500 trials were tested for each filter in this experiment. 146
3.34 Examples of the randomly selected real neural signal segment from the data bank. The last one is the template. 147

3.35 Detection accuracy of different filters for 100 signal segments randomly selected from the data bank. Each of the signal segments is inserted into 100 random pink noise trials. A total of 10,000 trials were tested for each filter in this experiment. The detection result shows that the pre-whitening filter improves the matched filter's performance in low SNR. 148
3.36 Comparison of the detection accuracy of different filters. The experimental results show that the matched filter with pre-whitening has the best performance. The matched filter with pre-whitening and a compression ratio of 64x still has a better performance than the matched filter without pre-whitening. 149

4.1 Matlab simulation of the membrane potential with dynamic simulation current based on the integrate-and-fire model. 154
4.2 Illustration of the electrode and electrolyte interface, modified from [20]. (a) The physical representation, and (b) a simplified electrical circuit model. 155
4.3 The measured impedance of two types of custom made tungsten electrodes used in this research. (a) and (b) shows the electrode with a diameter of 75μm and 50μm, respectively. Each figure shows an overlay of the measurements of 10 electrodes, and a fitting curve in red. The parameters of the fitting models are labeled in the figures. . . . 156
4.4 The design considerations and trade-offs of an electrical neural stimulator. 158
4.5 Illustration of a typical biphasic stimulation waveform with the parameters marked. I_S: stimulation current, I_R: reversal current, T_S: stimulation phase time, T_R: reversal phase time, T_D: discharging phase time, T_P: phase interval, T_I: pulse interval, T_L: pulse group interval. 162
4.6 Illustration of (a) monopolar and (b) bipolar stimulation methods. A voltage-regulated stimulation is used for illustration, but a current-regulated stimulation can be applied in the same way. 163
4.7 The block diagram of the neural stimulator. The stimulator includes a digital part designed in 1.8V, and an analog part designed in 1.8/5V. 168

4.8 The circuit schematic of the proposed multi-mode stimulator site. Each site consists of: i) a current-mode DAC which generates a reference for the output current, ii) a current driver including current sink and source output stages with high output impedance, and iii) high voltage switches with level-shifters. Each site demultiplexes to 4 channels, and provides near-simultaneous stimulation. 169
4.9 The circuit schematic of the level shifter. 170
4.10 The timing for generating a monopolar stimulation at the electrode X. The DAC # belongs to the site where the electrode X locates. (I_S: stimulation current, I_R: reversal current, T_S: stimulation phase time, T_R: reversal phase time, T_D: discharging phase time, T_P: phase interval. XXX means the DAC can be in any value.) 171
4.11 The timing for generating a bipolar stimulation between the electrode X and Y. The DAC # and DAC $ belong to the site where the electrode X and Y locates, respectively. Notice that they can be in a same site. Parameter definitions are the same as in Fig. 4.10. 172
4.12 The microphotography and the layout of the neural stimulator. The major building blocks are highlighted in the layout. 173
4.13 The measured stimulator output current versus output voltage. . . . 173
4.14 The measured output currents from the current source and current sink of the stimulator output stage. The non-linearity of the source and sink current is 0.31% and 0.37%, respectively. 174
4.15 The measured simultaneous stimulation output from four independent channels. The boxed window shows the measurement of a single pulse in high resolution. 175
4.16 (a) The traditional charge-balancing method matches the stimulation and reversal currents. The ideal charge curve on the electrode is plotted in a dashed line. The practical charge curve varies from the ideal curve due to the irreversible reaction and chemical products diffusion. (b) this work terminates the reversal phase based on the monitoring the net-zero charge point. $\phi 1$: stimulating phase, $\phi 2$: interval phase, $\phi 3$: reversal phase, $\phi 4$: discharge phase 178
4.17 The model for simulating the effects of non-ideal charge diffusion. . . 179

4.18 The simulation of the voltage across the blocking capacitor C_B (Fig. 4.17) during a symmetrical single pulse. Ideal current sources with equal stimulation current amplitude and time are used. 500kΩ, 1MΩ, and 5MΩ resistors are used to mimic the impedance between the stimulation location to the tissue ground. 179
4.19 The simulation of the voltage across the blocking capacitor of a 10 symmetrical pulse train without discharging. 500kΩ, 1MΩ, and 5MΩ resistors are used to mimic the impedance between the stimulation location to the tissue ground. The charges build up even using the ideal current sources with equal amplitude and time. 180
4.20 Illustration of the adaptive driving voltage stimulation. 181
4.21 The channel configuration of (a) the traditional stimulation, and (b) the proposed work. Arbitrary channel configuration is feasible without pre-calibration. 183
4.22 The block diagram of the net-zero charge neural stimulation system. The system consists of an analog core, a digital module, and off-chip power management units. 183
4.23 The working flow chart of the proposed stimulation strategy. 185
4.24 Circuit schematic of (a) basic current generator, (b) output stage with voltage-controlled transistor, modified from [21], (c) output stage with digital-set DAC, modified from [22]. 186
4.25 (a) The proposed output stage with current set current-mode DAC. (b) The OTA used in the work. 187
4.26 100 runs monte-carlo simulation of the different output stage architectures with mismatch and process variation. 188
4.27 Analysis of the delay of the comparator for determinating the zero-crossing point of the blocking capacitor. 189
4.28 The circuit schematic of the comparator consisting of a 3-stage pre-amplifier and latch. The pre-amplifier has auto-zero calibration, and the latch has a 4-bit DAC for calibration. 191
4.29 The circuit schematic of the switched capacitor circuit used to discharge the blocking capacitor. The circuit is also used to amplify the residue voltage for the calibration purpose. The amplified residue voltage is compared with two pre-defined safe voltage window. If the residue charge is out of the safe window, the calibration DAC of the comparator will be changed accordingly. 191

4.30 The micrograph of the fabricated stimulator chip. The occupied silicon area is 3×1.5mm^2. 192
4.31 The measured current from the output stage versus the output voltage. The embedded figure shows a zoom-in plot from 0 to 0.3V. 193
4.32 The measured generated stimulation waveforms with a high impedance load. Channel 1 and channel 2 measures the output of the WE and CE electrode, respectively. The Math channel calculates the difference between the two channels. Channel D3 to D6 show the states of the digital module. 194
4.33 (a) The measurement of a stimulation pulse during the driving voltage adjustment. The load model is given with the measurement points highlighted. (b) The measured stimulating and reversal currents during the driving voltage adjustment. 195
4.34 The measurement of biphasic stimulation in the saline solution using the traditional method (red) and the proposed method (blue). (a) shows a 5-min continuous stimulation without discharge. (b) and (c) show an overlay of 20 measurements of the voltage across the blocking capacitor and the derived stimulation current. 196
4.35 *In-vivo* experiment performed on a sedated rat. (a) Whisker movements, as measured by an optical micrometer, were reliably evoked. (b) Whisker displacements were a function of current intensity. 197

5.1 The circuit model for the bi-directional neural interface with (a) shared grounds and (b) separated grounds. 207
5.2 Configuration of typical multiple channels neural recording front-end. (a) single-ended recording configuration, and (b) differential recording configuration. 209
5.3 The configuration of typical neural stimulation back-ends. (a) shows a monopolar stimulation configuration, (b), (c), and (d) are bipolar configurations. (a) and (b) are using a Type-I stimulator, (c) and (d) are using a Type-II stimulator. 212
5.4 Two common ground configurations for the recorder and the stimulator: (a) dual-supply and (b) single-supply. In a dual-supply design, the grounds of all circuit modules are connected together, while in a single-supply design, the lowest supply in each module is connected together. The design is also limited by the bulk of the CMOS technology. . . . 213

5.5 The block diagram of the bench testing board. The supplies and grounds for the recorder, the stimulator and the digital modules are intentionally separated on the board. 214
5.6 The photo of the assembled bench testing board. The dimension is 216mm × 171mm. 215
5.7 Electrode setup for the stimulation artifacts experiments. 217
5.8 The measured stimulation artifacts of a Type-I stimulator with common ground between the recorder and the stimulator. (a) monopolar stimulation and single-ended recording, (b) monopolar stimulation and differential recording, (c) bipolar stimulation and single-ended recording, and (d) bipolar stimulation and differential recording. 219
5.9 The measured stimulation artifacts of a Type-I stimulator with separate grounds between the recorder and the stimulator. (a) monopolar stimulation and single-ended recording, (b) monopolar stimulation and differential recording, (c) bipolar stimulation and single-ended recording, and (d) bipolar stimulation and differential recording. 220
5.10 Stimulation artifacts of a Type-II stimulator with common ground between the recorder and the stimulator. The stimulator uses a current sink in (a) and (b), and a current source in (c) and (d). Recorder is configured to do single-ended recording in (a) and (c), and differential recording in (b) and (d). 221
5.11 Stimulation artifacts of a Type-II stimulator with separate ground between the recorder and the stimulator. Recorder is configured to do single-ended recording in (a) and differential recording in (b). 222
5.12 The photo of the setup of the *in-vivo* experiment. A female Long-Evans rat received two implants, one in the motor cortex, the other in the sensory cortex. The measurement was conducted using the testing board presented in section 5.2.4. 223
5.13 Bi-directional neural interface experiment in a Long-Evans rat. Local field potential was recorded in the motor cortex while stimulating the sensory cortex. (a) When the grounds of the recorder and stimulator were connected together, there was a large artifact and long recovery time; (b) when the grounds were separated, the artifact is minimized. An evoked potential was shown after the stimulus artifact. 224

5.14 Bi-directional experiment in a Rhesus macaque. (a) The grounds of the stimulator and the recorder are shorted together, the stimulation artifact saturates the amplifier, and takes hundreds of milliseconds to return to the pre-stimulus potential. (b) The grounds of the stimulator and recorder are separated, and only minor artifact appears in the recording. (c) A 200Hz filter is applied to the recording in (b), which completely removes the artifacts. 225
5.15 (a) The block diagram of a typical closed-loop control system. (b) The typical block diagram of a bi-directional closed-loop neural interface system. The neural recorder works as the sensor, and the neural stimulator works as the actuator. 228
5.16 The block diagrams of different closed-loop BMI applications. The functions of the proposed neural interface system are shown in red. (a) Standard bi-directional BMI for a prosthetic arm. (b) Same as (a) but with improved sensory encoding method using a PID controller. (c) Bi-directional BMI to re-animate paralyzed arm by decoding desired arm trajectory. (d) Same as (c) but decoding motor goal and implementing arm trajectory with a PID controller. 231
5.17 The block diagram of the overall system using the designed PID controller. 235
5.18 The circuit schematic of the designed PID controller. 236
5.19 The microphotograph and the layout of the PID controller module. . 237
5.20 The 2nd order RC ladder network used for testing the PID controller. 238
5.21 The measured transient response of the PID controller in different configurations. The ratio of the K_p, K_i, and K_d are (a) 1-2-0 (b) 2-1-0 (c) 1-0-1 and (d) 1-1-1. 239
5.22 Diagram of the testing configuration of the PID controller. The major blocks used are the PID controller, a stimulator, an action potential detector, a neuron model, and a lossy integrator for finding spike rate. The on-chip stimulator is configured in the test mode to output continuous current. The lossy integrator is implemented in the MCU. The neuron is modeled by another MCU. 239
5.23 The measured transient response of the dynamic neuronal clamp with different PID parameters. Relative values of P, I and D components are shown on the right. 241

5.24 The measured transient response of the dynamic neuronal clamp with different references. 242

6.1 The BMBI system include four kinds of devices: neural signal analyzer, neural stimulator, body-area sensor node and computer interface. All devices can be configured wirelessly in a computer GUI. Possible closed-loop operation between devices are shown. 247
6.2 Photograph of the PennBMBI neural signal analyzer (NSA) in front, rear, and side view. The wireless module and Micro SD card are not shown in the front view. 249
6.3 Architecture of the analog front-end. V_{CM} is the common mode voltage. A1∼A4 forms the instrumentation amplifier with high pass filter. A5 works as the second gain stage with configurable gain and low pass filter. The third gain stage has a programmable gain. The ADC digitizes the amplified neural signals at configurable sampling rates and resolution. A6 and A7 are shared by the four channels to drive the shield potential and tissue ground. 250
6.4 Photograph of the neural stimulator. The stimulator includes a current driving back-end, a MCU with integrated DAC and ADC, a wireless transceiver, a power management unit. 253
6.5 Circuit schematic of the high compliance voltage current source output stage. Arbitrary stimuli waveform is generated by a digital to analog converter. V-to-I gain is programmable to provide a high dynamic range. 254
6.6 The photograph of sensor node in comparison with a Quarter. 255
6.7 Matlab based Graphic User Interface (GUI). Six major panels are included in the GUI, which are: 1) PC configuration; 2) recording device configuration; 3) stimulator configuration; 4) body-area sensors configuration; 5) closed-loop configuration; 6) display windows. 256
6.8 Input referred noise spectrum of the analog front-end. 257
6.9 The measured frequency response of the analog front-end in different configurations. (blue) 10 to 200Hz with a gain of 66dB, (red) 300 to 6kHz with a gain of 66dB, (magenta) 10 to 200Hz with a gain of 78dB. 258
6.10 The measured output current for different loads (1kΩ, 2kΩ, ..., 8kΩ). (a) shows the standard deviation of the output current across different loads. (b) shows the average current mismatch between anodic and cathodic electrodes across different loads. 259

6.11 Measured stimulation pulses with different amplitudes in Sodium Chloride solution. 260
6.12 Measured wireless closed-loop operation from the neural signal analyzer and the stimulator. Zoomed-in view of (a) is shown in (b). 261
6.13 Measured wireless closed-loop operation from sensor node and deep brain stimulator. 262
6.14 Action potentials recorded by the neural signal analyzer. Detected spikes are marked by red triangles. 262
6.15 Comparison between data captured by the PennBMBI analyzer (black) and the RZ2 Neurophysiology Workstation (red). 263
6.16 Illustration of the developed perception augmentation experiment. (a) shows a rat wearing the developed wireless waterproof neuroprosthetic. The electrodes are chronically implanted in the somatosensory cortex. (b) shows the experimental setup. A rat navigates to a hidden platform using only the perception established from the stimulation. (c) illustrates the estimated rat's swimming traces with/without the simulation guidance. 265
6.17 The block diagram for the watermaze system. (a) shows the computer with graphic interface, camera, and a wireless dongle to the wireless neuroprosthetic device. 266
6.18 The typical closed-loop diagram for the developed perception augmentation experiment. 267
6.19 The circuit schematic of the first generation watermaze stimulator. . . 270
6.20 The 3D construction of the 1st generation of the watermaze stimulator board. (a-1) and (a-2) are the top boards, and (b-1) and (b-2) are the bottom board. 270
6.21 (a) The 3D construction of the 1st generation of the watermaze stimulator board. (b) the photo of the assembled stimulator board. Wires are for electrodes and battery charging. 271
6.22 The circuit schematic of the 2nd generation watermaze stimulator. . . 272
6.23 The 3D construction of the 2nd generation watermaze stimulator board. (a) is the top view and (b) is the bottom view. 272
6.24 The photo of the assembled 2nd generation watermaze stimulator board. Wires are for electrodes and battery charging. The whole device is coated with PDMS for waterproofing. 273

6.25 The circuit schematic of the 3rd generation watermaze stimulator. The tuning voltage V_{DAC} is generated by the microcontroller. 273
6.26 The 3D construction of the 3rd generation of the watermaze stimulator board. (a) is the top view and (b) is the bottom view. 274
6.27 The photo of the assembled 3rd generation watermaze stimulator board. Wires are for electrodes and battery charging. The whole device is coated with silicon for waterproofing. 274
6.28 (a) The equivalent circuit model for the electrode interface. (b) A typical stimulation waveform between the electrodes E_A and E_C. Compliance voltages at the beginning and end of the stimulation phase are measured for estimating the impedance. 275
6.29 The flowchart of the computer program. 277
6.30 The flowchart of the program implemented in the neuroprosthetic device. 280
6.31 Illustration of the timing of the stimulator in the neuroprosthetic. Each red dot is a stimulation pulse train, and the green dot indicates the current time in each subplot. (a) shows the delivered pulses when the stimulation interval is less the time per frame, and (b) shows the delivered pulses when the stimulation interval is larger than the time per frame. The finite state machine will correct the stimulation time interval according to the latest command. 281
6.32 Matlab based Graphic User Interface (GUI). 283
6.33 One frame of the captured video during the experiment with displayed information highlighted. The large green circle shows the submerged platform, the yellow dot shows the start location. The red curve shows the swimming trace. The left top corner shows the current distance to the platform and the total distance the rat has been traveled in this trial. 284
6.34 A rat wearing the wireless waterproof neuroprosthetic device. 286
6.35 The *in-vivo* experimental results. (a) and (b) are webcam captured frames during the experiments. The small yellow and large green circles indicate the start and platform locations, respectively. These were superimposed on the video frame and not visible to the rat. (a) shows the swimming trace of the rat without the simulation, and (b) shows the swimming trace with the simulation guidance. 286

6.36 A typical trial with stimulation. The animal receives a stimulation when it swims away from the platform. The stimulation pulses are marked by red vertical lines in this figure. It clearly shows the animal turned the direction when it received the stimulation. 287
6.37 Illustration of the four possible platform locations $P1 - 4$. In this experiment setup, the platform was randomly placed in one of them. The rat was initially set free at the center of the water tank in each trial. 288
6.38 In a total of 139 trials consisting of 124 stimulation trials and 15 catch trials, the percentage of trials in which the rat reaches the correct platform in its first visit is 65% with stimulation, and 20% without stimulation. 288
6.39 The chance for the rat's first visit is the platform location in the previous trial is 17.5% with stimulation, and 46.2% without stimulation. 289
6.40 The average times of platform visits are 1.8 times with simulation, and 3.7 times without stimulation guidance. The error bars show the standard deviation in the data. 290
6.41 Envisioned bidirectional clinical hand neuroprosthesis. Motor function is restored through brain-controlled electrical stimulation of hand muscles, and somatosensation is restored through sensor controlled electrical stimulation of the brain. The motor pathway replacement has been extensively studied, and this work is focused on the sensation restoration. 292
6.42 Illustration of the BMI device housed in a chamber (not to scale). The electrodes will be chronically implanted with universal nano-connecters through the skin, secured by dental cement. (*electrode type varies with different applications). 293
6.43 The photograph of (a) monkey D with one of the custom designed chamber and (b) monkey M with two of the custom designed chambers. Each chamber has a diameter of 30mm and a height of 40mm. 296
6.44 Photographs of several assembled devices for the chamber. (a) A bi-directional BMI device with micro-SD card, (b) a 32-channel wireless neural recorder, and (c) a 16-channel wireless bi-directional BMI. . . 297
6.45 The block diagram of the basic version of the bi-directional BMI ASIC. 298
6.46 Architecture of the bi-directional, closed-loop brain-machine interface system. The system includes a custom system-on-chip (SoC) and supporting electronics. 299

6.47 (a) Communication interface between the chip and general purpose MCU (not all pads are shown), (b) communication data format. The MCU (master) writes the gray sectors. 301
6.48 A micrograph of the basic version of the bi-directional neural interface SoCs in (a) On-Semi 0.5μm CMOS technology, and (b) IBM 180nm CMOS technology. Major building blocks are highlighted. 302
6.49 The micrographs of the bi-directional neural interface with the proposed energy-efficient neural feature extraction and PID closed-loop controller. (a) shows the first version with 12 channels with debugging and testing structures. (b) shows the second version with 16 channels. Major building blocks are highlighted. 303
6.50 (a) Illustration of the implanted depth electrode arrays. (b) MRI and CT image with the visible hippocampal array. 305
6.51 The power spectrums of the long-term recordings of the monkey D from the three electrode arrays: (a) hippocampus, (b) entorhinal cortex, and (c) medial septum. 306
6.52 The recorded spectrum of the recovery process from anesthesia in (a) monkey D and (b) monkey F. 307
6.53 The stacked plot of in total of 278 stimulation triggered evoked potentials recorded using the developed BMI device. 308
6.54 The average response waveform of the stimulation triggered evoked potentials in the hippocampus. 309
6.55 Recording of the stimulus-evoked potentials. Stimulus trains of 40 Hz and 60 Hz in different brain states are shown. The plots show an triggered average over 3 hours recording in total, with stimulation every 30 seconds. 309
6.56 Time course of EP peaks and pre-stimulus gamma power in across three behavioral states: sedated (red, black circles), recovery (gray circles), and awake (blue circles). Gamma (30-50 Hz) power was calculated in a 300-ms window preceding each stimulus. 310
6.57 Mean EPs during sedation (red, black) and awake (blue) for (a) monkey D and (b) monkey F. 95% confidence intervals on the mean are shown in gray. 311

6.58 (a) The power spectral density of the hippocampal recordings during recovery (gray) and awake (blue) states in monkey F. (b) Distribution of gamma amplitude across sedated (red, black) and awake (blue) recording sessions in monkey F. 312

Chapter 1

Introduction

1.1 Background and Motivation

Since the dawn of human civilization, people have started the attempts to study the brain, with the hope that it will give us answers to the fundamental questions like what makes us who we are, where is the consciousness from. However, even with the science and technology advancements nowadays, many mechanisms of the brain functions remain unclear. There are more than a hundred billion neurons in a human brain [23], approximating the number of stars estimated in our galaxy [24]. Each of the neurons establishes connections with ten thousands of other neurons, forming a massive neural network. Interestingly, the brain, or the neurons represent the information in terms of an electrical signal by the redistribution of the ions with different charges [25]. This gives electrical engineers a unique opportunity to design instrumentations for collecting the neural signal, and also generating electrical signals imitating the neural signal. The direct communication pathway between the brain and

the external world is named brain-machine interface (BMI), brain-computer interface (BCI) or neural interface [26].

The first BMI experiment was conducted by J. Vidal from the University of California, Los Angeles in 1973 [27], for an observation and detection of brain events in electroencephalogram (EEG). The first intracortical BMI was built by P. Kennedy from the Georgia Institute of Technology in 1987 [28]. The first demonstration of controlling a physical object using EEG signal was reported by S. Bozinovski in 1988 [29]. In 1999, researchers led by Yang Dan at the University of California, Berkeley decoded neuronal firings to reproduce images seen by cats [30]. The same year, J. Chapin and colleagues from the MCP Hahnemann School of Medicine and Duke University demonstrates the first direct control of a robotic manipulator by decoding an assembly of cortical neurons [31]. In 2000, Nicolelis and his colleagues from the Duke University developed BCIs that decoded brain activity in monkeys and used the devices to reproduce monkey movements in robotic arms [32]. The same year, Gerwin Schalk from the Wadsworth Center of the New York State Department of Health developed a general-purpose system for BCI research called BCI2000 [33]. In 2012, Leigh R. Hochberg and fellow investigators at the Brown University helped two people with tetraplegia to reach for and grasp objects in three-dimensional space using robotic arms from decoding the motor cortex [34]. In 2016, Sharlene N. Flesher and his colleagues from the University of Pittsburgh helped a paralyzed man experience the sense of touch in his mind-controlled robotic arm by intracortical microstimulation of the somatosensory cortex [35].

These achievements are promising and encouraging for the development of the future generation of BMI systems. However, several bottlenecks still need to be overcome before the technology can be extensively used in experimental and clinical

studies, such as the development of a robust bidirectional neural interface and close-loop operation [36]. The importance and motivation of a bidirectional closed-loop neural interface can be understood from three perspectives: i) the development of the neuroprosthetic device with sensory feedback, ii) the treatment of neural disorders, and iii) the study of neuroscience and neurology. Each perspective is analyzed as follows.

i) Firstly, a bidirectional closed-loop neural interface is important for the development of the neuroprosthetic and BMI device with sensory feedback. Sensations and actions are inextricably linked. Behavioral goals are achieved by sampling the environment with the available sensory modalities and modifying actions accordingly. Somatosensory feedback is especially important to the dexterous hand movement control. Recent developments in hand prosthetics with motor pathway replacement alone do not lead to the adequate use of a paralyzed hand [37]. Artificial sensation restoration is needed for this technology to meet the performance required for clinical adoption. The sensation may be restored with direct electrical microstimulation of the brain [38]. The cuneate nucleus (CN) in the dorsal brainstem carries fine touch and proprioceptive information from the upper body, and is a suitable sensory encoding site. Besides, its compact representations may be reliably activated artificially. Fig. 6.41 illustrates an envisioned bidirectional clinical hand neuroprosthesis with motor function restored through brain-controlled stimulation of hand muscles, and somatosensation restored through sensor controlled electrical stimulation of the brainstem.

ii) Secondly, a bidirectional closed-loop neural interface is important for the treatment of neural disorders. Deep brain stimulation (DBS) is an FDA approved treatment for essential tremor, Parkinson's disease, dystonia, and obsessivecompulsive

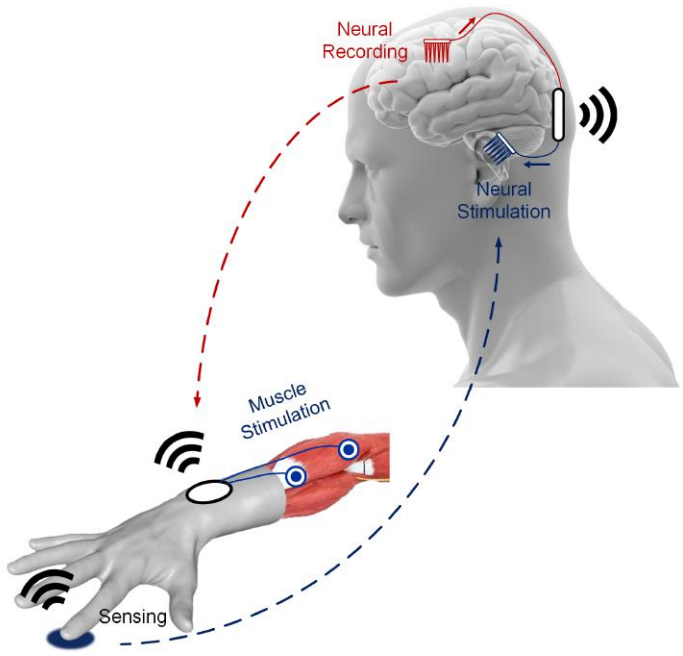


Figure 1.1: Envisioned bidirectional clinical hand neuroprosthesis. Motor function is restored through brain-controlled electrical stimulation of hand muscles, and somatosensation is restored through sensor controlled electrical stimulation of the brain. The motor pathway replacement has been extensively studied, and this work is focused on sensation restoration.

disorder (OCD) [39]. Despite the long history and success in the clinical use of the DBS, the underlying mechanism remains not clear [40]. However, recent research has shown that a closed-loop stimulation can achieve better performance than standard open-loop treatment. In 2011, the research conducted by B. Rosin and his colleagues from the Hebrew University-Hadassah Medical Association School of Medicine shows that the closed-loop stimulation has a greater effect than the conventional open-loop stimulation paradigms, and have the potential to be effective in other brain disorders [41]. A. Bernyi *et al.* from the Rutgers University presented the closed-loop control of epilepsy by transcranial electrical stimulation in 2012 [42]. The seizure-triggered feedback transcranial electrical stimulation can effectively reduce the pathological brain pattern while leaving the other aspects of brain functions unaffected. In 2013, J. Paz

and colleagues from the Stanford University showed that a closed-loop optogenetic control of thalamus can immediately interrupt electrographic and behavioral seizures [43]. All of these successful implementations encourage the implementation of a fully integrated closed-loop bidirectional neural interface for the clinical treatment of neural disorders.

iii) Last but not least, a bidirectional closed-loop neural interface is an essential approach for the study of neuroscience and neurology. The goal of the neuroscience research is to better understand the operational principles of the brain. The brain activity consists of complex interactions of both internal state and external stimuli [44]. This is reflected from the single neuron level to the recurrent neuronal network level, and is important for both *in-vitro* and *in-vivo* studies [45]. Examples of studies using bidirectional BMI include bridging lost biological connection [46], generating synaptic plasticity and strengthening weak synaptic connections [47], reinforcing the activity that generates the stimulation [48].

1.2 Review of Prior Work

This section presents a comprehensive survey and review of the bidirectional brain-machine interface designs published to date. The BMI systems can be categorized from several different perspectives: i) based on the electrodes' location, the BMI systems can be classified into non-invasive systems and invasive systems; ii) based on the signal and control flow, the BMI systems can be classified into one-directional BMIs (recording or stimulation alone), and bidirectional BMIs (both recording and stimulation); iii) based on the study and characterization approaches, the BMI systems can be classified into open-loop BMI and closed-loop BMI. Fig. 1.2 shows the historical trend for the publications of BMI systems over the past 15 years. Specific categories

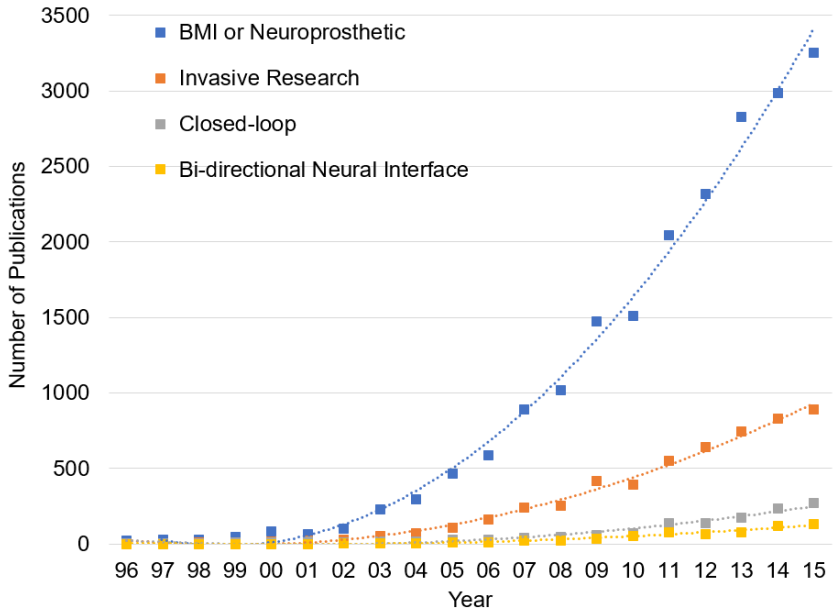


Figure 1.2: The historical trend for publications of BMIs over the past 15 years. Specific categories of BMIs including invasive approach, closed-loop study, and bidirectional interface are plotted for comparison.

of BMIs including the invasive studies, the closed-loop studies, and the bidirectional

neural interface are plotted for comparison. The data was retrieved from the Scopus database [49].

Non-invasive systems mainly exploit the EEG to control external computers and devices, while the invasive systems are based on intracranial recording and stimulation of single or multiple neurons. The non-invasive approach doesn't require surgery for electrode implants and thus has a significant less safety concern. However, the invasive approach gives a more direct interaction with neurons, resulting in advantages in recording resolution and bandwidth, as well as the stimulation effectiveness and accuracy. This work mainly focuses on the study of invasive BMI systems. In addition, most existing neuroscience research and BMI circuits and system development are based on one-directional signal flow and open-loop approach: either neural signal recording or neural stimulation. Fig. 1.3 illustrates the percentage of BMI research papers and designs using bidirectional signal flow and closed-loop approach. Second

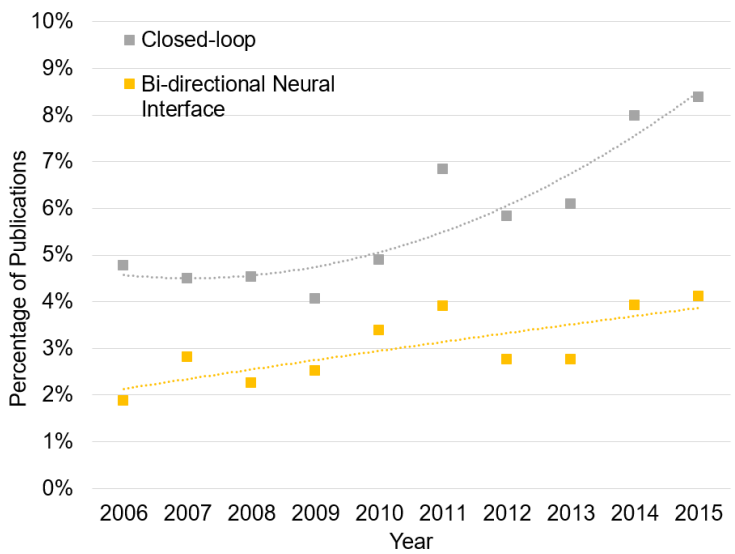


Figure 1.3: The trend for the percentage of closed-loop and bidirectional designs among all BMI publications. Second order polynomial fitting curves are plotted for showing the trend.

order polynomial fitting curves are plotted for showing the trend. There is a clearly increasing trend for the closed-loop approach and bidirectional BMI designs. It is even more convincing given the exponential increase of the overall publications of BMIs. However, the overall percentage is still low. The work in this thesis mainly focuses on the study of bidirectional and closed-loop approach for BMI design.

In order to have a comprehensive understanding of the progress of bidirectional BMI designs, especially from the electrical engineering perspective, a survey of bidirectional BMI designs are given below. Table 1.1 and Table 1.2 together list the bidirectional BMI designs with key design features. The tables include five of the publications from this thesis. The selected features include the recording and stimulation channel counts, neural feature extraction processing, closed-loop operation, wireless communication, target application, and the animal experiments or validation.

Among all bidirectional BMI designs, work that involves high channel count has been reported [56, 64, 69]. It should be noticed that most of the high channel count designs were for in-vitro study, which has less restraint on the power consumption. Neural feature extraction has been performed in a computer [53, 54], in a general-purpose microcontroller [6, 7, 57], or on an ASIC [11, 60, 70, 76]. Commonly used neural features include: neural energy in specific frequency bands [7, 11, 53, 55, 68], action potentials [6, 50, 54, 63], freq-time wavelet domain features [70, 76], entropy [10, 57], and phase synchrony [60]. Commonly used closed-loop methods include: simple trigger [50, 53, 66], linear mapping [63, 67, 79], classifiers [8, 57, 70], PI or PID controllers [11, 68, 73]. The wireless communication modules integrated in the BMI systems include commercial solutions [55, 57, 61] and ASIC designs [10, 11, 79]. The target applications of these papers include generalized neuroscience research [58, 62, 62, 65, 71], in-vitro neuronal study [9, 64, 69], deep brain stimulation (DBS) treatment

Table 1.1: Survey of Bidirectional Neural Interface Designs (Part I)

Ref.	Publication	Affilication	Lead Author	Rec.	Stim.
2005 [50]	J. Neurosci. Methods	Univ. of Washington	J. Mavoori	1ch	1ch
2006 [51]	JSSC	ETH Hoenggerberg	F. Heer	128ch	128ch
2007 [52]	TCAS-I	Georgia Tech.	R. Blum	16ch	16ch
2009 [53]	TBE	UC Berkeley	S. Venkatraman	16ch	16ch
2009 [54]	EMBC	Emory Univ.	J. Rolston	60ch	60ch
2010 [55]	JSSC	Univ. Michigan, Ann Arbor	J. Lee	8ch	64ch
2010 [56]	TBioCAS	Univ. of Toronto	F. Shahrokhi	128ch	128ch
2010 [57]	EMBC	National Cheng Kung Univ.	S. Liang	1ch	2ch
2011 [6]	TNSRE	Univ. of Washington	S. Zanos	3ch	3ch
2011 [7]	JNE	Washington Univ.	A. Rouse	4ch	8ch
2011 [9]	JSSC/TBE	Case Western Reserve Univ.	M. Azin	4ch x2	1ch x2
2011 [58]	TBioCAS	Univ. of Cagliari	D. Loi	8ch	8ch
2012 [8]	TNSRE	Medtronic/MIT	S. Stanslaski	12ch	8ch
2013 [59]	ISCAS	Univ. of Ulm	U. Bihr	1ch	1ch
2013 [60]	JSSC	Univ. of Toronto	K. Abdelhalim	64ch	64ch
2014 [61]	ISCAS	Univ. of Penn	**This work**	4ch	2ch
2014 [10]	JSSC	National Chiao Tung Univ.	W. Chen	8ch	1ch x2
2014 [62]	CICC	Case Western Reserve Univ.	K. Limnuson	1ch	1ch
2014 [63]	Scientific Reports	Italian Institute of Tech.	G. Angotzi	8ch	8ch
2014 [64]	JSSC	ETH Zurich	M. Ballini	1024ch	1024ch
2014 [11]	JSSC	Univ. Michigan, Ann Arbor	H. Rhew	4ch	8ch
2014 [65]	ESSCIRC	Medtronic/Washington Univ.	P. Cong	8ch	32ch
2014 [66]	JNE	Imec	T.K.T. Nguyen	32ch	1ch
2015 [67]	TBioCAS	Univ. of Penn	**This work**	4ch	2ch
2015 [68]	BioCAS	Univ. of Penn	**This work**	12ch	12ch
2015 [62]	AICSP	Case Western Reserve Univ.	K. Limnuson	1ch	1ch
2015 [69]	TBioCAS	Univ. of Toronto	R. Shulyzki	256ch	64ch
2015 [70]	JSSC	Masdar Inst. of Sci. Tech.	M. Altaf	16ch	1ch
2015 [71]	VLSI	Univ. Michigan, Ann Arbor	A. Mendrela	8ch	4ch
2015 [72]	JSSC	UC Berkeley	W. Biederman	8ch	2ch
2016 [73]	TBioCAS	Univ. of Penn	**This work**	16ch	16ch
2016 [74]	Microelectronics J	Seoul National Univ. Sci. Tech.	A. Abdi	1ch	1ch
2016 [75]	JSSC	Univ. Michigan, Ann Arbor	A. Mendrela	8ch	4ch
2016 [76]	VLSI	Cal Tech./ UCLA	M. Shoaran	16ch	1ch
2016 [77]	Sensors	Wuhan Univ.	Y. Su	32 ch	4 ch
2016 [78]	BioCAS	Univ. of Ulm	M. Haas	1ch	1ch
2017 [79]	ISCAS	Univ. of Penn	**This work**	16ch	16ch

Table 1.2: Survey of Bidirectional Neural Interface Designs (Part II)

Ref.	Neural Feature Ex.	Closed-loop	Wireless	Application	Animal Exp.
2005 [50]	Spike	Trigger	No streaming	Generalized	Free behaving
2006 [51]	-	-	-	Generalized	-
2007 [52]	-	-	-	Generalized	-
2009 [53]	Energy(1)	Trigger	-	Generalized	Awake
2009 [54]	Spike(1)	Trigger	-	Generalized	Free behaving
2010 [55]	Energy	Trigger	Off-chip	DBS treatment	Anesthetized
2010 [56]	-	-	-	Generalized	Anesthetized
2010 [57]	Entropy/spectrum(2)	Classifier	Zigbee	Seizure ctrl.	Free behaving
2011 [6]	Spike(2)	Classifier	-	Generalized	Free behaving
2011 [7]	Energy(2)	Classifier		Generalized	-
2011 [9]	Spike	Trigger	-	Neuronal Study	Anesthetized
2011 [58]	-	-	-	Generalized	Anesthetized
2012 [8]	Spectrum	Classifier	-	Generalized	Awake
2013 [59]	-	-	-	Generalized	-
2013 [60]	Phase synchrony	Trigger	UWB	Seizure ctrl.	Anesthetized
2014 [61]	Energy/spike	Trigger	2.4GHz(3)	Generalized	Awake
2014 [10]	Entropy/spectrum	Trigger	OOK	Seizure ctrl.	Awake
2014 [62]	-	-	-	Generalized	-
2014 [63]	Spike(2)	Mapping	2.4GHz(3)	Generalized	Free behaving
2014 [64]	-	-	-	Neuronal study	In Vitro
2014 [11]	Energy	PI ctrl.	Back-scattering	DBS treatment	-
2014 [65]	Spectrum	Unknown	-	Generalized	Awake
2014 [66]	Spike	Trigger	-	Generalized	Awake
2015 [67]	Spike/Energy	Mapping(2)	2.4GHz(3)	Generalized	Free behaving
2015 [68]	Spike/Energy	PID ctrl.	2.4GHz(3)	Generalized	Free behaving
2015 [62]	-	-	-	Generalized	-
2015 [69]	-	-	-	Neuronal Study	Anesthetized
2015 [70]	Freq-time	Classifier	-	Seizure ctrl.	-
2015 [71]	-	Unknown	-	Generalized	-
2015 [72]	Spike	Trigger	-	Generalized	Anesthetized
2016 [73]	Spike/Energy	PID ctrl.	2.4GHz(3)	Generalized	Free behaving
2016 [74]	-	-	-	Generalized	-
2016 [75]	-	-	-	Generalized	Anesthetized
2016 [76]	Freq-time	Trigger	-	Seizure ctrl.	Free behaving
2016 [77]	-	-	2.4GHz(3)	Generalized	Free behaving
2016 [78]	Spectrum	-	-	Generalized	-
2017 [79]	-	Mapping	UWB	Sensory Encoding	Awake

(1) Off-chip, in computer or workstation
(2) Off-chip, in commercial micro-controller
(3) Off-the-shelf electronic solution

[11, 55] especially neural disorder control [57, 60, 70, 76], and sensory encoding study [79]. The animal experiments and validation can be categorized as conducted in: anesthetized animals [55, 56, 60], awake but restrained animals [10, 65, 66], and freely behaving animals [73, 76, 77]. Anesthetized and restrained awake animal experiments can be conducted using wire connected instrumentation, which has much less concern than those experiments conducted in freely behaving animals [73].

In addition to the general survey, a few key designs, which can be considered as milestones in the development of BMIs, are reviewed. These works are from Brown University, Case Western Reserve University, Duke University, National Chiao Tung University, Medtronic Inc., Stanford University, University of California, Berkeley, University of Michigan, Ann Arbor, University of Toronto, University of Washington, and Washington University, St Louis. The selected works are focused on the design from an electrical engineering perspective. The system architecture and circuit implementation of these papers are very helpful in understanding bidirectional BMI designs. The major innovations and contributions are highlighted, while the limitations are also mentioned. The first and corresponding authors with associated labs mentioned in this session have years of experience in BMI development, thus are very valuable resources for tracking the trends of BMI designs.

Stavros Zanos and Eberhard E. Fetz *et al.* from the University of Washington designed an autonomous head-fixed computer (the Neurochip-2) for recording and stimulating in freely behaving monkeys in 2011 [6]. The first generation of the device developed in this group was published in 2008 [50]. The block diagram of the major components and signal routing are shown in Fig. 1.4. The device has three recording and three stimulating channels. Digital filtering and action potential discrimination can be performed in the hardware, and action potential triggered stimulation was

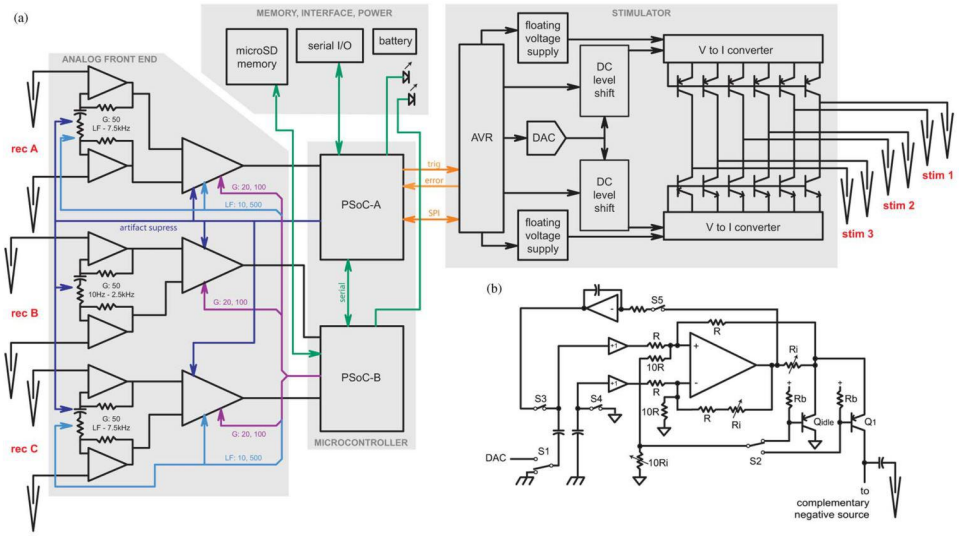


Figure 1.4: The block diagram and signal routing of the Neurochip-2 system developed by S. Zanos and Eberhard E. Fetz *et al.* from the University of Washington in 2011 [6]. Reused with permission, copyright 2011, IEEE.

demonstrated. An accelerometer was integrated into the system. The device has a wireless interface for uploading data and setting device configuration, but real-time data streaming was not supported. An 8MB on-board memory was used to store the recorded data. Though with a limited number of channels, this work is among the early demonstrations of long-term bidirectional recording and stimulation in freely behaving monkeys.

A. G. Rouse and T. J. Denison *et al.* from Washington University, St. Louis and Medtronic Inc. designed a chronic generalized bidirectional brain-machine interface in 2011 [7, 8]. The system incorporated neural recording and processing subsystem into a commercial neural stimulator. The block diagram of the system is shown in Fig. 1.5. The system performs spectral analysis, algorithm processing, and event-based data logging. A three-axis accelerometer was also included in the system. The prototype underwent verification testing to ensure reliability. The system includes

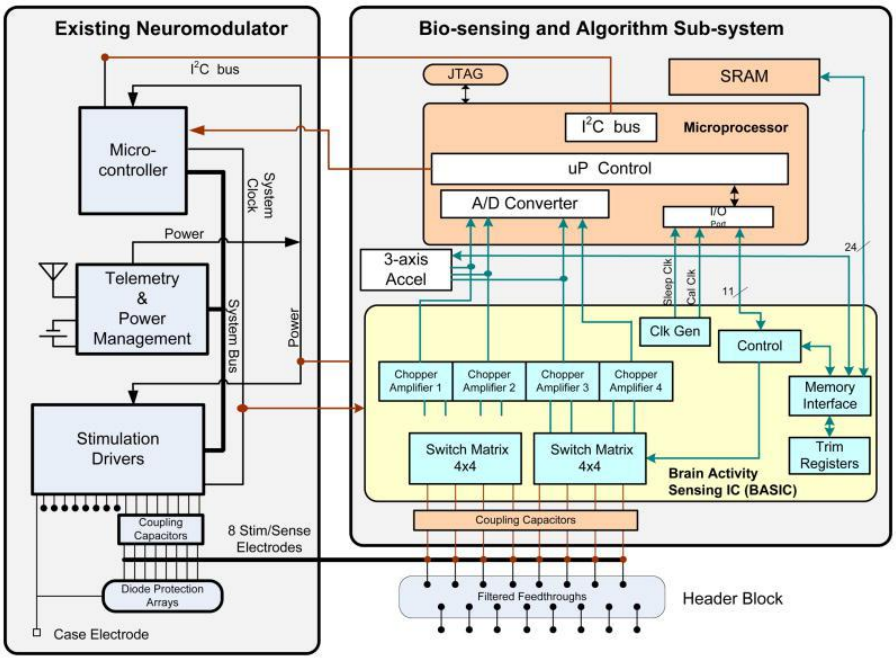


Figure 1.5: Electrical device architecture and system partition for the prototype developed by A. G. Rouse and T. J. Denison *et al.* from Washington University, St. Louis and Medtronic Inc. in 2011 [7, 8]. Reused with permission, copyright 2011, IEEE.

a wireless link for data upload and configuration, but real-time data streaming was not supported. The device integrates an 8MB SRAM for storing data. Concurrent bidirectionality of the device was not tested in the non-human primate model, and the closed-loop experiment was not demonstrated in this work.

Subramaniam Venkatraman and Jose M. Carmena *et al.* from the University of California, Berkeley designed a system for neural recording and closed-loop intracortical microstimulation in awake rodents [53]. This work also demonstrates the first real-time whisker tracking system. The system employed commercial recording and stimulation instrumentation and custom PCB interface board. An on-board circuit was designed to reduce stimulus artifacts. This work performs the signal processing on a computer and doesn't support wireless communication, thus has limited usage in

freely behaving animal experiments. William Biederman *et al.* from the same group proposed a fully-integrated neuromodulation SoC in 2015 [72]. This work consists of 64 acquisition channels and dual stimulation channels. The work also features on-chip digital compression and presents the lowest area and power for the highest integration complexity achieved to date.

Meysam Azin and Pedram Mohseni *et al.* from the Case Western Reserve University designed a battery-powered activity-dependent intracortical microstimulation IC in 2011 [9]. The chip consists of two modules, each module integrates 4 recording

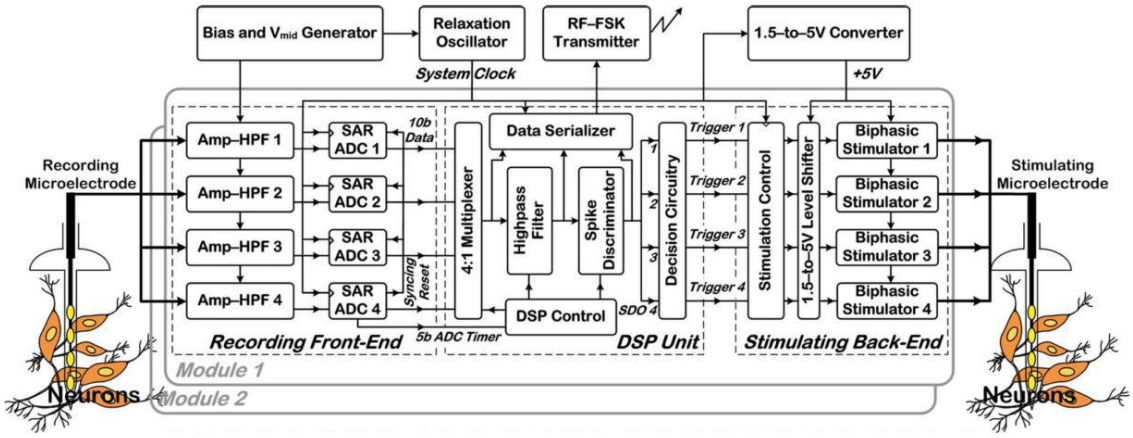


Figure 1.6: The system architecture of the activity-dependent cortical microstimulation (ICMS) chip proposed by M. Azin and Pedram Mohseni *et al.* from the Case Western Reserve University in 2011 [9]. Reused with permission, copyright 2011, IEEE.

channels and 1 stimulating channel. The chip was designed and fabricated in 0.35 μm CMOS technology, powered by a 1.5V battery and provided a stimulation voltage up to 5.05V. This design is among the early demonstration of on-chip action potential discrimination and spike-triggered stimulation. Follow up works from the same group adds an on-chip stimulation artifact rejection feature [62].

Farzaneh Shahrokhi and Roman Genov *et al.* from the University of Toronto designed a 128 channel fully differential digital integrated neural recording and stimulation interface in 2010 [56]. The chip was designed and fabricated in 0.35 μm CMOS technology. The chip doesn't perform feature extraction or closed-loop operation. The same group developed a 320-channel bidirectional neural interface chip in 2015 [69]. The seizure onset detector was implemented off-chip. Wireless communication was not supported in both chips.

Wei-Ming Chen *et al.* from the National Chiao Tung University, Taiwan, designed a fully integrated closed-loop neural prosthetic CMOS SoC for real-time epileptic seizure control in 2014 [10]. The block diagram of the SoC is shown in Fig. 1.7. The SoC consists of 8 recording channels, 1 stimulating channel, digital seizure detection processor, and wireless transceiver. The SoC was fabricated in 0.18 μm CMOS technology. The system and the seizure detection algorithm was verified in Long-Evans rats.

Hyo-Gyuen Rhew and Michael P. Flynn *et al.* from University of Michigan, Ann Arbor designed a fully self-contained logarithmic closed-loop deep brain stimulation SoC in 2014 [11]. The overall system block diagram of the proposed SoC is shown in Fig. 1.8. This work is the first reported implantable SoC with an on-chip closed-loop DBS algorithm. Logarithmic ADC and logarithmic filters were used in this work. A digital PI controller was implemented as the closed-loop controller. This work also integrates an ultra-low-power backscattering wireless transceiver. Adam E. Mendrela *et al.* from the same group developed a bidirectional neural interface circuit with active stimulation artifact cancellation in 2016 [75]. This work also features cross-channel common-mode noise suppression.

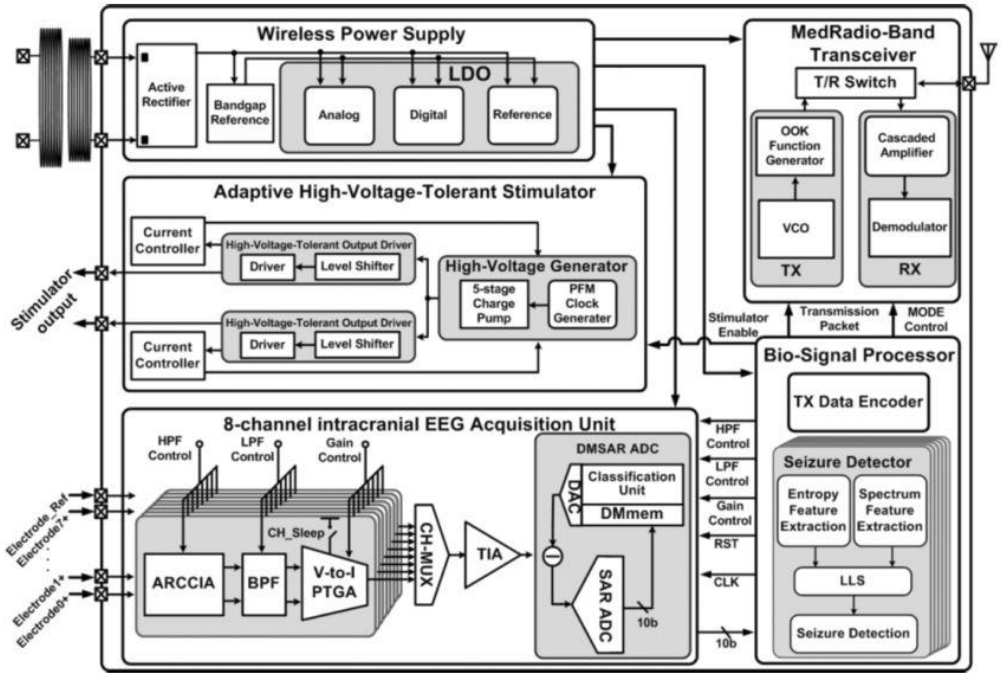


Figure 1.7: The system architecture of the SoC proposed by Wei-Ming Chen *et al.* from Nation Chiao Tung University in 2014 [10]. Reused with permission, copyright 2014, IEEE.

There are also a few publications that describe custom designs for freely behaving animal experiments. These papers address a lot of design challenges of the housing and connection of the electronics. Krishna V. Shenoy *et al.* from Stanford University developed a wireless recording system for freely behaving animals, namely Hermes system, reported in 2007 [80], 2009 [81], 2010 [82], and 2012 [83]. The most recently reported HermesE system features 96-channel full data rate direct neural signal recording (no stimulation). Ming Yin, and Arto V. Nurmikko *et al.* from Brown University developed a wireless neurosensor for full-spectrum electrophysiology recording during free behavior in 2014 [12]. This work supports 96 channel full-spectrum data wireless streaming in short distance. The wireless data rate is up to 200 Mbps. The architecture of the developed neurosensor is shown in Fig. 1.9 (a). David A. Schwarz

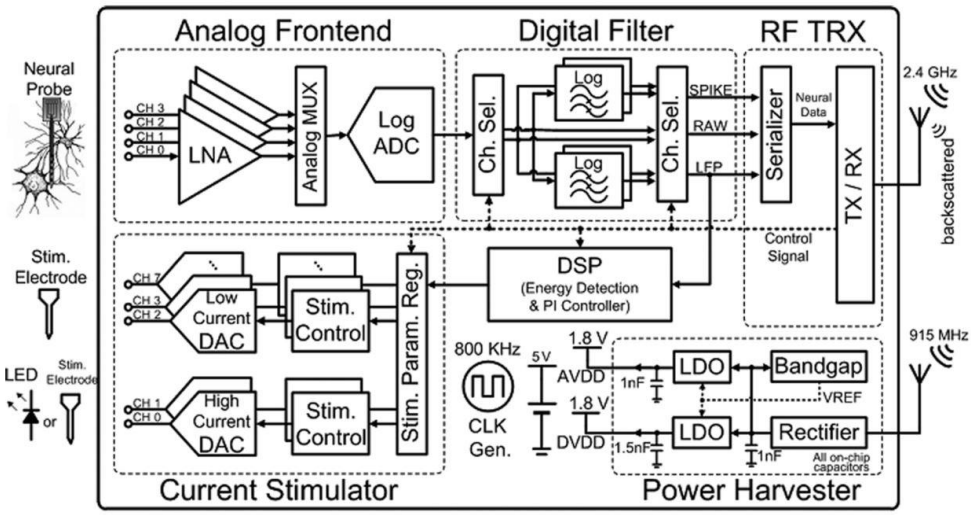


Figure 1.8: The overall system block diagram of the SoC proposed by Hyo-Gyuen Rhew and Michael P. Flynn *et al.* from the University of Michigan, Ann Arbor in 2014 [11]. Reused with permission, copyright 2014, IEEE.

and Miguel A. L. Nicolelis *et al.* from the Duke University developed a chronic wireless recording system for freely behaving monkeys in 2014 [13]. The architecture of the developed device is shown in Fig. 1.9 (b). This work features three-dimensional multielectrode implants and is capable of isolating up to 1,800 neurons from an animal. The design has been validated in several monkeys, and this work reports the highest number of neurons wireless recorded from freely behaving animals to date.

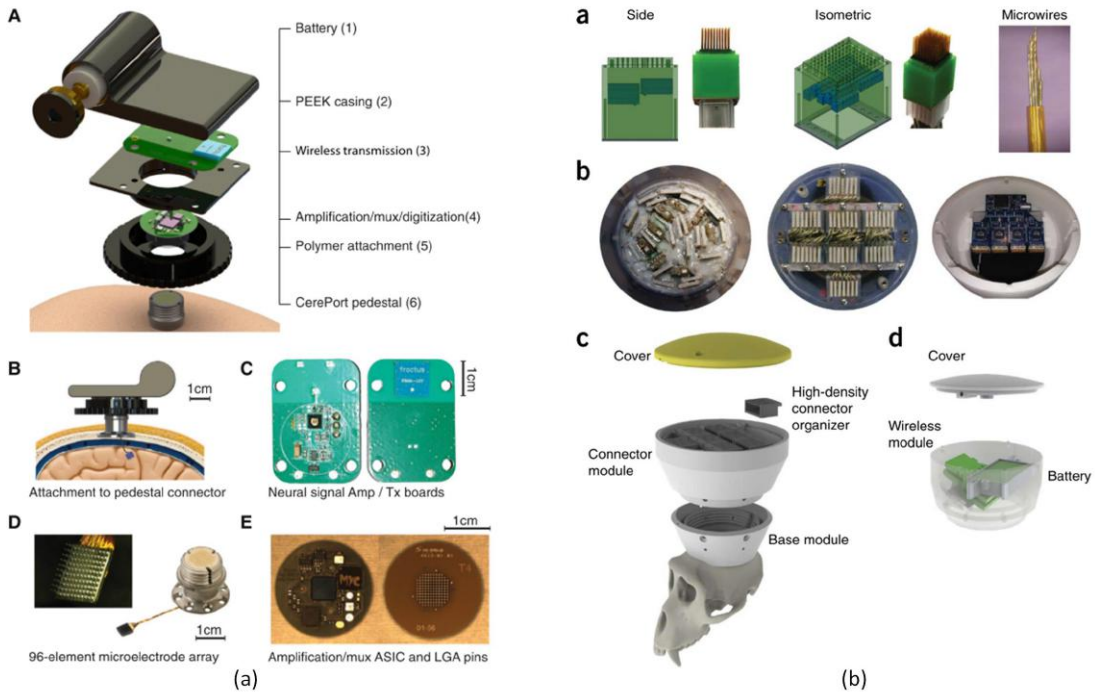


Figure 1.9: The architecture of (a)the wireless neural recording device developed by Ming Yin, and Arto V. Nurmikko *et al.* from the Brown University [12], and (b) the chronic wireless recording system developed by David A. Schwarz and Miguel A. L. Nicolelis *et al.* from the Duke University [13]. Both systems have been validated in animal behaving experiments.

1.3 Overview of Bidirectional Closed-loop Brain Machine Interface

This section gives an overview of the design and implementation of the developed bidirectional BMI system for the closed-loop neuroscience experiments, especially for the use with freely behaving animals. Ideally, the BMI device should be optimized for safety, reliable operation, rich function, small dimension, and long-term operation. To achieve this goal, design optimizations are performed from the neuron-electronics interface up to the system level. The key design requirements are summarized as

follows:

(1) **Safety**: Both implanted electrodes and the stimulation and recording electronics must have a minimal damage to the tissue. This requires the design of the neural interface electronics to have proper input and output impedance, proper stimulation power density given the electrode material and surface area, stimulation charge balancing, and so on. In addition, the design of the packaging and the housing for the electronics and batteries, poses important safety requirements;

(2) **Performance & Reliability**: Both performance and reliability are very important for a BMI device. A reliable performance includes a robust signal recording quality, a reliable wireless data link or data storage, reliable signal processing, feature detection and closed-loop algorithm, reliable electrode connection and electronic assembling, and so on. The dependable and robust performance in recording and stimulation is of great importance for neuroscience investigation and neuroprosthetics development. The reliable on-chip processing is critical for closed-loop operation;

(3) **Interfaces**: The BMI device should provide multiple functional interfaces for neural signal recording, neural stimulation, and various sensing, including body area sensors and supervision. The interface should also include the user interface for the researchers and investigators to use the BMI device for experiments and data analysis;

(4) **Flexibility**: The BMI device should have programmable recording and stimulation configurations including recording gain, bandwidth, channel, sampling rate, and simulation parameters. In addition, the device may also offer programmable

neural feature extraction, machine learning, modulation algorithms to support real-time closed-loop operation;

(5) **Portable**: One of the key requirement for the animal behaving experiment is the ability to record and stimulate wirelessly while the animal is freely moving, including locomotion, social interaction. Conventional rack-mounted BMI instrumentation doesn't support these experiments. The custom BMI device should be lightweight and can be carried by the animals without disturbing their normal behavior;

(6) **Low Power**: A sufficient battery life is important for studying animal behavior in long-term experiments, such as during sleep. It is also important for building plasticity which requires a consistent closed-loop operation. For implantable devices, low power is important to minimize the tissue damage due to the generated heat. Usually, the wireless transceiver and data storage (flash memory programming) consumes most of the power consumption in the BMI device. A stimulation with a high current with a high compliance voltage also consumes high peak power.

The key features and specifications of the proposed BMI device are listed in Table 1.3. The neural recording front-end of the proposed system is designed for invasive recording, including local field potential (LFP) and action potential (AP) signals. The major building blocks of the recording front-end include low-noise neural amplifiers, programmable neural filters, programmable gain amplifiers, and an analog-to-digital converter. On-line data compression, including compressed sensing and spike detection can be used to reduce the wireless data rate. The stimulator back-end is designed for functional electrical stimulation (FES). A high compliance

Table 1.3: Key Features of the Proposed Bidirectional BMI

Analog Front-end		**Stimulator Back-end**	
Channel count	16	Channel count	16
Input referred Noise	<5μV	Stimlus Current	0-255μA/4mA
Bandwith (LFP)	1 - 200Hz	Current Resultion	6 bit
Bandwith (AP)	300 - 10kHz	Pulse width	1 - 255μs
Gain	1000 - 8000	Time interval	8ms - 2s
ADC resolution	10 - 12bit	Compliance voltage	5V
Neural Feature Ex.		**Closed-loop Operation**	
Local field potential	Energy	Feature Extraction	LFP/AP
Action potential	Detection & discrim.	Closed-loop controller	On-chip
Others	Matched filter	Machine Learning	Off-chip
Wireless		**Power**	
Wireless protocol	Bluetooth/FSK	Chip power	<1mW
Wireless datarate	2Mbps	System poower	<30mW
Micro-SD Card	FAT32	Total battery life	>12h

voltage is required for high impedance electrodes. On-chip neural feature extraction for both local field potentials and action potentials are implemented. The closed-loop operation is supported using on-chip PID controller and off-chip general purpose microcontroller. Various body-area sensors are to be used to monitor animal behavior and sensory inputs. The complete battery powered device should able to support continuous operation over 8 hours, or over 12 hours if experiments during animals sleep are required. The data loss over the wireless link should be minimized, and the on-board flash memory would need a file system to support the data storage during the long-term recording. Other features including custom packaging of the device are also important parts of the system design.

The block diagram of the proposed general purpose BMI system is illustrated in Fig. 1.10. The main building blocks of the system include: BMI device, the sensors,

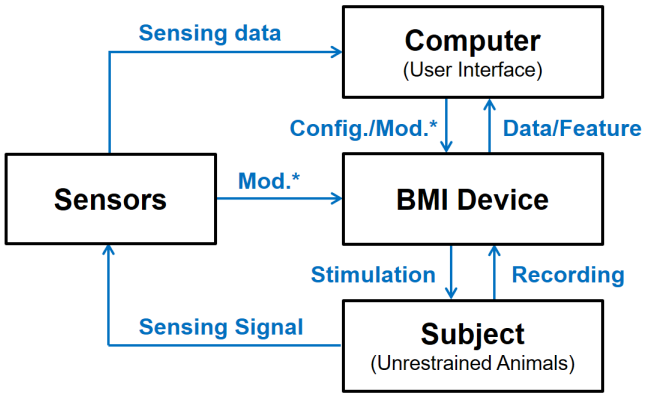


Figure 1.10: The block diagram of the proposed generalized closed-loop, bidirectional BMI system

and the computer with a user interface. The signal flows are marked in the figure. As the core part the BMI system, the BMI device features a bidirectional neural interface and a duplex wireless communication with the computer. The bidirectional neural interface enables both neural signal recording and electrical neural stimulation. The bidirectional wireless communication allows the BMI device to send data back to the computer, and to read commands from the computer or sensor nodes. In addition, the BMI device can process certain neural feature extraction and some pre-defined closed-loop algorithms. The BMI device will be housed in a secure chamber fixed on the animals skull or housed in a customized jacket the animal can wear. Sensors are also important elements in the system. There are two types of sensors: wearable sensors and surveillance sensors. Wearable sensors may include a pressure sensor, flex sensor, accelerometers, goniometer, etc. A sensor node is built using commercial sensors and a wireless transceiver. Surveillance sensors include a video recorder and a motion tracking sensor, which can be designed in CMOS technology. A computer station

provides the user interface for data display, device configuration, and also performs closed-loop algorithms in certain applications. A standard Bluetooth module or a custom designed wireless computer interface board with high-speed USB 2.0 can be used with any computer to control the BMI system. The graphic user interface is designed based on MATLAB. Complex closed-loop modulation algorithms or online data processing can be performed in MATLAB with a certain processing latency. The wireless communication between all blocks uses a customized command and parameter protocol to avoid on-the-air conflict and reduce the cost of system upgrading.

It should be noticed that not all the components are necessary for an experiment. The system can be configured to work in various closed-loop operating modes. Fig. 1.11 shows four commonly used configurations. For examples, if the sensor is

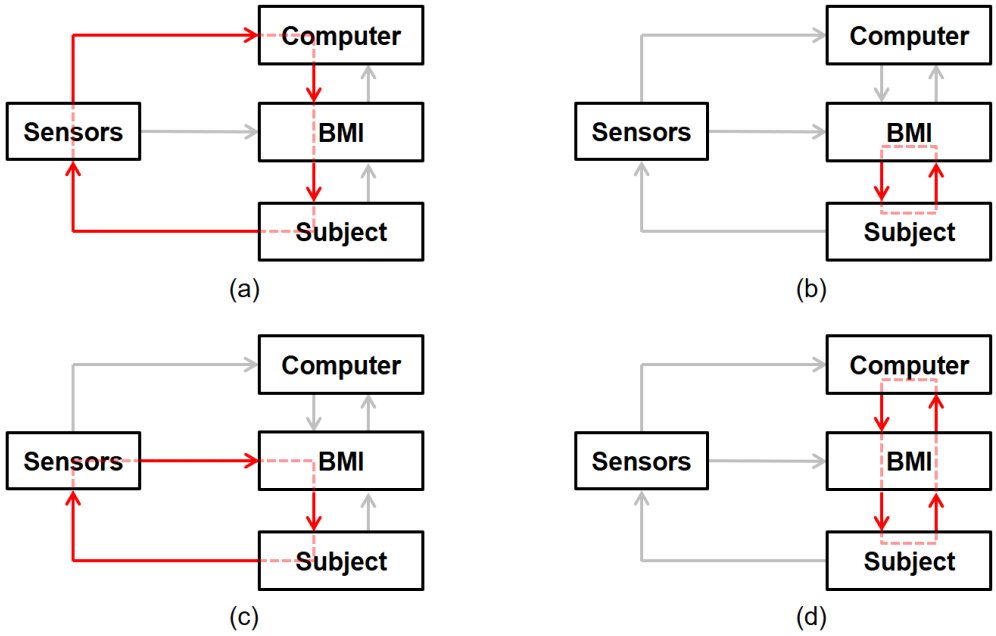


Figure 1.11: Various closed-loop configurations of the proposed BMI system.

a camera, the Fig. 1.11 (a) shows the operation of the watermaze sensory augmentation experiment presented in Chapter 6.3. The camera tracks the animal, sends the location information to the computer, while the computer performs the mapping, then sends the stimulation command to the wireless neuroprosthetic. Fig. 1.11 (b) shows the operation of the bidirectional recording and stimulation experiments in freely behaving monkeys described in Chapter 6.4, and the same BMI system can also be configured to perform the operation as in Fig. 1.11 (c) and (d).

1.4 Thesis Outline and Contributions

The thesis presents the design and analysis of a bidirectional closed-loop BMI system, with emphasis on the use for experiments in freely behaving animals. Fig. 1.12 highlights the major blocks in a typical bidirectional BMI system. The corresponding chapters in this thesis are marked in the figure.

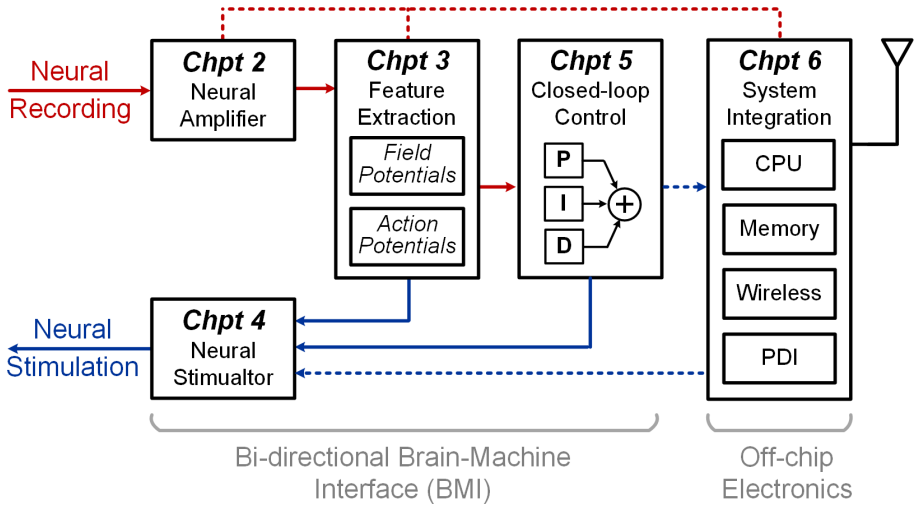


Figure 1.12: The building blocks of a typical bidirectional closed-loop BMI system and the organization of this thesis.

The thesis is organized as follows. Chapter 2 presents the analysis and design of the neural recording front-end. The prior work is reviewed including an analysis of key trade-offs. A general-purpose low-noise amplifier and low-power SAR ADC are designed, followed by a novel pre-whitening neural front-end design for a boosted dynamic range. A neural signal acquisition system that features compressive sensing has been developed for long-term recording in freely behaving animal experiments.

Chapter 3 discusses the neural feature extraction from three perspectives: the energy extraction, the action potential detection, and the matched filtering. Several novel circuits and algorithms have been proposed to improve the performance and power efficiency of the commonly used neural feature extraction. The proposed matched filter in a combination of the pre-whitening neural front-end improves the neural feature detection accuracy.

Chapter 4 presents the analysis and design of a high efficiency electrical neural stimulator. The background of neural stimulation and physicochemical properties of the electrode-electrolyte interface are reviewed, and the key design requirements are summarized. Next, a general purpose neural stimulator design is presented, followed by a novel net-zero charge stimulation mechanism.

Chapter 5 addresses the design issues with the bidirectional neural interface from two perspectives: the stimulation artifacts and the closed-loop operation mechanism. A study of stimulation artifacts in different BMI configuration is presented with *in-vitro* and *in-vivo* experimental results. The mechanisms of different closed-loop neural interface system are summarized, and a commonly used PID controller is designed and tested.

Chapter 6 presents the BMI system integration with a focus on the experiments with freely behaving animals. A general purpose experimental platform, namely the PennBMBI, was presented, featuring wireless recording, stimulation, and sensing ability, with a custom communication protocol and user-friendly computer interface. A watermaze experiment is designed and conducted for the study of augmenting

perception through modulated electrical stimulation of somatosensory cortex. A waterproofed wireless neural stimulator and a complete animal tracking and neuromodulation experimental system are presented. Finally, custom BMI devices are developed and used in long-term bidirectional experiments in freely behaving monkeys. A study in the hippocampal gamma-slow oscillation coupling using the developed system was presented.

To the best of my knowledge, this thesis presents the *first* comprehensive study of bidirectional BMI design for freely behaving animal experiments featuring closed-loop operation. The thesis gives the *first* and most complete survey of the bidirectional BMI designs published to date. The survey gives insights to the progress in terms of key features for BMI systems with emphasis on the on-chip feature extraction, closed-loop operation, and validation in animal experiments; The key contributions of the thesis are summarized as follows:

(1) **System Level**: A complete wireless bidirectional BMI system capable of on-chip neural feature extraction and closed-loop operation has been developed. This work is the *first* reported portable system to provide all necessary hardware for a closed-loop sensorimotor neural interface. The thesis also gives the *first* comprehensive study of stimulation artifacts in different BMI configurations, which is a critical issue in bidirectional BMI design. In addition, the thesis presents the *first* review and summary of the mechanisms for closed-loop BMI operation.

(2) **Circuit Level**: Novel circuits have been proposed in this work to improve the state-of-the-art designs. Several innovative designs are highlighted here: a pre-whitening recording front-end was proposed to improve the dynamic range of

the recording front-end; a natural logarithmic domain neural energy extraction unit was designed to improve the efficiency; a matched filter was proposed to be used in combination with the proposed pre-whitening front-end to improve the neural feature detection accuracy; a novel net-zero-charge neural stimulator was designed for safety and power efficiency. Moreover, custom circuits were developed and optimized in support the system integration and closed-loop operation.

(3) **Application & Experiment Level**: Research and investigation have been conducted using the developed bidirectional BMI system. Novel animal experiment paradigms and methods were proposed and implemented. The presented watermaze experiment is the *first* wireless sensory encoding experiment conducted in freely swimming rat. Bidirectional neuroscience experiments were conducted in macaques using the developed device, including the *first* study to directly compare the hippocampal field potentials in sleep to sedation.

In summary, the methods, circuit techniques, system architecture, and experimental paradigms proposed in this work can be used in a wide range of neurophysiology research and neuroprosthetics development, especially experiments in freely behaving animals.

Chapter 2

Neural Recording Front-end Design

2.1 Introduction

Neural signal recording revolutionizes our understanding of the human brain. Since the first extracellular recording pioneered by the investigators Ward and Thomas in the 1950s [84], neural recording has revealed the fundamental structure and organization of the brain. The number of simultaneously recorded neurons doubled approximately every seven years [85]. The exponential growth in the recording ability is to a large extent driven by the innovations in CMOS technology, circuits and systems design, microelectrode fabrication, and bio-compatible packaging techniques.

The large-scale neural recording also provides a unique opportunity for the research in brain-machine interface (BMI), which builds the interface between the brain

and the artificial devices [84, 86]. However, recent studies estimate that a simultaneous invasive recording of 100,000 neurons is needed for decoding the full-body movements [13], which is beyond the recording ability of the cutting-edge BMI devices. At the same time, the multi-channel recording from freely behaving animals in a natural environment is important for both neuroscience and neuroprosthetic research. However, most of the research to-date still relies on rack-mount instrumentation with restrained cables. The requirement of recording high bandwidth neural signal from multi-channel, in multi-brain areas, via wireless miniature devices places a significant challenge on existing electronic technology and design techniques. The design optimization of a fully integrated neural recording front-end is thus highly desirable.

In the last two decades, a large number of neural recording front-end designs have been reported with improvements from many different aspects [3, 4, 84, 87]. The major innovations have come from the novel circuit and system topologies [16, 88, 89], low-noise design techniques [90–93], large number of the channel-count designs [90, 94–96], energy efficient designs [88, 90, 96], and wireless interfaces including ISM band FSK [90, 97, 98], FM [99, 100], UWB [82, 95], and backscattering [96, 101, 102]. In addition, several systems have been used in freely behaving animal experiments [12, 13], and some of the prototype devices are fully integrated and potentially implantable [94, 96, 103].

This chapter presents the design and analysis of the neural recording front-end. Several novel circuit designs are proposed to improve the state-of-the-art. The chapter is organized as follows. Section 2.1 introduces the characteristics of the neural signal. The design specifications of neural recording circuit and systems are summarized. Section 2.2 reviews the prior work and analyze the key trade-offs in the neural amplifier design, followed by a prototype design of a general-purpose low-noise neural

amplifier. Section 2.3 proposes a novel pre-whitening neural amplifier design, which exploits the frequency characteristics of the neural signal to relax the dynamic range and linearity requirement of the recording front-end. Section 2.4 presents the design of a 10-bit low-power SAR ADC for neural signal digitization. Section 2.5 presents the design of a complete neural signal acquisition front-end with compressive sensing for long-term neural signal recording in freely behaving animals.

2.1.1 Signal Characteristics

The neural signal is recorded via invasive or noninvasive electrodes. Fig. 2.1 shows the most common types of neural signals and the corresponding electrode placement. The

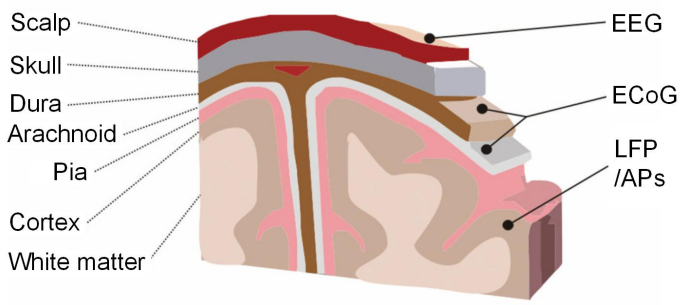


Figure 2.1: The sources of the neural signals and their locations relative to the brain, modified from [14].

electroencephalography (EEG) is the electrical brain activity recorded from the scalp; the electrocorticography (ECoG) is the electrical brain activity recorded beneath the skull; local field potential (LFP) and action potential (AP) are electrical signals recorded within the parenchyma. The AP is the individual neuron activity, and the LFP is the summed activities from multiple nearby neurons.

Fig. 2.2 shows the amplitude and frequency features of different types of neural signal. Main noise sources are also marked in this figure, including the thermal

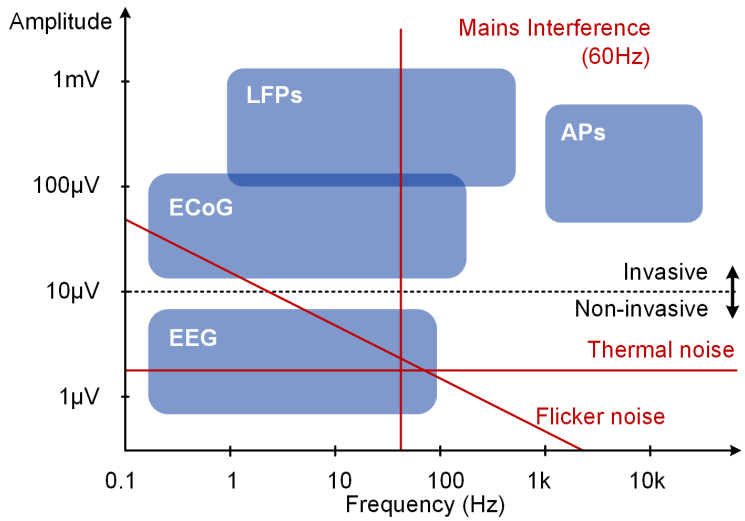


Figure 2.2: The amplitude and frequency characteristics of the neural signal, in comparison with main noise sources. This figured is modified from [15].

and flicker noise from the electrodes and electronic recording device, and the mains interference.

2.1.2 Design Specifications

The key requirements for a neural recording front-end include: low input-referred noise, sufficient dynamic range, high input impedance, high linearity, high common-mode rejection ratio (CMRR) and power-supply rejection ratio (PSRR). The minimum requirements are summarized and listed in Table 2.5, cited from International Electro-Technical Commission (IEC) medical electrical equipment standard 60601-2-47. It should be noticed that the requirement for a specific application will usually be higher than the general standard.

Table 2.1: Summary of Specifications for Neural Recording [1]

Requirement	**Range**	**Unit**
Input Dynamic Range	10	mVpp
Electrode Offset	± 300	mV
Input Impedance	>10	Mohm
Common Mode Rejection	60, at 50-60Hz 30, at 100-120Hz	dB
Gain Accuracy	Error <10% and <±10	μV
Gain Stability over 24h	<3	%
Noise	50	μV
Crosstalk	<0.2 <5	mV %
Timing Accuracy	<30 (over 24 hour)	Sec
Temporal alignment	Error <20	ms

In addition to the requirement in the neural front-end, there are several key requirements of a successful chronic invasive neural recording system:

i) Longevity requirement: safe electrode interface, minimum tissue damage, and infection;

ii) Noise, bandwidth, and channel count requirement for the target signal;

iii) Sufficient battery life to support long-term recording;

iv) Reliable data storage or wireless transmission. In addition, the research of BMI usually requires the front-end to be highly programmable, wireless compatible with commercial equipment and sensors, and also easy to upgrade. All of these features together are required for a practical recording system for neuroscience research and BMI development. A balance between the requirements of each system block needs to be carefully considered.

There are several figure-of-merits (FoM) commonly used for evaluating and comparing different neural recording front-ends. The most important FoM for noise and power performance is the noise efficiency factor (NEF). The NEF was first proposed

by M. Steyaert *et al.* from the Katholieke Universiteit Leuven in 1987 [104], and was resurrected by R. Harrison *et al.* from the University of Utah in 2003 [16].

$$NEF = \overline{V_{ni,rms}}\sqrt{\frac{2I_{tot}}{\pi\Phi_t \cdot 4kT \cdot BW}} \tag{2.1}$$

where $V_{ni,rms}$ is the input-referred rms noise voltage of the amplifier, I_{tot} is the total amplifier supply current. Because for bipolar device, the input-referred rms noise is:

$$\overline{V_{ni,rms}} = \sqrt{\frac{4kT\Phi_t}{I_{tot}} \cdot \frac{\pi BW}{2}} \tag{2.2}$$

So the NEF of a single bipolar transistor is 1 (the lower the better) [16]. Paper [16, 104] predicted that all practical circuits must have a NEF greater than 1, however, later developed techniques overcame this limitation [10, 105].

It should be noticed that the NEF leaves the supply voltage out of the trade-off. So two amplifier designs with different supply voltages but same supply current will have the same NEF. To mitigate this tissue, R. Muller *et al.* from the University of California, Berkeley proposed a power efficiency factor (PEF) in 2012 [106]. The PEF is defined as:

$$\begin{aligned} PEF &= NEF^2 V_{DD} \\ &= \frac{V_{ni,rms}^2 \cdot P_{tot}}{\pi \cdot kT/q \cdot 4kT \cdot BW} \end{aligned} \tag{2.3}$$

The PEF gives a direct trade-off between power and noise, and two amplifiers with the same input rms noise and power consumption should have the same PEF. Reducing supply voltage significantly reduce the power consumption and optimized the PEF, however, at the cost of lowering the dynamic range. In order to compare the overall

system efficiency, D. Han *et al.* from the Nanyang Technological University further proposed a system efficiency factor (SEF) in 2013 [107]. The SEF is defined as:

$$SEF = \frac{PEF}{DR_{out}} \tag{2.4}$$

where

$$DR_{out} = 10log\frac{V_{amp,max}^2}{2 \cdot G_{AFE}^2 V_{ni,rms}^2} \tag{2.5}$$

where $V_{amp,max}$ is the maximum voltage swing of the amplifier, and G_{AFE} is the voltage gain of the amplifier. SEF takes the noise, power, and dynamic range performance into account, thus is more suitable for system level comparison.

2.2 Design of a Low-Noise Neural Amplifier

2.2.1 Review of Prior Work

There have been numerous designs of neural amplifier reported. The motivation of this section is not to give a comprehensive survey of the prior work, but to analyze the key design trade-offs with featured examples. Review, tutorial, and comprehensive surveys for neural amplifier designs can be found in [3, 4, 84, 87].

A. System Topology

The classical instrumentation amplifier uses a 3-opamp topology. The 3-opamp instrumentation amplifier has a high input impedance, a good CMRR, but at a low power efficiency. The commonly used low-power CMOS neural amplifiers use capacitor and resistor elements to set the closed-loop gain. Typical block diagrams are shown in Fig. 2.3. Using a capacitive gain element, the design is inherently AC

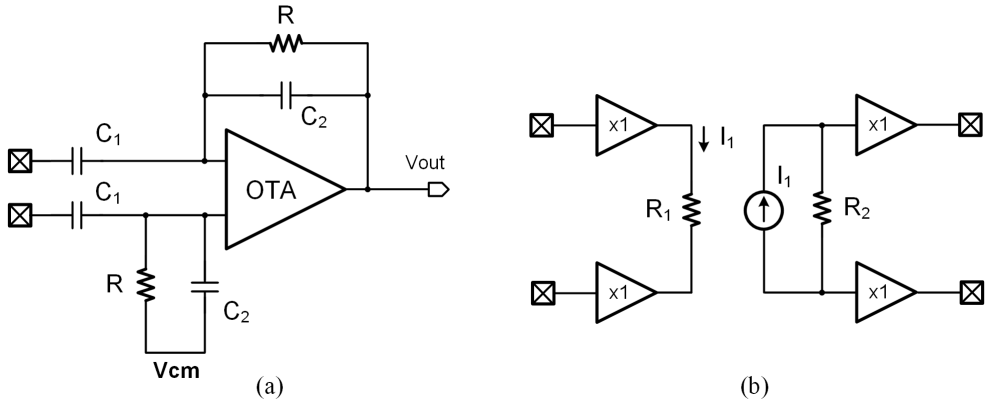


Figure 2.3: The block diagram of the typical neural amplifiers with (a) capacitive gain element [16] and (b) resistive gain element [17]. The x1 symbol is for a unity gain buffer.

coupled 2.3 (a). Thus the input common-mode range is from ground to VDD, limited by the ESD protection circuits. The only active component is the operational transconductance amplifier (OTA). The CMRR is mainly limited by the mismatch of the capacitors. The input impedance is limited by the size of the input capacitor and is frequency dependent. Using a resistive gain element, the design is inherently DC coupled 2.3 (b). The input common-mode range is less than VDD. The DC headroom is VDD/Gain, which is very small. The input impedance is limited by the parasitic capacitance, so it is much higher than the capacitive counterpart. The CMRR is mainly limited by the mismatch of the analog buffers. Unlike the capacitive gain elements, the resistors also contribute to the overall noise. Several key features of the capacitive and resistive amplifiers are listed in Table 2.2. In summary, the capacitive gain element topology enjoys several inherent advantages over the resistive counterpart. Although a lot of techniques have been reported to successfully address these problems [2, 17, 88], the capacitive gain element topologies are the mainstream designs for neural amplifiers.

Table 2.2: Comparison of capacitive and resistive gain elements [2]

Gain Element	Capacitive	Resistive
Gain	C_1/C_2	R_2/R_1
Noise PSD	$v^2_{Thermal}$,$v^2_{Flicker}$	$v^2_{Thermal}$,$v^2_{Flicker}$,v^2_R
Input Impedance	$1/j\omega C_1$	$1/j\omega C_p$
DC headroom	VDD	VDD/Gain
Input CM Range	VDD	<VDD

Besides using capacitive and resistive gain elements, other active feedback topologies can also be used to shape the frequency response [106, 108]. With active feedback,

large input capacitor can be replaced by a small integrating capacitor, and high input impedance can be achieved. However, the active feedback adds to power consumption and contributes to the overall noise of the system.

In addition to topologies using passive or active components to set the closed-loop gain, open-loop amplifiers have also been reported in literature [109, 110]. Compared with the closed-loop topologies, open-loop amplifiers have higher power efficiency, but usually suffer from poor linearity. But since the neural signal has a small amplitude, it may remain in the linear range of the amplifier.

B. Low-Noise OTA

As the core of a low-noise neural amplifier, the OTA design is reviewed in this section. Fig. 2.4 shows the most commonly used low-noise OTA structures. Single-ended topologies are used for illustration. The current mirror OTA (Fig. 2.4 (a)) has two stages with the dominant pole located at the second stage. No compensation capacitor is required to maintain stability. The detailed noise analysis is presented in [16]. However, the current mirror OTA has limited gain, and there exists a trade-off between the noise and the phase margin. A gain boosting circuit can be used to enhance the gain of this amplifier. The two-stage OTA with Miller compensation capacitor (Fig. 2.4 (b)) is also widely used as a low noise amplifier [111, 112]. The folded cascode OTA (Fig. 2.4 (c)) can achieve a high gain in a single stage, at the price of a higher power consumption. Paper [91] describes the detailed analysis in choosing the parameters in a folded cascode OTA design for better power-noise efficiency. The telescopic OTA (Fig. 2.4 (d)) can achieve the highest gain in a single-stage, but with limited input range and voltage swing.

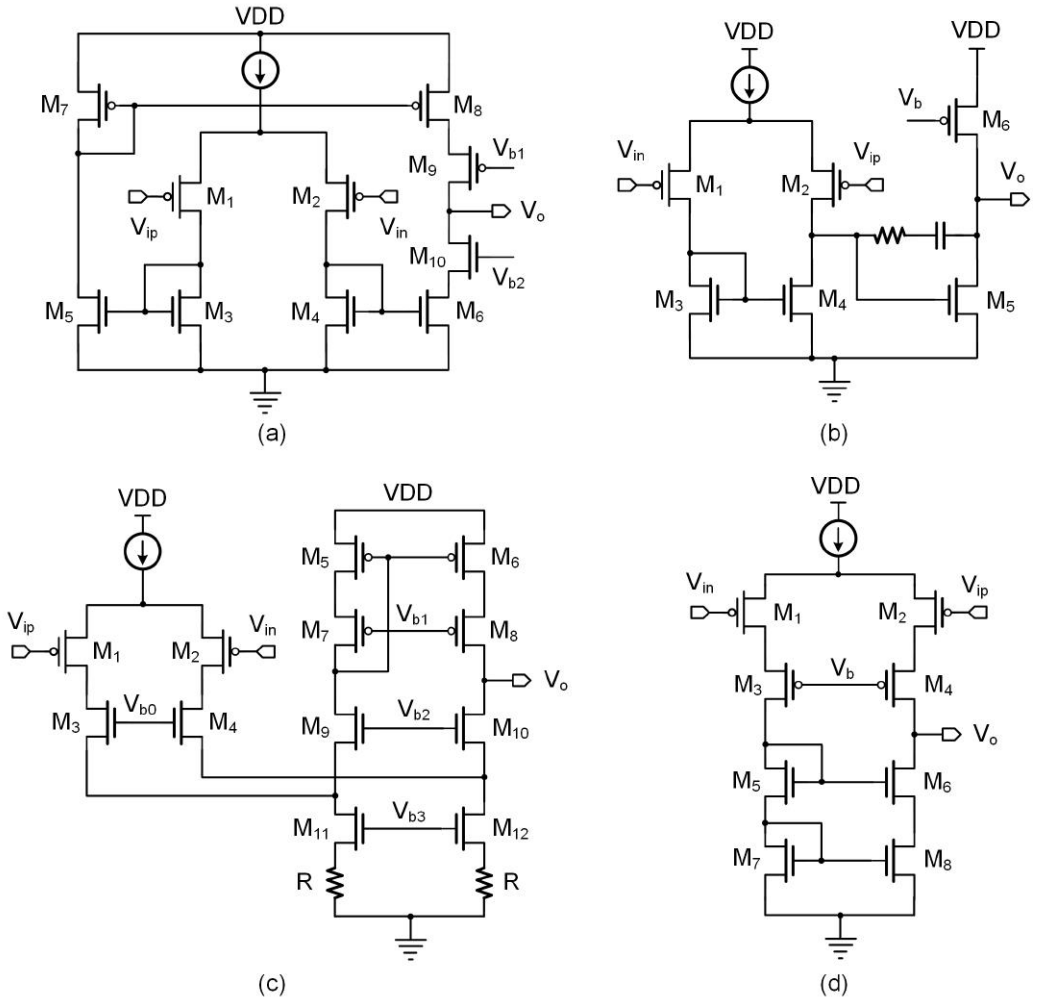


Figure 2.4: The circuit schematic of commonly used low noise OTAs: (a) the current mirror OTA, (b) the Miller OTA, (c) the folded cascode OTA, and (d) the telescopic OTA.

In summary, all of the designs have pros and cons, and differently adapted versions have been widely reported to achieve improved power-noise efficiency. The voltage gain and input-referred noise of these OTA topologies are summarized in Table 2.3. For the thermal noise, increasing the transconductance of the input devices is critical in lowering the noise. Thus maximizing the transconductance for a given

supply current is important for better power-noise efficiency. Besides, supply current can be programmed for achieving the optimal power efficiency in different noise conditions [10].

Table 2.3: Summary of commonly used low-noise OTA [3, 4]

OTA topology	Voltage gain	Input-referred thermal noise
Current mirror 2.4(a)	$g_{m1}(g_{m9}r_{o9}r_{o8}\|\|g_{m10}r_{o10}r_{o6})$	$\frac{16kT}{3g_{m1}^2}(g_{m1}+2g_{m3}+g_{m7})$
Miller 2.4(b)	$g_{m1}(r_{o2}\|\|r_{o4})\cdot g_{m7}r_{o7}$	$\frac{16kT}{3g_{m1}}(g_{m1}+g_{m3})$
Folded cascode 2.4(c)	$g_{m1}\alpha[(g_{m10}r_{o10}\cdot g_{m12}r_{ro12}R)$ $\|\|(g_{m8}r_{o8}r_{o6})]$[1]	$\frac{16kT}{3g_{m1}}(g_{m1}+\frac{2}{R}+g_{m6})$
Telescopic 2.4(d)	$g_{m1}(g_{m4}r_{o3}r_{o4}\|\|g_{m6}r_{o6}r_{o8})$	$\frac{16kT}{3g_{m1}}(g_{m1}+g_{m8})$

C. Other Noise Reducing Techniques

Many circuit techniques have been proposed in the literature to further reduce the noise in the amplifier circuits. Commonly used low-noise techniques include chopping [92, 93, 113], auto-zeroing [114], digital assisted trimming [10, 115], analog or digital filtering, and so on.

For example, chopping is a very popular technique among neural amplifier designs, especially for EEG recordings. Fig. 2.5 illustrates the concept of chopping. Before the amplification, the input signal is modulated by a chopping frequency f_{chop}, which is much higher than the signal frequency. The modulated signal is then located at a frequency higher than the filter noise. After the amplification, the signal is converted back to the baseband frequency, and the flicker noise is up-converted to the

[1]The parameter α depends on the biasing currents' ratio, as calculated in [91].

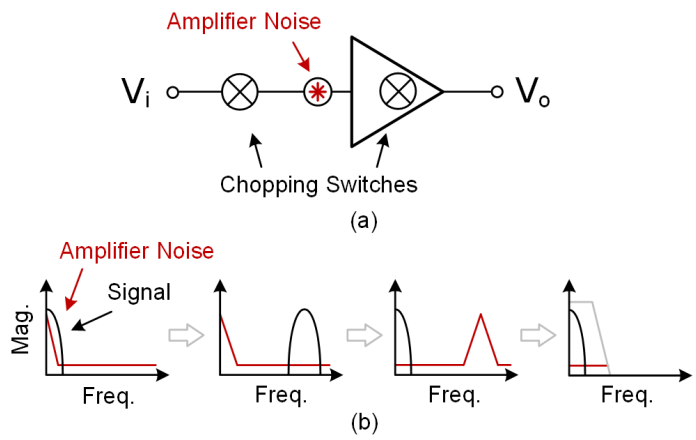


Figure 2.5: (a) The block diagram of a chopping amplifier. (b) Illustration of the signal and noise spectrum before and after the chopping.

chopping frequency, which can be further removed by a lowpass filter. The chopping technique reduces both flicker noise and the DC offset, and the circuits after the chopper switches can achieve an excellent CMRR. It should be noticed that the chopping also causes extra non-idealities, including offset, ripple, charge injection, clock feed-through, switch noise, and so on. Many techniques have been proposed to suppress the problems, including chopping within the feedback loop [92], chopping at the virtual ground [113], ripple reduction [93], input impedance compensation [116], offset cancellation [96, 117], and so on.

2.2.2 Circuit Implementation

The design and analysis of the circuit implementation of the neural amplifier are presented in this section. Fig. 2.6 shows the high-level block diagram of the neural recording front-end. The neural recording front-end includes a low-noise neural amplifier, a programmable gain amplifier (PGA), a multiplexer, an ADC, and a control module. The neural amplifier uses a fully-differential, capacitor feedback topology.

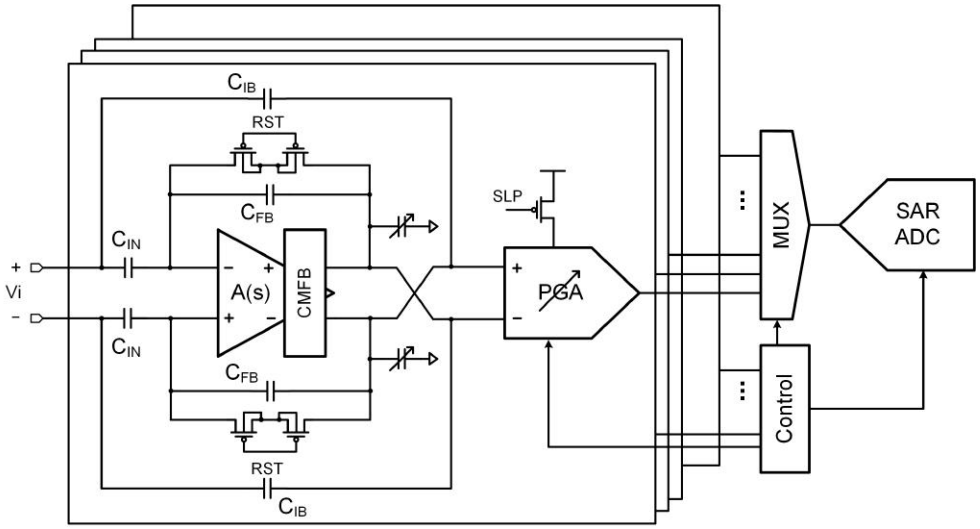


Figure 2.6: The block diagram of the neural recording front-end.

The input capacitors block the electrode offset and the half-cell potential from the electrode-tissue interface. The closed-loop differential gain is set by $C_{IN}/C_{FB}$to be 40dB to relieve the noise requirement for the following stages. A large MOS pseudo-resistor is used to set the low-frequency cut-off. The circuit schematic of the pseudo-resistor is shown in Fig. 2.7. Compared with the MOS-bipolar resistor implemented

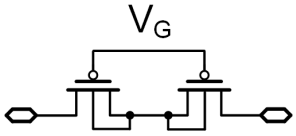


Figure 2.7: The circuit schematic of the MOS pseudo-resistor.

in [16], this resistor has a higher linear range. Besides, setting the gate voltage to ground can short the feedback loop and force the input gate to mid-supply. A simulation of the MOS resistor in IBM 180nm CMOS technology is shown in Fig. 2.8. The W/L of the MOS used in this simulation is 2μm/2μm. The simulated impedance is in the order of 100GΩ. The cut-off frequency is usually set to be much lower than the

signal requirement to prevent the resistor noise from rolling into the signal frequency band. A tunable highpass or bandpass filter can be implemented in the following stage.

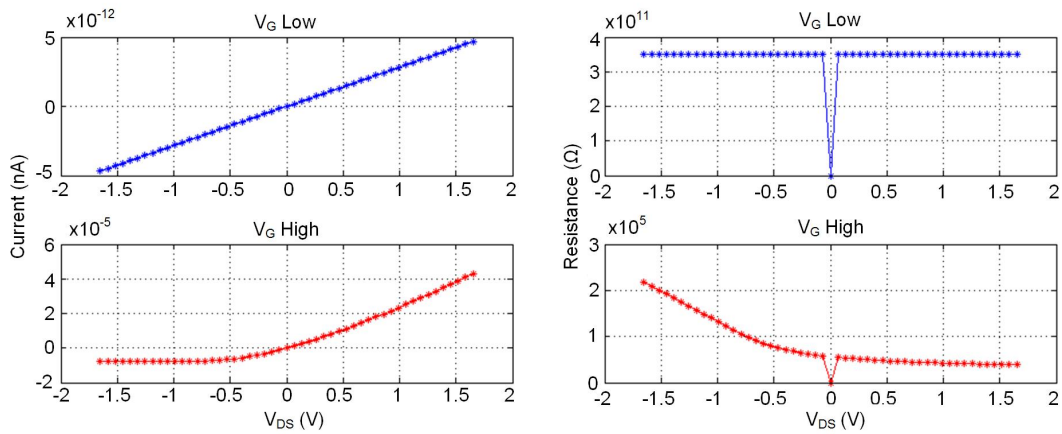


Figure 2.8: Simulation of the resistance of the MOS pseudo-resistor.

The circuit schematic of the OTA is shown in Fig. 2.9. The OTA has been designed to maximize the noise and power efficiency. A single-stage amplifier with a high gain is used to avoid the stability compensation in two-stage structures. The transistors' parameters are listed in Table 2.4.

Table 2.4: The Transistors' Parameters for the OTA Design

Device	$M_{1,2}$	$M_{3,4}$	$M_{5,6}$	$M_{7,8}$
W/L (μm)	20/1	20/1	20/0.8	20/0.8
Multiplicity	30	30	1	1
Finger	4	4	4	4

The overall gain of the amplifier is given by:

$$A_v = (g_{m1} + g_{m3})(g_{m5}r_{o5}r_{o1}||g_{m7}r_{o7}r_{o3}) \tag{2.6}$$

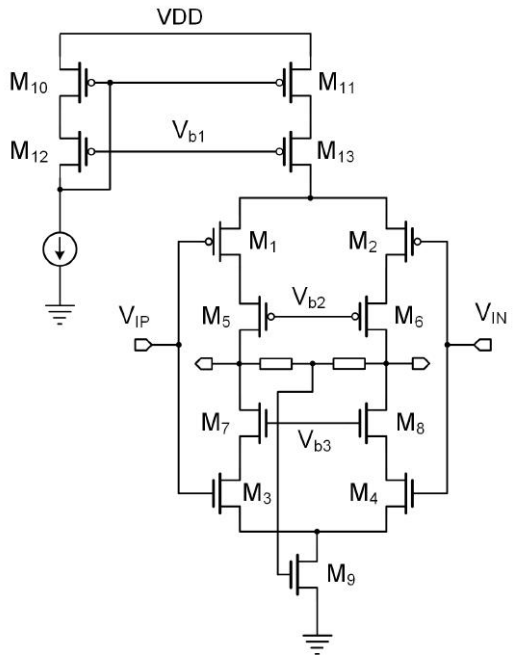


Figure 2.9: The circuit schematic of the fully differential low-noise OTA with complementary input stage.

where g_{mX} is the transconductance of the transistor M_X, and r_{oX} is the output resistance of the transistor M_X. The output thermal noise is:

$$\overline{i_{no}^2} = 4kT\gamma(g_{m1} + g_{m2} + g_{m3} + g_{m4})\Delta f \tag{2.7}$$

where $k = 1.38 \times 10^{-23} J/K$ is the Boltzmann constant. The input-referred thermal noise is:

$$\overline{v_{ni}^2} = \frac{4kT\gamma(2g_{m1} + 2g_{m4})}{(g_{m1} + g_{m3})^2}\Delta f \tag{2.8}$$

Taking the flicker noise into account, the total input-referred noise power of the OTA can be expressed as:

$$\begin{aligned}\overline{v_{ni,tot}^2} = \frac{1}{(g_{m1}+g_{m3})^2}[8kT\gamma(g_{m1}+g_{m3})+ \\ 2(\frac{K_N g_{m3}}{C_{ox,N} f W_N L_N} + \frac{K_P g_{m1}}{C_{ox,P} f W_P L_P})]\Delta f\end{aligned} \tag{2.9}$$

The flicker noise can be reduced by increasing the size of the input transistors or implementing compensation techniques like chopping. If only thermal noise is considered in the following design optimization, the input-referred noise voltage equals to

$$\overline{V_{ni,rms}} = \sqrt{\frac{8kT\gamma}{g_{m1}+g_{m3}}\frac{\pi}{2}BW} \tag{2.10}$$

The noise efficiency factor (NEF) [104] for this amplifier can be derived as:

$$\begin{aligned}NEF &= \overline{V_{ni,rms}}\sqrt{\frac{2I_{tot}}{\pi\Phi_t \cdot 4kT \cdot BW}} \\ &= \sqrt{\frac{8kT\gamma}{g_{m1}+g_{m3}}\frac{\pi}{2}BW\frac{2I_{tot}}{\pi\Phi_t \cdot 4kT \cdot BW}} \\ &= \sqrt{\frac{2\gamma I_{tot}}{(g_{m1}+g_{m3})\Phi_t}}\end{aligned} \tag{2.11}$$

Thus, a lower NEF (the lower the better) can be expected if a higher power efficiency (g_m/I_{tot}) is achieved.

In this work, since complementary input devices are used, the overall transconductance can be approximately doubled without increasing the quiescent current. Besides, all the input transistors are biased in the sub-threshold region to achieve a

high-efficiency [118]. In the above threshold operation:

$$g_m = \sqrt{2\mu C_{ox}\frac{W}{L}I_D} \quad \text{and} \quad g_m \propto \sqrt{I_D} \tag{2.12}$$

In the sub-threshold operation:

$$g_m = \frac{\kappa I_D}{\Phi_t} \quad \text{and} \quad g_m \propto I_D \tag{2.13}$$

where Φ_t is the thermal voltage. Thus, the sub-threshold operation gives a higher transconductance than the above threshold operation for the same drain current. Notice that the sub-threshold operation has a limited bandwidth due to the large parasitic of the large device dimension. But since the neural signal has a low bandwidth in nature, it is usually not a limiting factor in the neural amplifier design. A simulation result shows that 98% of the noise is from the four input transistors, and that the flicker noise contributes more than the thermal noise in the frequency range from 1 to 10kHz.

The cascode transistors (M5-M8) are used to increase the voltage gain. However, this is usually at the price of limiting the voltage headroom, and can be a challenge in low-supply voltage with advanced CMOS technology. The simulation shows an open-loop gain of 90dB is achieved in this OTA under 1μA. In a later design, a second stage is used to provide sufficient gain, which will be discussed in section 2.5.4.1. The common mode feedback (CMFB) loop is merged in the main current path to further reduce the total current. Pseudo resistors are used to get the common mode voltage without loading the amplifier. However, the common-mode output of this design also depends on the threshold voltage of the tail transistor. A later design addresses this problem by introducing an additional CMFB circuitry with low current.

A high input impedance reduces the signal attenuation. In a practical neural recording experiment with a multi-channel electrode array, the recording electrode and the reference electrodes are usually not the same type of electrode, and may have large impedance difference. Thus, even if the neural amplifier achieves a perfect common mode rejection, it cannot reject the conversion of the common-mode signal to differential-mode due to the electrode mismatch. This problem can be relieved by increasing the input impedance of the neural amplifier. Positive current feedback [93] can be used to boost the input impedance by providing the driving current required at the input stage. A post-layout simulation of the input impedance boosting is shown in Fig. 2.10.

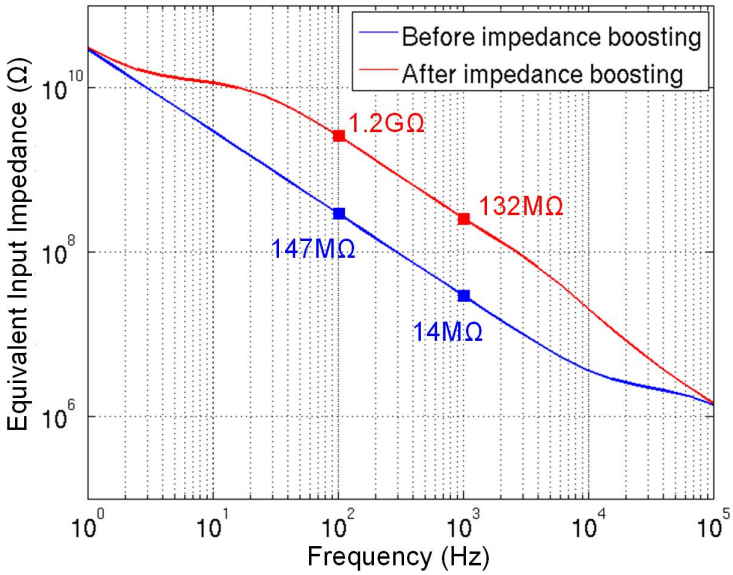


Figure 2.10: Simulation of the input impedance boosting.

A programmable capacitor array (C_L) is put at the output of the low-noise OTA. The bandwidth of the closed-loop amplifier is given by:

$$BW = \frac{g_m}{C_L}\frac{C_2}{C_1} \tag{2.14}$$

The C_L can be programmed using two bits. The bandwidth of the OTA can also be tuned by changing the biasing current.

The programmable gain amplifier (PGA) implemented in this work (2.6) is a classical 3-opamp amplifier. The gain is set by the resistors' ratio and can be chosen from 7, 10 and 19. Thus the maximum gain of a recording channel is 1,900. Additional analog buffers are added to the debugging points to drive the IO pads directly.

2.2.3 Measurement Results

The design has been fabricated in IBM 180nm CMOS technology. The micrograph of the chip is shown in Fig. 2.11. The occupied silicon area of the full chip is $4.5\times1.5\text{mm}^2$, including IO pads.

Bench testing was conducted to verify the function and performance of the fabricated chip. Fig. 2.12 shows a measurement of the neural amplifier's output with a 1kHz sinusoidal signal. A resistor divider consists of a 2kΩ and 1Ω was applied at the output of the function generator, giving a gain of 1/2001. The neural amplifier was configured to have the maximum gain of 1,900. The measured real gain was 1892.94, which corresponds to an absolute gain error of 0.37%.

The measured differential-model and common-mode frequency response of the low noise amplifier is shown in Fig. 2.13. The closed-loop gain is set to be 60dB. The

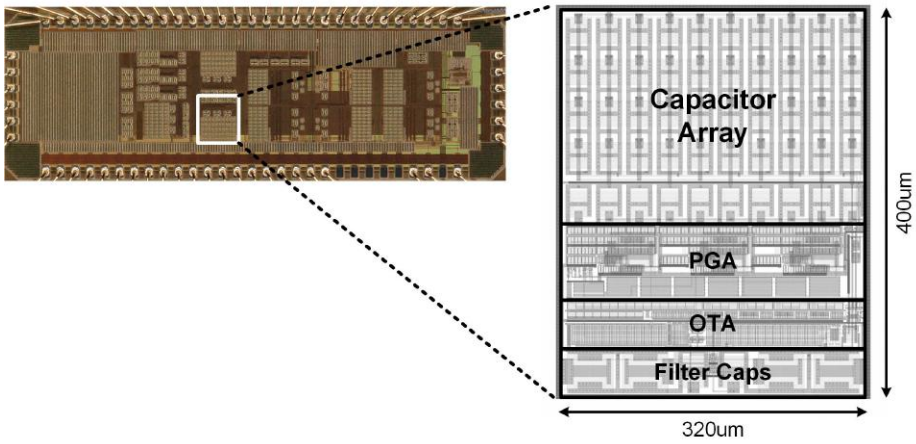


Figure 2.11: The microphotography and layout of one channel of the neural recording front-end. The major building blocks are highlighted.

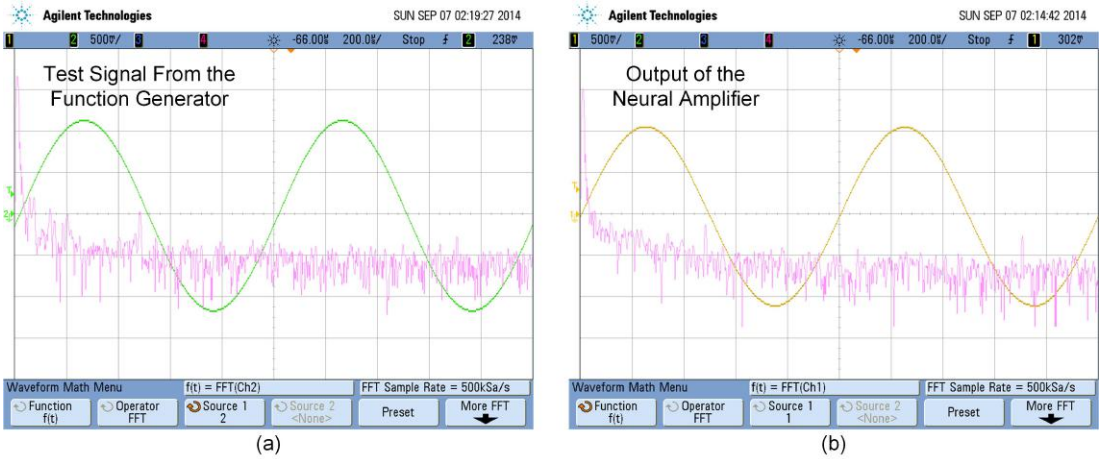


Figure 2.12: The measured response of the neural amplifier with a 1kHz sinusoidal signal. A resistor divider of 2kΩ and 1Ω is applied at the output of the function generator, and the amplifier is configured with the maximum gain of 1900. The gain error is 0.37%.

low-frequency cutoff f_L is approximately 0.5Hz. The measurement shows a CMRR above 110dB.

The input-referred noise spectrum is shown in Fig. 2.14. An integration under this curve from 1Hz to 7kHz yields a rms noise voltage of 2.55μV. This noise level

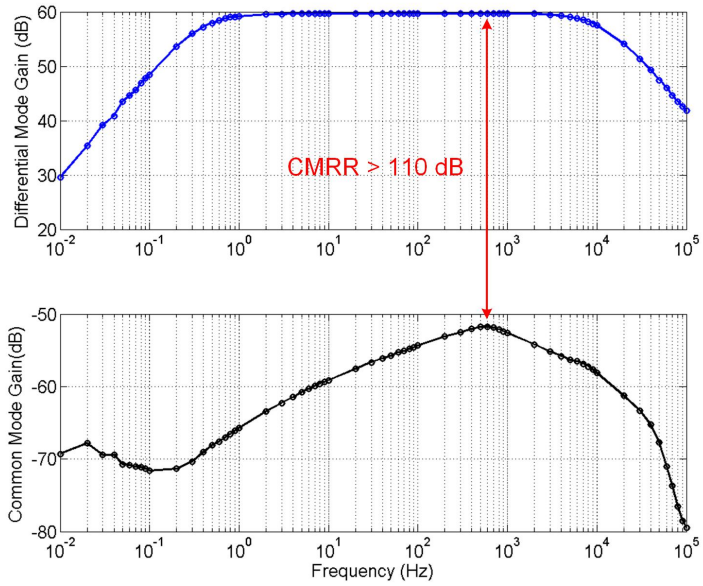


Figure 2.13: The measured differential-model and common-mode frequency response of the low-noise neural amplifier.

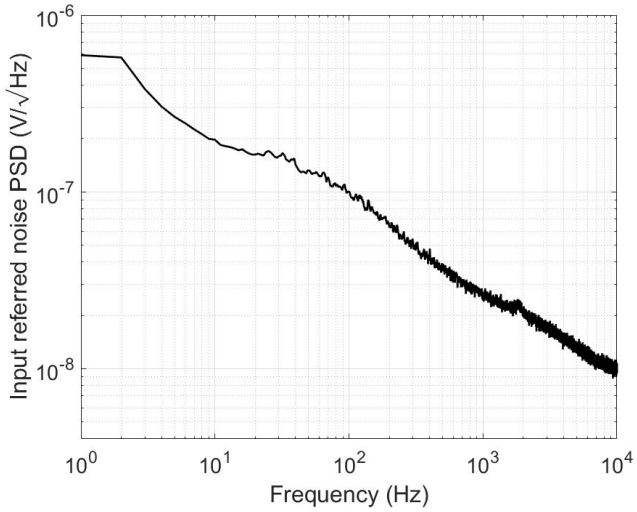


Figure 2.14: The measured input-referred voltage noise spectrum. An integration under this curve from 1Hz to 7kHz yields a rms noise of 2.55μV.

was measured with a closed-loop gain of 60dB, and the input electrodes were shorted

using an internal switch. The noise density was calculated by:

$$NoiseDensity = \frac{V_{rms}}{\sqrt{BW\pi/2}} \tag{2.15}$$

The noise density in the 7kHz bandwidth is 24.3nV/rtHz. The calculated NEF is 1.68, and the PEF is 9.38.

The summarized measured specifications of the design is listed in Table 2.5.

Table 2.5: The Neural Front-end Specifications Summary

Parameter	**Value**
Process	180nm CMOS
Supply voltage	3.3V
LNA current	2μA (biasing current not included)
Closed-loop gain	40dB
Gain error	0.37%
Bandwidth	1 $\sim$ 7kHz
Integrated noise	2.55μV
Noise Density	24.3nV/rtHz
NEF (Eq. 2.1)	1.68
PEF (Eq. 2.2)	9.38
Input range	4mV
CMRR	>110dB

In summary, this work presents the design of a general-purpose wideband low-noise neural amplifier. The design achieves a low noise floor, an accurate gain, a good CMRR in a good power efficiency. The design was later used in *In-Vivo* neural signal acquisition.

Table 2.6 compares the measured performance of this work with prior published neural recording front-end design. This achieves a comparable performance among the state-of-the-art designs.

Table 2.6: Comparison with Prior Works

Work	'03 [16]	'07 [92]	'07 [88]	'10 [113]	'13 [117]	'14 [10]	'14
Publication	JSSC	JSSC	JSSC	JSSC	JSSC	JSSC	This work
Technology	1.5um	0.8um	0.5um	180nm	180nm	180nm	180nm
Noise (μV)	2.2	0.95	2.26	1.3	0.91	5.23	2.55
BW (Hz)	0.025 -7.2k	0.05 -100	0.5 -1k	100	100	7k	1-7k
Current (μA)	16	1	11.1	3.5	NA	0.97	2
Supply (V)	5	1.8-3.3	3	1	1	1.8	3.3
NEF	4.03	4.6	9.2	9.4	5.1	1.77	1.68
PEF [2]	81.2	38.1	253.9	88.4	26.2	5.6	9.3

[2]Not provided by the author, but calculated using Eq. 2.2.

2.3 A Pre-whitening Neural Amplifier

2.3.1 Introduction

The power spectrum of electrocorticography (ECoG) and local field potential (LFP) have a characteristic $(1/f)^n$ drop with frequency [53]. This phenomenon has been observed in multiple species including humans [119]. At frequencies around 1Hz, the signal amplitude can be as large as a few millivolts, and attenuates at $1/f^2$ until 80Hz, then attenuates at $1/f^4$ [120]. At the same time, the noise power density of the CMOS front-end is usually inversely proportional to the frequency [121].

$$\overline{V_n^2} = \frac{K}{C_{ox}WL} \cdot \frac{1}{f} \tag{2.16}$$

where K is a process-dependent parameter on the order of 10^{-25} V^2F. This suggests that the SNR of the recording front-end improves as the frequency decreases, as illustrated in Fig. 2.15 (a). Intuitively, if a wideband recording front-end is designed

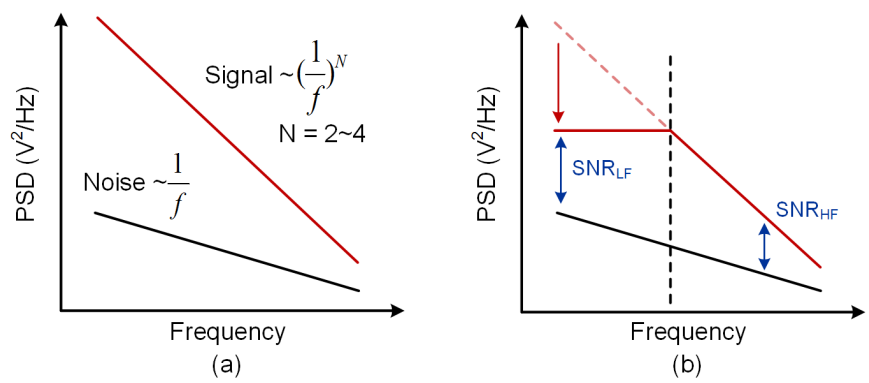


Figure 2.15: Illustration of the prewhitening filter. (a) The neural signal displays a $1/f^n$ power characteristic, while the recording front-end has a $1/f$ noise power characteristic. (b) The pre-whitening filter shapes the frequency response of the recording front-end to reduce the overall dynamic range requirement, while still preserves a sufficient SNR.

to achieve the voltage swing requirement for the low-frequency signal, at the same time preserves the SNR for the high-frequency signal, it needs to have an ultra-high dynamic range. The high dynamic range wideband low-noise amplifier and high-resolution ADC design are challenging and will cost high power consumption.

In this work, a pre-whitening filter is proposed to address this problem. The basic idea of the pre-whitening processing is illustrated in Fig. 2.15 (b). If we use a lower gain for the low-frequency signal, a sufficient SNR may still be preserved for the recording purpose, and the dynamic range requirement of the system can be significantly reduced. Since the frequency shaping processing is similar to a whitening filter, which turns the signal more like a white signal, the filter is named pre-whitening filter in this work. The simplest way to implement this pre-whitening filter is via a highpass filter. If the cut-off frequency of the highpass filter is known, the original signal can be recovered during the post-recording processing. It should be noticed that it is important to decide the cut-off frequency in practice, in order to preserve sufficient SNR for the low-frequency signal. A programmable filter is helpful in this case.

In summary, a pre-whitening neural recording front-end is proposed. In the pre-whitening front-end, the frequency response of the neural amplifier is shaped according to the characteristic of the neural signal. The design significantly reduces the dynamic range requirement of the neural amplifier and the ADC resolution without sacrificing the signal quality. In the following sections, the possible circuit implementations of the pre-whitening neural amplifier are analyzed, the key design trade-offs are described, and the simulation and experimental results of the proposed design are presented.

2.3.2 Analysis of the Pre-whitening Neural Amplifier Design

The thermal noise power spectral density of a resistor is given here for convenience:

$$\overline{V_{nR}} = \sqrt{4kTR} \tag{2.17}$$

where k = 1.38 $\times 10^{-23}$ is the Boltzmann's constant, T is the absolute temperature in Kelvin. If a recording electrode has an impedance of 100kΩ, it should have a noise density of $40.7nV/\sqrt{Hz}$. Assume the frequency interest of a wideband neural signal is from 1Hz to 10kHz, the electrode gives an integral thermal noise of 4.07μV in this frequency range.

Consider the simplest case of a 1st order RC highpass filter. Fig. 2.16 shows the circuit and the noise source.

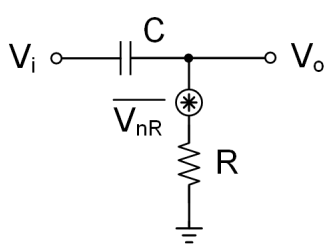


Figure 2.16: A 1st order RC highpass filter with noise source.

The equivalent output noise of the RC highpass filter is given by:

$$\begin{aligned} \overline{V_{n,o}} &= \frac{1}{1+sRC} \cdot \overline{V_{nR}} \\ &= \frac{\sqrt{4kTR}}{1+sRC} \end{aligned} \tag{2.18}$$

While the input-referred noise of the RC highpass filter is given by:

$$\begin{aligned} \overline{V_{n,i}} &= \frac{1}{sRC} \cdot \overline{V_{nR}} \\ &= \frac{\sqrt{4kTR}}{sRC} \end{aligned} \tag{2.19}$$

The input-referred noise increases with decreasing frequency, as the signal attenuates. This is an important observation and provides some intuition for the following analysis. Fig. 2.17 shows a simulation of a 1st order RC highpass filter. Both the output

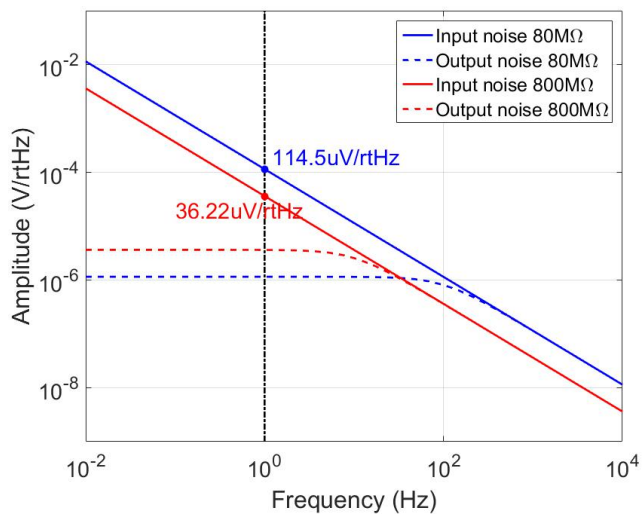


Figure 2.17: Noise simulation of RC highpass filters with frequency corners at 10Hz and 100Hz. The capacitor value is set to be 20pF.

and the input-referred noise are plotted. The capacitor value is set to be 20pF, and the resistor values are set to be 800MΩ and 80MΩ, and the cut-off frequency is 10Hz and 100Hz, respectively. The input-referred noise densities at the 1Hz are marked in the figure. Notice that using a larger capacitor value with the same cut-off frequency can achieve a lower noise density. However, large capacitors take a lot of silicon area, thus is not suitable for multiple channel recording front-end integration.

In summary, a simple RC filter is not suitable for implementing the proposed pre-whitening filter. In the following section, two methods of implementation are discussed: i) filtering after a wideband LNA, and ii) filtering at the direct neural interface.

A. Pre-whitening Filter After a Wideband Low-noise Amplifier

The block diagram of a pre-whitening filter after a wideband low-noise amplifier is shown in Fig. 2.18. Since the filtering is implemented after the wideband amplifier,

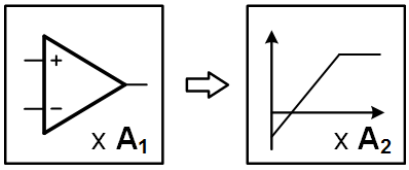


Figure 2.18: The block diagram of a pre-whitening filter after a wideband low-noise amplifier.

the noise from the filter will be attenuated by the gain of the wideband amplifier. Again, assume using a simple RC filter, Fig. 2.19 shows the noise source of the RC filter after the wideband amplifier.

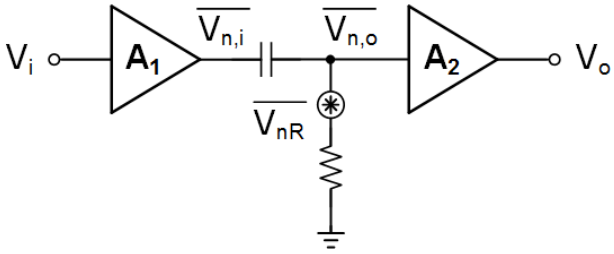


Figure 2.19: The circuit schematic of a pre-whitening filter after a wideband low-noise amplifier. Noise source from the resistor is shown.

The input-referred noise of the recording front-end from the filter is then given by:

$$\begin{aligned}\overline{V_{i,rms}} &= \frac{1}{A_1}\sqrt{\int_{f_L}^{f_H} \overline{V_{n,i}^2} \cdot df} \\ &= \frac{1}{A_1 \pi C}\sqrt{\frac{kT}{R}(\frac{1}{f_L} - \frac{1}{f_H})}\end{aligned} \tag{2.20}$$

If the frequency range of interest is 1Hz to 10kHz, and the 1st stage wideband neural amplifier has a gain A_1 of 100. If we set the highpass frequency of the second stage to be 10Hz, the integral noise is 0.11μV, and if we set the highpass frequency to be 100Hz, the integral noise is 1.14μV. In both cases, the integral noise is lower than the thermal noise of an electrode with an impedance of 100kΩ (Section 2.3.1).

Several active highpass filters can achieve lower input-referred noise than the simple RC filter. Consider the circuit with capacitive feedback in Fig. 2.20. The

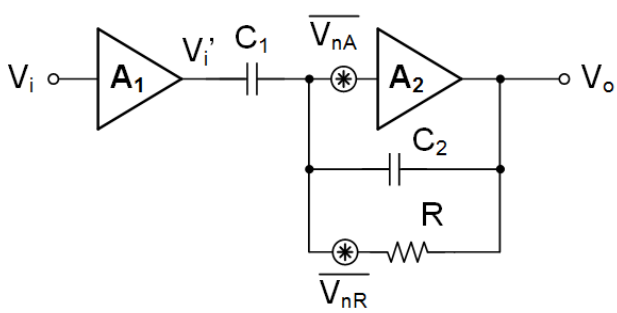


Figure 2.20: An implementation of an active highpass filter. Noise sources from the resistor and the 2nd stage amplifier are shown.

signal transfer function can be expressed as:

$$\begin{aligned}H_{sig}(s) &= \frac{sRC_1}{sRC_2 + 1} \\ &= \frac{C_1}{C_2} \cdot \frac{s}{s + \frac{1}{RC_2}}\end{aligned} \tag{2.21}$$

The midband gain of the amplifier A_{CL} is $\frac{C_1}{C_2}$, and the highpass frequency is set by $\frac{1}{RC_2}$. The amplifier A2's noise transfer function can be expressed by:

$$\begin{aligned} H_{nA}(s) &= 1 + \frac{sRC_1}{sRC_2 + 1} \\ &= 1 + \frac{C_1}{C_2} \cdot \frac{s}{s + \frac{1}{RC_2}} \end{aligned} \tag{2.22}$$

Since the amplifier's noise transfer function and the signal's noise transfer function has the same highpass frequency $\frac{1}{RC_2}$, the amplifier's noise is shaped in the same way as the signal. Thus, the noise increase from the amplifier won't cause frequency dependent SNR degradation.

Let's look at the resistor's noise transfer function:

$$\frac{V_o - V_{nR}}{R} = sC_2V_o = 0 \tag{2.23}$$

$$H_{nR}(s) = \frac{1}{sRC_1} = \frac{1}{A_{CL}} \frac{1}{sRC_2} \tag{2.24}$$

where $\frac{1}{RC_2}$ is the signal highpass frequency, and A_{CL} is the closed-loop gain of the second stage. So compared with the implementation in Fig. 2.19, the noise is further suppressed by the gain of the 2nd stage. The overall input-referred noise density from the resistor given by:

$$\overline{V_{nR,i}} = \frac{1}{A_1 A_{CL}} \frac{\sqrt{4kTR}}{sRC_2} \tag{2.25}$$

In a practical design, the total of $\overline{V_{nR,i}}$ and the input-referred noise density of the 1st stage should be lower than the noise and the SNR requirement of the recording

system. Again, assume the cut-off frequency of the pre-whitening amplifier is 10Hz, the capacitor C_2 is 2pF, the closed-loop gain of the 1st and 2nd stage is 100 and 40, respectively. The $\overline{V_{nR,i}}$ at 1Hz is $28.6nV\sqrt{Hz}$, which is lower than the thermal noise density of a $100k\Omega$ electrode (Section 2.3.2). If the cut-off frequency of the pre-whitening amplifier is 100Hz, the noise density $\overline{V_{nR,i}}$ at 1Hz is $90.5nV/\sqrt{Hz}$, which is still lower than most low-noise neural amplifier designs at 1Hz, and is sufficient for the SNR requirements in most intracortical neural recordings.

B. Low-noise Neural Amplifier with Integrated Pre-whitening Filter

This section discusses the possible integration of the pre-whitening filter into the 1st stage low-noise amplifier. It is more challenging to design pre-whitening filter at the 1st stage because of the noise increase with decreasing frequency due to the filter's response. However, there are also advantages. The electrode interface usually has a slowly varying offset or half-cell potential up to several hundred millivolts, as reviewed in section 2.1. The recording amplifier will need to reject this large offset, typically accomplished by using a highpass filter with a cut-off frequency below 1Hz. However, it is difficult to implement such a large time-constant filter on-chip. One solution is to use transistor pseudo resistor, as described in section 2.2.2. But the pseudo resistors have reliability problem for the use in an implanted medical device, and they are susceptibility to electromagnetic interface and degradation over time [53]. If the pre-whitening filter can be integrated into the first stage amplifier, the sub-Hertz filter can be avoided.

Consider the capacitor-coupled neural amplifier in Fig. 2.21. The signal transfer

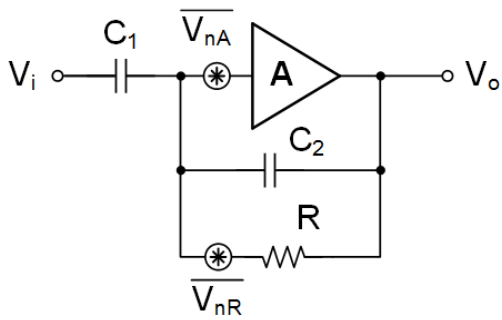


Figure 2.21: Capacitor coupled neural amplifier. Noise sources are marked in the figure.

function of this amplifier is:

$$H_{sig}(s) = \frac{C_1}{C_2} \cdot \frac{s}{s + \frac{1}{RC_2}} \tag{2.26}$$

Thus the highpass cut-off frequency is determined by $1/RC_2$. The transfer function of the amplifier noise is:

$$H_{nA}(s) = 1 + \frac{C_1}{C_2} \cdot \frac{s}{s + \frac{1}{RC_2}} \tag{2.27}$$

And the resistor noise's transfer function is:

$$H_{nR}(s) = \frac{1}{sRC_1} \tag{2.28}$$

The input-referred noise density from the resistor is:

$$\overline{V_{nR,i}} = \frac{1}{A_{CL}} \frac{\sqrt{4kTR}}{sRC_2} \tag{2.29}$$

Compared with Eq. 2.25, the only difference is that this input-referred noise is no longer attenuated by preamplifier. If A_{CL} is designed to be the product of the gain of the two stages in previous section, it can achieve the same noise performance. But it is difficult in practical designs.

There are other circuit topologies to implement the highpass frequency response in the low-noise amplifier. One type of topology is to use a DC servo loop. A typical example is shown in Fig. 2.22.

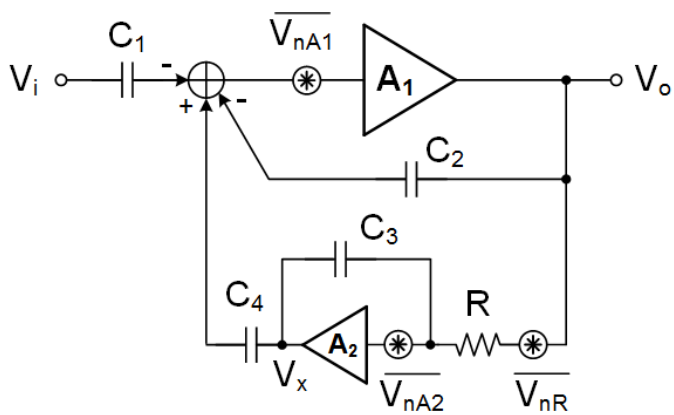


Figure 2.22: Capacitor coupled instrumentation amplifier with a DC servo loop.

$$sC_1V_i(s) + sC_2V_o(s) - sC_4V_x(s) = 0 \tag{2.30}$$

$$Vx(s) = -\frac{1}{sRC_3}V_o(s) \tag{2.31}$$

Signal transfer function can be expressed as:

$$\begin{aligned} H_{sig}(s) &= \frac{sRC_1C_3}{sRC_2C_3 + C_4} \\ &= \frac{C_1}{C_2} \cdot \frac{s}{s + \dfrac{C_4}{RC_2C_3}} \end{aligned} \tag{2.32}$$

The mid-band gain of this circuit is $\frac{C_1}{C_2}$. Compared with Eq. 2.26, the high pass frequency corner is $\frac{C_4}{C_2}\frac{1}{RC_3}$, where $\frac{1}{RC_3}$ is the frequency corner of the integrator in the feedback loop.

The noise transfer function of the amplifier A_1 is:

$$C_1V_{nA1} = C_2(V_o - V_{nA1}) + C_4(V_x - V_{nA1}) \tag{2.33}$$

$$H_{nA1}(s) = \frac{C_1 + C_2 + C_4}{C_2} \cdot \frac{1}{1 + \dfrac{C_4}{sRC_2C_3}} \tag{2.34}$$

The noise transfer function of the amplifier A_2 is:

$$\frac{V_o - V_{na2}}{R} = sC_3(V_{na2}) = sC_4Vx = sC_2V_o \tag{2.35}$$

$$H_{nA2}(s) = \frac{sRC_3 + 1}{sRC_3\dfrac{C_2}{C_4} + 1} \tag{2.36}$$

The noise transfer function of the resistor is:

$$\frac{V_o - V_{nR}}{R} = sC_3V_x \tag{2.37}$$

$$sC_4V_x = sC_2V_o \tag{2.38}$$

$$H_{nR}(s) = \frac{1}{1 + sRC_3\frac{C_2}{C_4}} \tag{2.39}$$

The input-referred noise density from the amplifier A_1 is:

$$\overline{V_{nA1,i}} = \frac{C_1 + C_2 + C_4}{C_1}V_{nA1} \tag{2.40}$$

The input-referred noise density from the amplifier A_2 is:

$$\overline{V_{nA2,i}} = \frac{C_4}{C_1}(1 + \frac{1}{sRC_3})V_{nA2} \tag{2.41}$$

The input-referred noise density from the resistor is:

$$\overline{V_{nR,i}} = \frac{C_4}{sRC_1C_3}\sqrt{4kTR} \tag{2.42}$$

Compared with Eq. 2.29, the noise contribution from the resistor also depends on the ratio of C_4/C_3. However, reduce the ratio of C_4 and C_3 decreases the input voltage headroom, which makes the trade-off difficult.

The resistor can be further replaced by a switched capacitor circuit. A simplified circuit schematic is shown in Fig. 2.23. In the switched capacitor circuit, the time-constant can be better controlled by the ratio of the capacitors and the switching frequency. However, in order to achieve the required large time-constant, a large capacitor tank is to be implemented. Several techniques have been proposed to reduce the required size of the capacitors [92, 122].

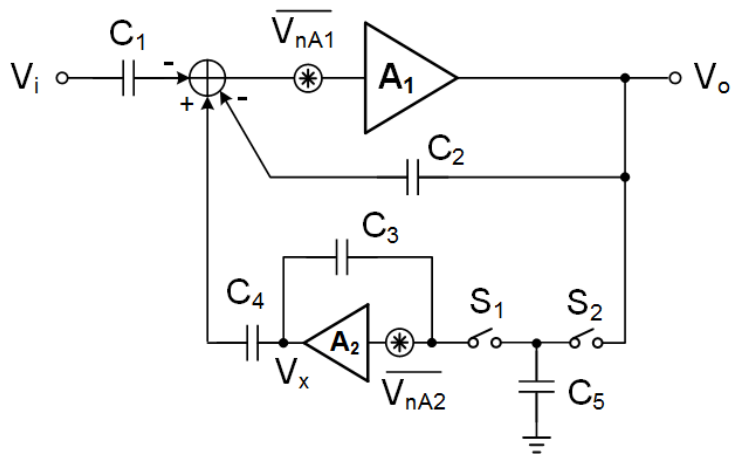


Figure 2.23: Capacitor coupled instrumentation amplifier with DC servo loop implemented by switched capacitor circuits.

If ultra low-noise is required for the low-frequency signal component, the chopping technique can be combined in the pre-whitening amplifier design. An example of a chopping pre-whitening amplifier is shown in Fig. 2.24. With chopping, the flicker

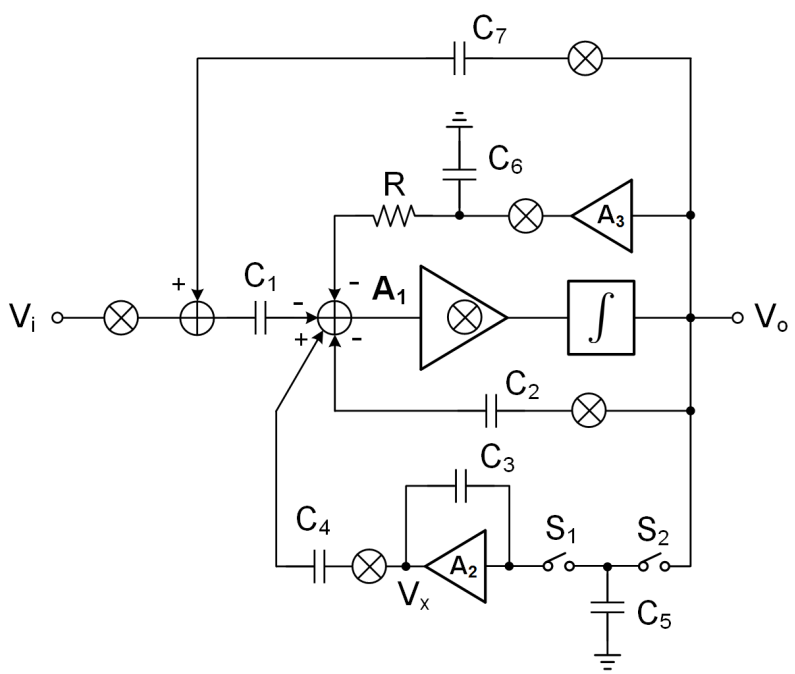


Figure 2.24: A capacitor coupled chopping amplifier with DC servo loop and input impedance boosting.

noise can be removed, and the amplifier can guarantee a good SNR for the signal even with the lower gain at the low frequency. However, there are also many trade-offs involved with the chopping amplifier design [93, 113, 116]. The effects it takes may counteract the benefits from the pre-whitening.

2.3.3 Circuit Implementation

A pre-whitening amplifier prototype of the architecture presented in section 2.3.2 is designed to demonstrate the idea. The circuit schematic of the designed pre-whitening amplifier is shown in Fig. 2.25. A single-ended architecture is used. A conventional

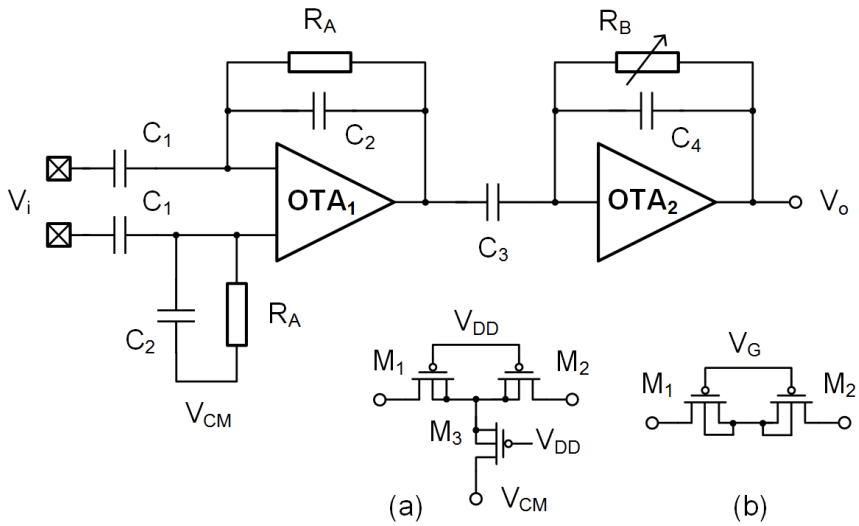


Figure 2.25: The circuit schematic of the designed pre-whitening amplifier. The Pseudo resistors R_A and R_B used in the 1st and 2nd stage is shown in subplot (a) and (b), respectively.

low-noise current mirror OTA is used in both stages [16]. A T-connected pseudo-resistor (TPR) proposed in [10] is used as the feedback resistor in the 1st stage. If the equivalent resistance of transistor X is R_X, the total equivalent resistance of the TPR

is $R_1+R_2+R_1 \cdot R_2/R_3$. The Pseudo resistor in the 2nd stage is the same as the one used in amplifier presented in the previous section 2.2. The gate voltage can be used to tune the resistance over a large range, which is used to tune the cut-off frequency of the pre-whitening filter. The 1st stage has a closed-loop gain of 100, and the 2nd stage has a closed-loop gain of 40.

2.3.4 Measurement Results

The design has been fabricated in IBM 180nm CMOS technology. The micrograph of the chip and the layout of one recording channel are shown in Fig. 2.26. The full chip

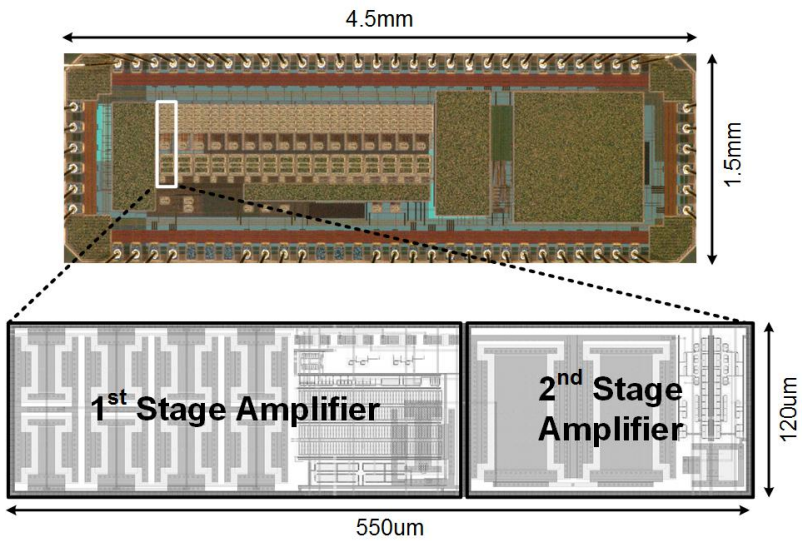


Figure 2.26: The microphotography and layout of one channel of the pre-whitening amplifier.

occupies a silicon area of $4.5\times1.5\text{mm}^2$, including IO pads. One recording channel has a dimension of $550\mu\text{m}\times120\mu\text{m}$.

Bench testing was conducted to evaluate the performance of the fabricated chip. Fig. 2.27 shows both the amplitude and phase frequency response of the pre-whitening

amplifier. The simulated frequency response is also plotted in dashed lines for com-

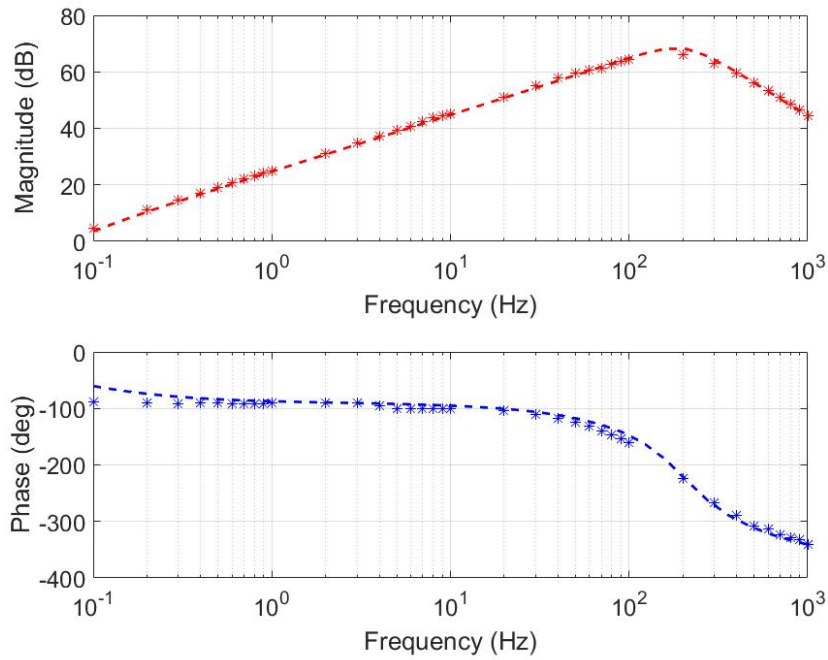


Figure 2.27: The measured frequency response of the pre-whitening amplifier in comparison with the simulation result.

parison. With the information of the frequency response, the original signal can be recovered from the pre-whitened recording.

The synthetic neural signal was generated using an arbitrary function generator 33521A from Agilent to the test the pre-whitening amplifier. A 1-min neural signal contains local field potentials is used for testing. The signal was recorded using RZ2 workstation from Tucker-Davis Technologies. The signal was sampled at 24.4kSps in a resolution of 24-bit. A resistor divider consists of 2kΩ and 1Ω was applied at the output of the function generator, gives a gain of 1/2001. The neural amplifier was configured to have a maximum gain of 4,000.

The designed amplifier can be configured to do both conventional wideband recording or frequency shaping pre-whitening recording. The power spectral density (PSD) was calculated for 24 channel of the LFP recording. Fig. 2.28 shows a comparison of the PSD of the conventional wideband recording and the pre-whitening recording. The result clearly shows that the spectrum of the pre-whitening record-

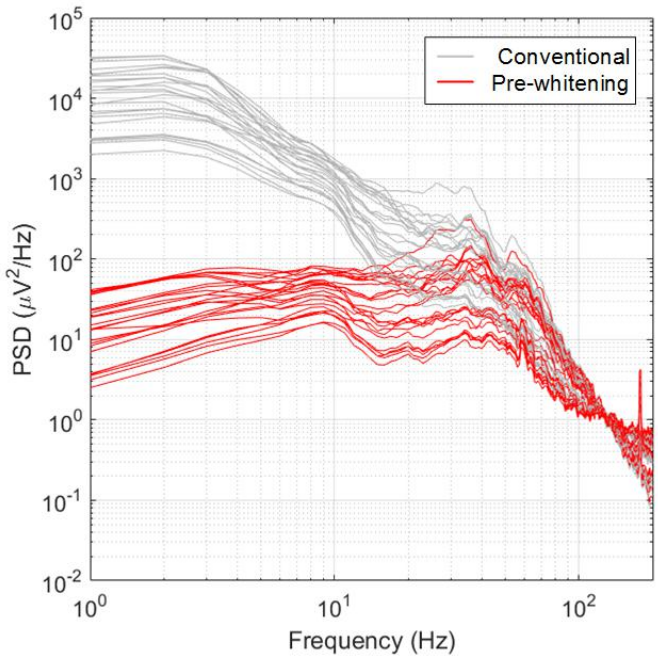


Figure 2.28: Comparison of the PSD of the pre-whitening amplifier and the original signal.

ing was flattened at the low-frequency range, which saves the voltage headroom by more than an order of magnitude. The reduction in the dynamic range relaxes the requirement of the linear range of the low-noise amplifier and the ADC design.

The reconstruction of the signal was performed in Matlab. Fig. 2.29 shows a comparison of 10 seconds of the conventional recording, the pre-whitening recording, and the reconstruction from the pre-whitening recording. Pearson correlation coeffi-

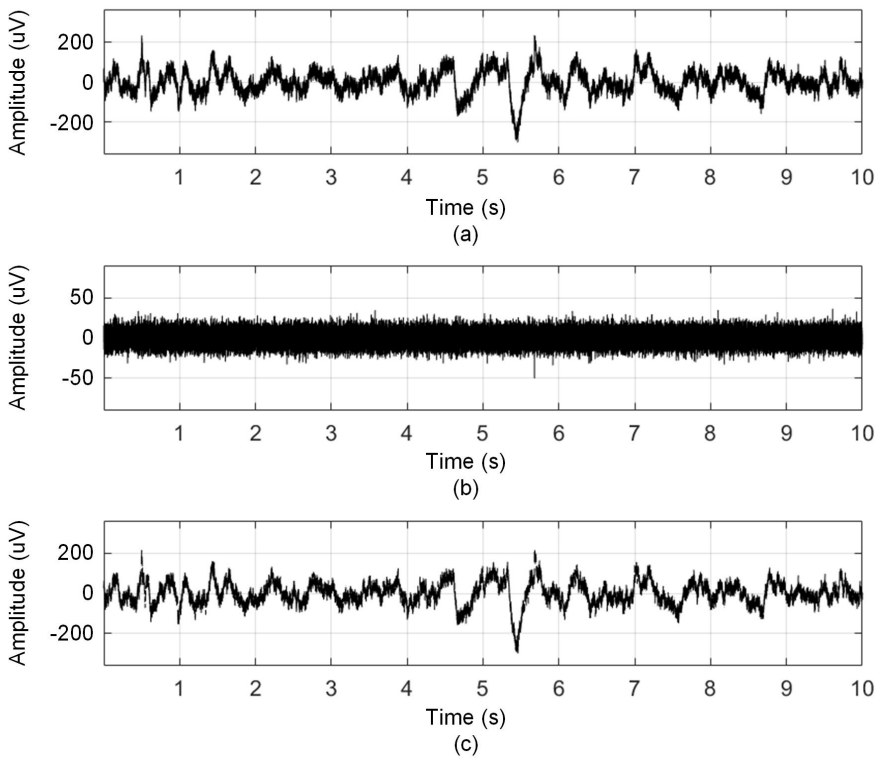


Figure 2.29: Comparison of (a) the wideband signal, (b) the measured output of the pre-whitening amplifier, and (c) the reconstructed signal from the pre-whitening amplifier's recording.

cient is used here to evaluate the accuracy of the reconstruction [123]. The correlation coefficient is defined as:

$$\rho(x,y) = \frac{1}{N-1}\sum_{i=1}^{N}(\overline{\frac{x_i-\mu_x}{\sigma_x}})(\overline{\frac{y_i-\mu_y}{\sigma_y}}) \tag{2.43}$$

where μ_x and σ_x are the mean and standard deviation of the signal x, and μ_y and σ_y are the mean and standard deviation of the signal y. The result shows a correlation coefficient of 97.6%, which indicates a faithful recovering of both phase and amplitude. Noticed that a high-order zero-phase digital filter of 1 to 200Hz was applied before the comparison. But even in this case, both phase and amplitude will need to be

recovered at the same time to completely reconstruct the original signal.

Power spectral density estimation was calculated by periodgram for both the original signal and the reconstructed signal from the pre-whitening recording, as shown in Fig. 2.30. The result indicates that the pre-whitening processing can provide a

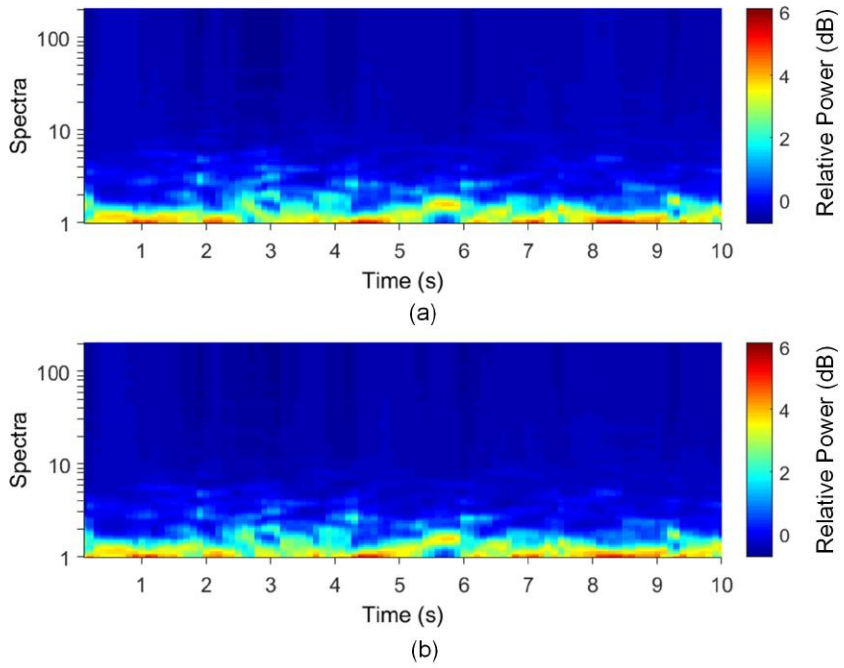


Figure 2.30: Comparison of the spectrum of (a) the original signal and (b) the reconstructed signal from the pre-whitening amplifiers recording.

faithful reconstruction of the spectrum content.

In summary, the design of the pre-whitening amplifier takes advantage of the characteristics of the neural signal. Since the power density of the neural signal including ECoG and LFP drops faster with frequency than the filter noise of the CMOS recording front-end, there is an opportunity to design a recording front-end with less gain at low frequency, at the same time, preserve sufficient SNR for the wideband signal. The design significantly reduces the dynamic range and linearity

requirement of the low-noise amplifier and the ADC. The circuit implementation of the pre-whitening front-end is analyzed in this section with a detailed noise analysis. A prototype was designed and fabricated, and experimental results are presented in comparison with simulation and theoretical computation. Since the pre-whitening amplifier provides an opportunity to improve the performance of neural recording-end without a power penalty, it can be advantageous to include it into a high channel-count neural recording front-end system.

2.4 Design of a Low-Power Analog-to-Digital Converter

2.4.1 Introduction

A low-power analog-to-digital converter (ADC) is an essential component in a neural interface system. In a typical bi-directional neural interface system, ADCs can be used for digitization of the neural signal, the extracted neural features, the sensory signal, and the stimulation compliance voltage. Among all ADC topologies, successive approximation register (SAR) ADCs have an advantage in power efficiency for a moderate sampling rate. Firstly, a SAR ADC does not require a high gain and high bandwidth opamp for high accuracy and linearity. Secondly, SAR logic mainly consists of digital circuits, the speed and power of which scales down with the deep sub-micron CMOS technology. Thirdly, if a capacitive DAC is used, no static power is consumed, thus the power scales with the sampling rate. Comprehensive reviews and tutorials of SAR ADC design can be found in [124–126].

Recently, a lot of techniques for power-efficient SAR ADC designs have been reported. Among these techniques are split capacitor array [127, 128], monotonic capacitor switching [18], partial floating capacitor switching [129], step-charging design [130], reference free design [131], asynchronous timing [132], and so on. In addition to techniques for general-purpose SAR ADC designs, several techniques have been reported to optimize power consumption, especially for neural or sensory signal digitization. Among them are:

- Adaptive resolution or dynamic range: including changing ADC resolution [133], or adding additional programmable gain amplifier before the ADC [134]
- Data dependent or data-driven sampling: for example, combine action potential detection and digitization together [135]. Besides, the sampling rate can also be adapted to the activity using continuous time level-crossing sampling [136].
- Delta difference sampling: since the neural signal has both slow and fast oscillations over time, normal sampling during slow activity period is not energy efficient. So digitize only the difference [10, 137], using a bypass window [138], or using LSB-first approach [139] can achieve better power efficiency.

In this section, the design of a voltage-mode 10-bit SAR ADC is presented. Specifications are analyzed, circuit design details are described, and measurement results are presented.

2.4.2 Circuit Implementation

The architecture of the 10-bit voltage-mode SAR ADC is shown in Fig. 2.31. The major building blocks are: i) comparator, ii) SAR logic, iii) DAC, and iv) sample and hold switch. 14 clock cycles are used to finish one conversion, allowing 4 clock cycles for sampling.

A commonly used capacitive DAC is employed in this SAR ADC. Since the required capacitor in a conventional binary capacitor array can be very small without compromising the ENOB, custom designed capacitors are often used to achieve a minimum total input capacitance and thus ultra low-power [18, 140]. However, these

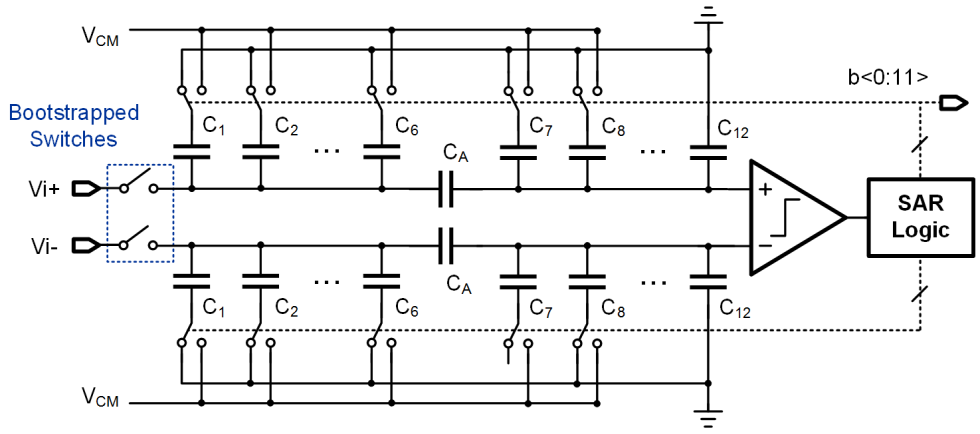


Figure 2.31: The architecture of the 10-bit voltage-mode SAR ADC.

designs usually require custom characterization for a specific fabrication process. In this work, a split capacitor array is adopted to reduce the total capacitance, lowering the power consumption and area. The capacitors are realized as a standard metal-insulator-metal (MIM) structure.

A monotonic switching procedure is applied to minimize the power consumption from unnecessarily charging and discharging of the capacitor array [18]. A comparison of the waveforms of the conventional switch to the V_{CM} procedure and the monotonic switching procedure is shown in Fig. 2.32. In the monotonic switching procedure,

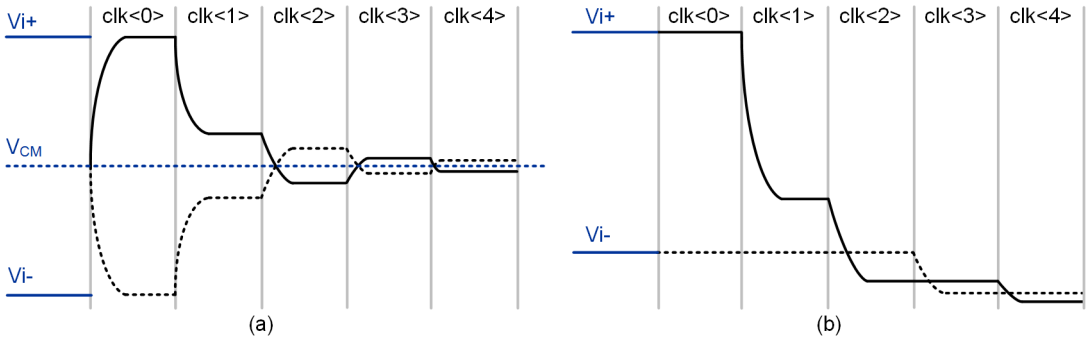


Figure 2.32: A comparison of the waveforms of (a) the conventional switching to V_{CM} procedure, and (b) the monotonic switching procedure. Modified from [18].

the first comparison is performed without switching, and the total capacitance is half of the conventional capacitive SAR ADC's DAC array [18].

Fig. 2.33 shows the circuit schematic for the timing generation module. A global

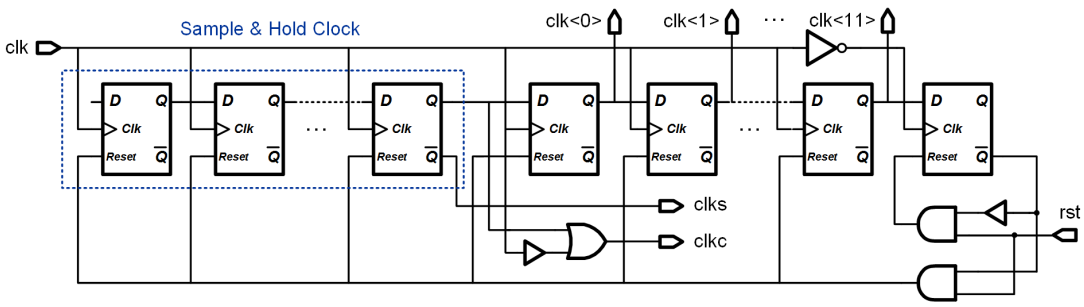


Figure 2.33: The circuit schematic for the SAR timing generation module. *clk* is the input clock, *clks* is the signal for the sampling switch, *clkc* is the clock for the comparator, and *clk*[x] is for the bit[x] of the DAC.

reset signal is used to synchronize the start of the conversion, and the control logic generation is cyclic.

The sample and Hold (S/H) circuit is critical in achieving good SFDR for an ADC design. The bootstrapped switch is commonly used since it provides a constant small on-resistance [141]. The circuit schematic of the bootstrapped switch implemented in this work is shown in Fig. 2.34. The gate to source voltage of the switch transistor is fixed at the supply voltage by the capacitor C_S.

Fig. 2.35 shows the circuit schematic of the comparator. The comparator consists of a pre-amplifier and a dynamic latch. Since the input voltage has a range from the ground to V_{cm}, the comparator uses a PMOS input stage. The current source NMOS are used in parallel with the diode-connected NMOS for increasing the gain [121]. The pre-amplifier provides moderate gain to reduce the equivalent mismatch due to the latch. The latch consumes no static current. When *clkc* (as shown in Fig. 2.33)

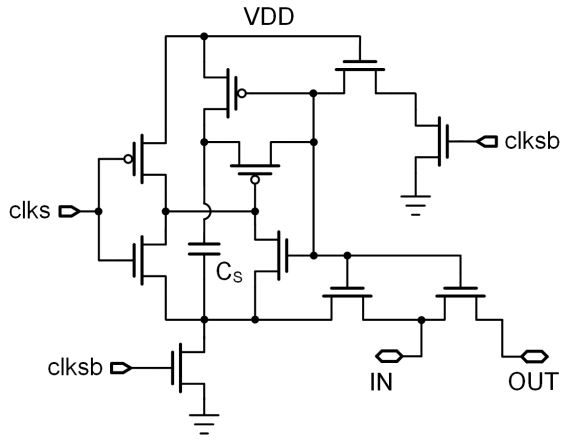


Figure 2.34: The circuit schematic of the bootstrapped switch.

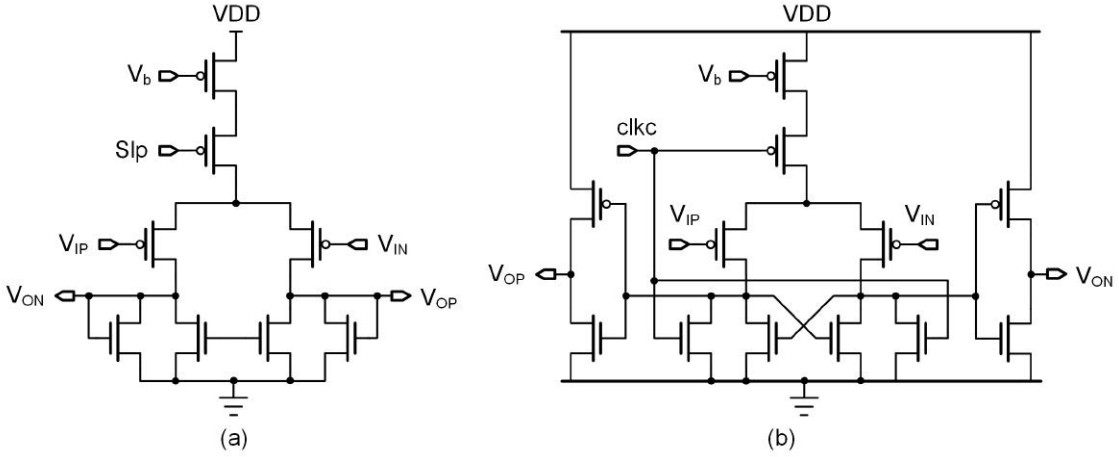


Figure 2.35: The circuit schematic of the comparator. (a) A pre-amplifier, and (b) a dynamic latch.

is high, the outputs are reset to high. When *clkc* goes to low, the regeneration latch forces one output to high and the other to low. The SAR logic only takes the V_{OP} and generates an inverted signal V'_{ON} to avoid the metastability problem.

A Class-AB output stage has been designed to drive the sample-hold circuits of the following ADC stage. To digitize a single-ended signal, a single-to-differential converter (S2D) can be integrated. An example of the S2D circuit is shown in Fig.

2.36. The resistor values are designed to be R_1=R_3=R_4, and the voltage gain is

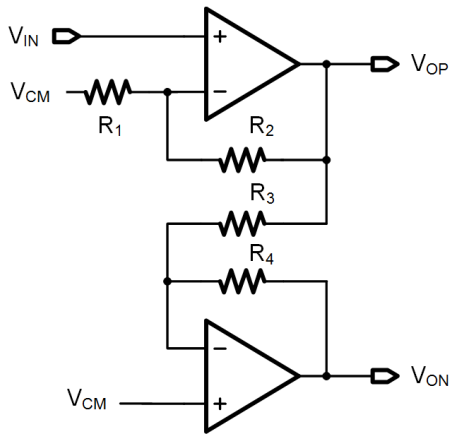


Figure 2.36: The circuit schematic of a single-to-differential converter.

$2(1+R_2/R_1)$. R_2, which can be programmed by a shift register.

2.4.3 Measurement Results

The ADC has been fabricated in IBM 180nm CMOS technology. The layout of the 10-bit SAR ADC is shown in Fig. 2.37 with major building blocks highlighted. The total occupied silicon area is 220μm $\times$ 190μm. The measurement results of the prototype are presented below.

The differential nonlinearity (DNL) and integral nonlinearity (INL) of the ADC were measured using slow ramps. The result is shown in Fig. 2.38. The peak DNL and INL are -0.49/+0.56LSB, and -0.82/+0.77LSB, respectively.

The SAR ADC's dynamic performance was measured with a low-frequency input tone and a near Nyquist input tone. The output spectrums are shown in Fig. 2.39 and Fig. 2.40, respectively.

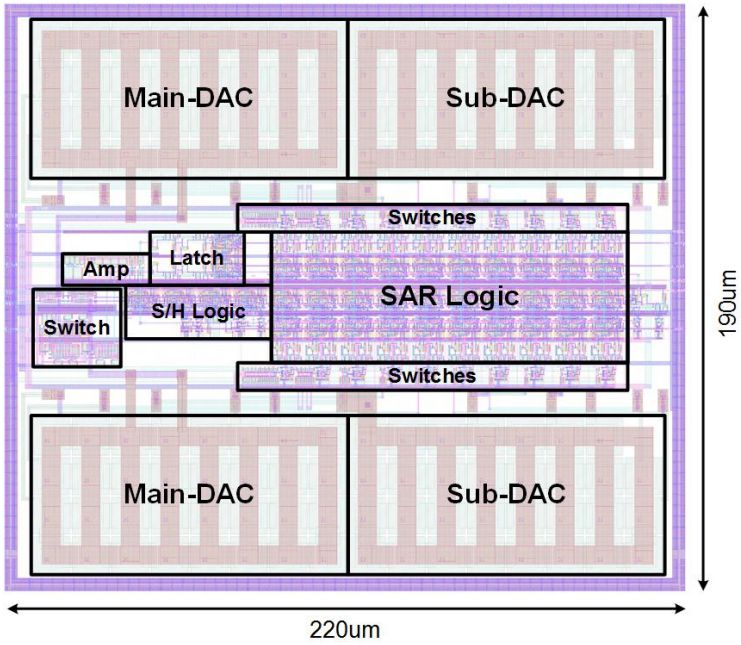


Figure 2.37: The layout of the 10-bit SAR ADC with major building blocks highlighted.

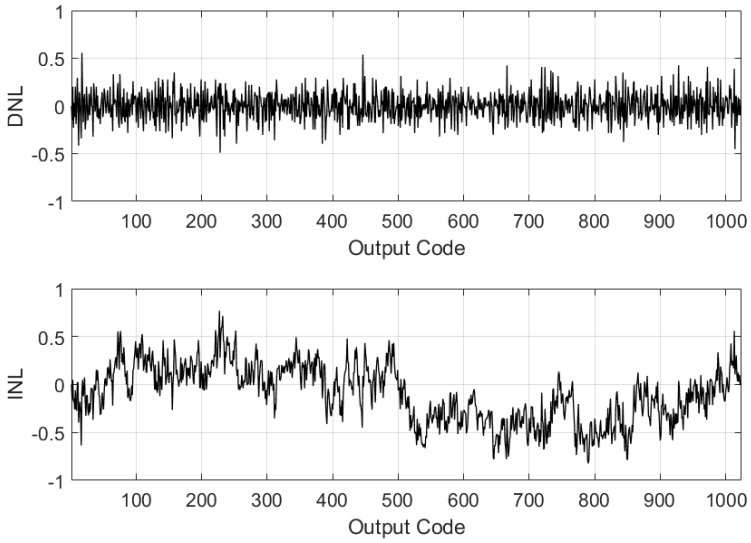


Figure 2.38: Measured DNL and INL of the 10-bit SAR ADC. The worst DNL is -0.49/+0.56LSB, and the worst INL is -0.82/+0.77LSB.

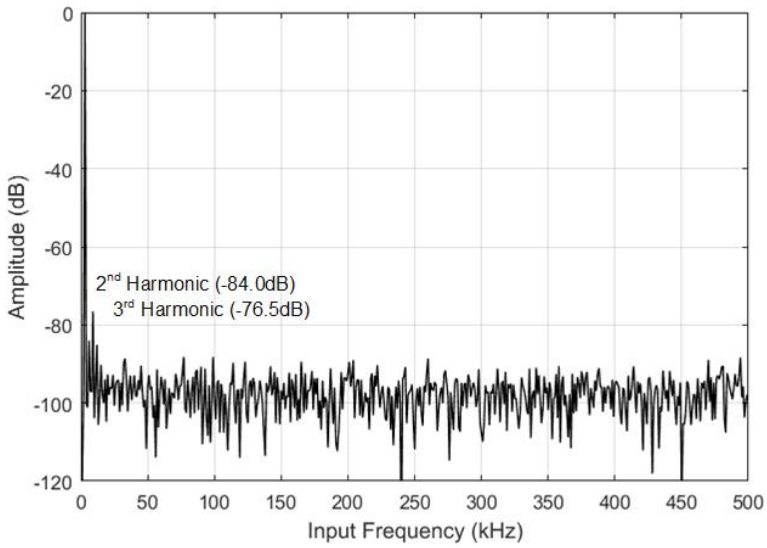


Figure 2.39: Measured FFT spectrum at 1MS/s with an input tone of 3kHz. The SFDR is 76.54dB and the SNDR is 56dB. The ENOB at 3kHz is 9.01.

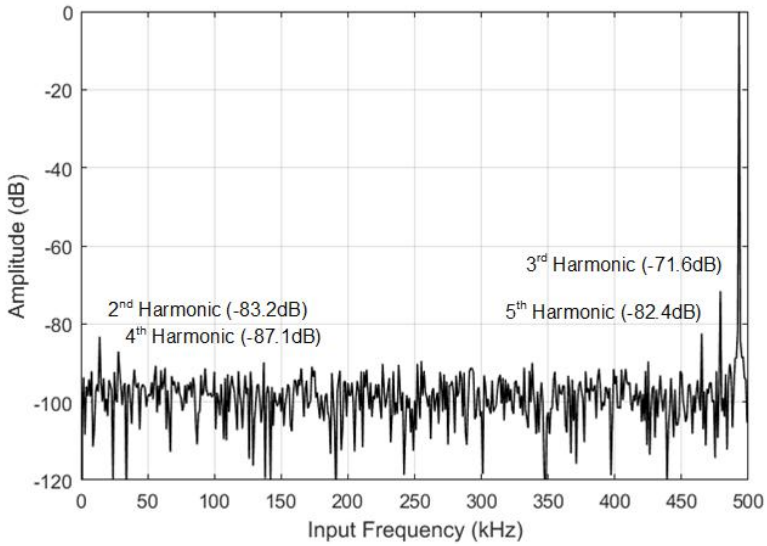


Figure 2.40: Measured FFT spectrum at 1MS/s with an input tone of 493kHz. The SFDR is 71.6dB and the SNDR is 54.6dB. The ENOB at 493kHz is 8.77.

The spurious-free dynamic range (SFDR) achieved in these tests was 76.54dB and 71.6dB, respectively. The signal-to-noise and distortion ratio (SNDR) was measured

to be 56dB and 54.6dB, respectively. The effective number of bit (ENOB) is defined as:

$$ENOB = \frac{SNDR - 1.76}{6.02} \tag{2.44}$$

The ENOB of the designed ADC was measured to be 9.01 and 8.77, respectively. The figure-of-merit is calculated using:

$$FoM = \frac{Power}{2^{ENOB} \times f_s} \tag{2.45}$$

The FoM of the ADC is 98fJ/conv-step at 1MSps with a supply of 1.8V. The specifications of the ADC was summarized in Table 2.7.

Table 2.7: Specification Summary of the 10-bit SAR ADC

Specification	**Measurement Result**
Technology	180 nm
Supply Voltage	1.8 V
Input Range	3 Vp-p
Sampling Rate	1 MSps
Active Area	0.042 mm2
INL	-0.82/+0.77 LSB
DNL	-0.49/+0.56LSB
SNDR	54.6 dB
SFDR	71.6 dB
ENOB	8.77
FoM	98fJ/conv-step

In summary, a 10-bit SAR ADC was presented in this section. A prototype was fabricated in 180nm CMOS technology. The design uses an energy efficient switch procedure and a split-capacitor array. The measurement results successfully meet the design specifications, with comparable performance among the state-of-the-art ADC designs for the neural recording purpose. As a part of the neural interface system,

the power consumption of the ADC was usually not the bottleneck. So this work didn't seek to aggressively minimize the ADC's power using techniques like charge recycling [129], asynchronous timing [132], or step charging [130]. The supply voltage was kept at 1.8V to be compatible with the neural recording front-end. In the future, the ADC's performance can be further optimized based on the characteristics of the neural signal.

2.5 A Compressed Sensing Neural Signal Acquisition System

2.5.1 Introduction

Wireless telemetry is usually the power bottleneck of a neural recording system [142]. On-chip data compression is an effective solution to reduce the power consumption reducing the data rate. Various on-chip data compression techniques for neural signal have been proposed. For single or multi-units action potential recording, spike detection [143], and spike sorting [144] are the most effective ways to reduce the recording data rate, and can also be used to drive the BMIs directly. The hardware implementation of the spike detection can be as simple as a comparator with a pre-defined threshold. A compression ratio higher than 100x can be achieved with little power consumption [67]. However, the spike detection based compression drops most of the raw waveform, and is vulnerable in long-time recording since the spike waveform may change due to the change of electrode impedance or electrode displacement. For EEG, ECoG, or LFP, wavelet transformation is an effective solution, given its high compression ratio and good reconstruction quality [145, 146]. However, the hardware implementation of wavelet transformation is non-trivial and usually takes considerable area and power. Moreover, the custom design for a specific signal type and sampling frequency significantly limits the applications of these recording systems.

Compressed sensing is an emerging signal processing technique that enables sub-Nyquist sampling and near lossless reconstruction of a signal [147]. Since it was introduced in 2006 [148], the compressed sensing technique has also been successfully applied to rapid MRI [148], computational image sensors [149], biomedical sensors

[142, 150], high frequency receivers [151], and other applications. Compressed sensing is especially attractive to neural signal recording given its minimum hardware cost in the front-end, favoring the power constraint of implanted devices.

Prior research shows the sparsity of neural signals in different frequency bands [150, 152–154]. Since an on-chip transformation using random matrix usually achieves sufficient incoherence and restricted isometry property (RIP) [155], a general-purpose recording device can be designed without the knowledge of the target signal. In addition, the compressed sensing measurements can also be used in signal processing (e.g. machine learning classifiers) [156], or driving BMI directly. Without a full reconstruction of the raw signal, the processing in the compressed domain can be easily implemented in a low-power embedded system.

Fig. 2.41 shows a survey of publications in data compression of neural signals. Compressed sensing shows a fast growth trend, and plays an increasingly important

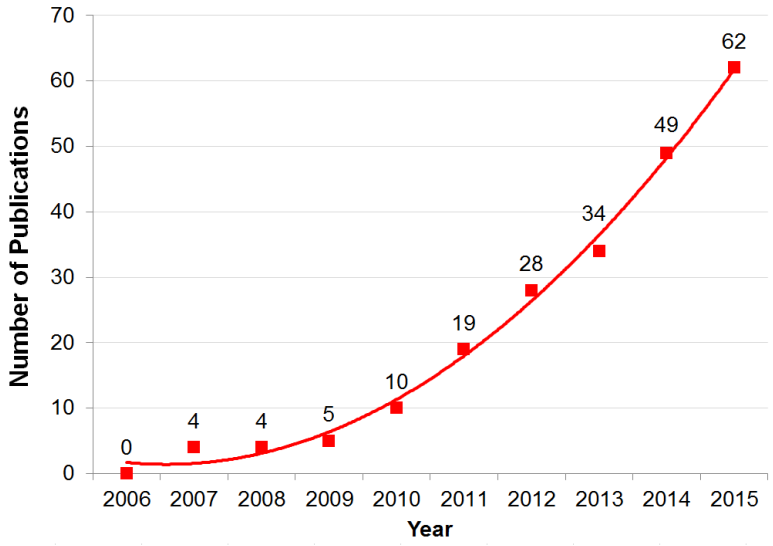


Figure 2.41: Historical trend for publications using compressed sensing technique in biomedical signal acquisition in the past decade. Data retrieved from Web of Science.

role in the hardware design of neural signal acquisition systems.

2.5.2 A Brief Background of Compressed Sensing

Compressive sensing (CS) is a signal processing technique that enables sub-Nyquist sampling and near lossless reconstruction of a signal with sparsity in a certain domain. The technique is particularly appealing for low-power high channel count neural signal recording. This section gives a brief introduction to the compressive sensing theory. Detailed explanation and rigid mathematical proof can be found in [147, 148, 155].

A. Compression Process

Assume the digitized signal $\mathbf{x}$ has a dimension of N, demoted by $\mathbf{x} \in \mathbb{R}^{N\times 1}$. Consider a general linear measurement process that computes $\mathbf{y}$ with a full row-rank matrix denoted by $\Phi \in \mathbb{R}^{M\times N}$, and $M \ll N$

$$\mathbf{y} = \Phi\mathbf{x} \tag{2.46}$$

where $\mathbf{y}$ is the compressive sensing data, and Φ is the sensing matrix. Notice that the sensing matrix is fixed and known to the reconstruction algorithm. The signal x can be expressed as:

$$\mathbf{x} = \sum_{i=1}^{N} s_i \Psi_i \tag{2.47}$$

where $\mathbf{s}$ is the representation of the signal in the Ψ domain. The signal $\mathbf{x}$ is $K-Sparse$ if only K of the $\mathbf{s}$ coefficients are non-zero. The signal is compressible in this case. Thus $\mathbf{y}$ can be written as

$$\mathbf{y} = \Phi\Psi\mathbf{s} \tag{2.48}$$

B. Reconstruction Process

The signal reconstruction process is to use the $\mathbf{M}$ measurements in $\mathbf{y}$, the measurement matrix Φ, and the basis Ψ to reconstruct the signal $\mathbf{x}$, or equivalently, its sparse representation $\mathbf{s}$. Since $M \ll N$, the equation is underdetermined, which means there are infinite $\mathbf{x}$ (or $\mathbf{s}$) that satisfy the condition. Therefore, the signal reconstruction process is to find out the signal sparse coefficient vector.

The classical approach is to find the vector in the translated null space with the smallest ℓ_2 norm by solving

$$\hat{\mathbf{s}} = argmin||\mathbf{s}'||_2 \quad such\ that \quad \Phi\Psi\mathbf{s}' = \mathbf{y} \tag{2.49}$$

However, the ℓ_2 minimization usually has difficulty in finding a K-Sparse solution. ℓ_0 norm can recover a K-Sparse signal exactly with high probability.

$$\hat{\mathbf{s}} = argmin||\mathbf{s}'||_0 \quad such\ that \quad \Phi\Psi\mathbf{s}' = \mathbf{y} \tag{2.50}$$

Unfortunately, solving Eq. 2.50 is both numerically unstable and NP-complete. While ℓ_1 norm can exactly recover K-Sparse signal and closely approximate the signal with high probability.

$$\hat{\mathbf{s}} = argmin||\mathbf{s}'||_1 \quad such\ that \quad \Phi\Psi\mathbf{s}' = \mathbf{y} \tag{2.51}$$

This is a convex optimization problem and can be conveniently reduced to a basis pursuit problem, with a computational complexity about $O(N_3)$.

C. Reconstruction Evaluation Criteria

Several numerical derivations are used to evaluate the performance of individual reconstruction algorithms and dictionaries. The commonly used criteria include compression ratio and signal-to-noise and distortion ratio.

The Compression Ratio (CR) is defined as:

$$CR = \frac{N}{M} \tag{2.52}$$

The signal-to-noise and distortion ratio (SNDR) is defined as:

$$SNDR = 20 \times \log \frac{||\mathbf{x}||_2}{||\mathbf{x} - \hat{\mathbf{x}}||_2}.$$

2.5.3 System Overview

The paradigm of the hypothetical chronic wireless neural signal acquisition system is illustrated in Fig. 2.42. The system has a dedicated implantable subsystem and a flexible external subsystem. The implantable subsystem contains the proposed compressed sensing neural recording SoC, an inductive charging module, and a super capacitor. The device will need to be sealed in a biocompatible package. The device can be placed under the skin, above the skull bone. The recording electrode can be placed in any brain area of interest. The external subsystem consists of a standard wireless transceiver, a rechargeable battery, and a coil. The external subsystem powers the implanted device and collects data back through back-scattering.

The advantages of the proposed system are three-folds: i) the implanted wireless device leaves the skin intact, which reduces the risk of infection, ii) the battery is left externally so that the device's lifetime will not be limited by the battery's recharging

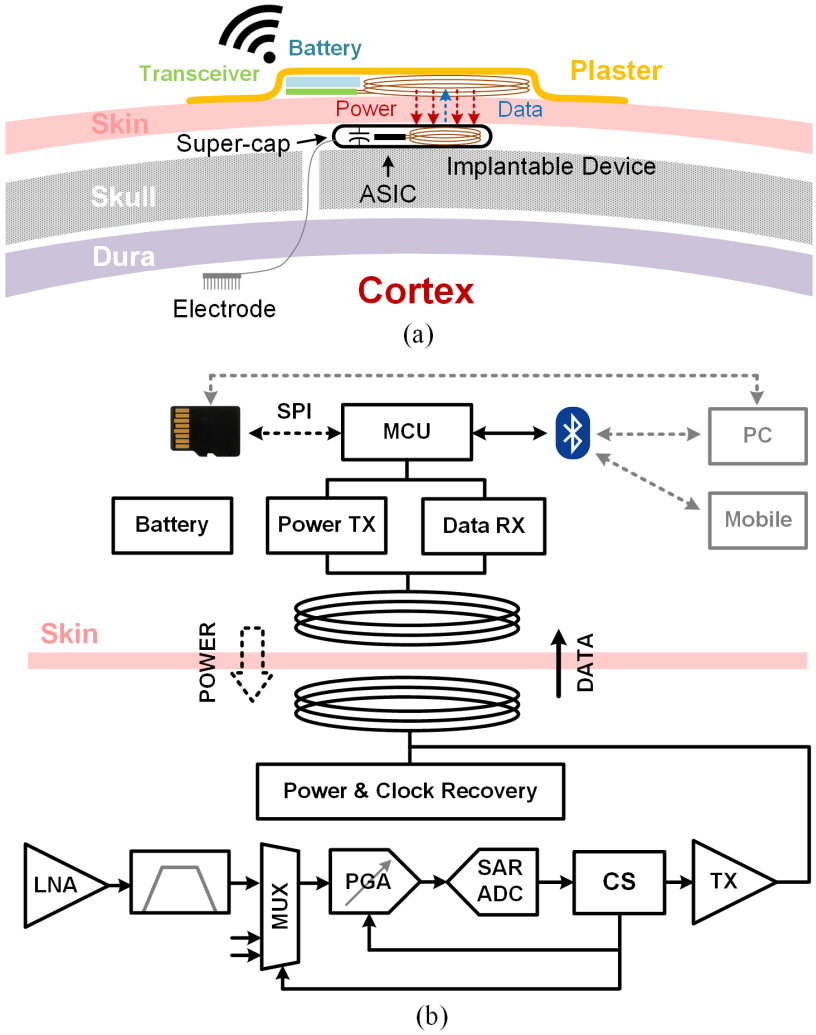


Figure 2.42: (a) Illustration of the hypothetical chronic neural signal recording system using the fully integrated compressed sensing chip, and (b) the architecture of the chip.

cycles, and the toxicity associated with batteries will not be a potential danger to the subject, iii) the external transceiver makes the system flexible and versatile, for instance, different wireless solutions or flash memory can be used for different situations. The upgrading of the system is also much easier, since the chronic implant can be used for years or even decades while the digital and wireless electronics develop

much faster than the analog recording interface.

A single pair of coils is used for both power delivery and data read back. A carrier frequency of 13.56MHz is chosen given the trade-off between the power transfer efficiency and the data rate. Compressed sensing reduces the data rate of the wireless uplink, which is especially helpful for the multiple channel recordings.

2.5.4 Circuit Implementation

2.5.4.1 Energy Efficient Analog Front-end

The block diagram of each analog recording channel is shown in Fig. 2.43. A fully differential low-noise instrumentation amplifier (IA) is used to amplify the neural signal. The following Gm-C based high pass filter stage (HPF) conditions the signal with a tunable cut-off frequency. The next stage (LFP) is an operational transconductance amplifier (OTA) that converts the voltage signal into a current in a programmable low-pass frequency corner.

The IA in this work is a fully differential capacitor-coupled neural amplifier, which amplifies the weak neural signal in a wide frequency band. The input capacitors block the large electrode offset and half-cell potential from the interface, giving a maximum input range. The closed-loop differential gain is set to be 34dB to relieve the noise requirement for the following stages. The core of the IA is a low-noise OTA, as shown in Fig. 2.43 (a-1). The OTA has been designed to maximize the noise and power efficiency. Compared with the design presented in Section 2.2.2, a two-stage topology is used to provide a sufficient open-loop gain. The major consideration here is the use of the cascode topology limits the voltage headroom. A complementary input

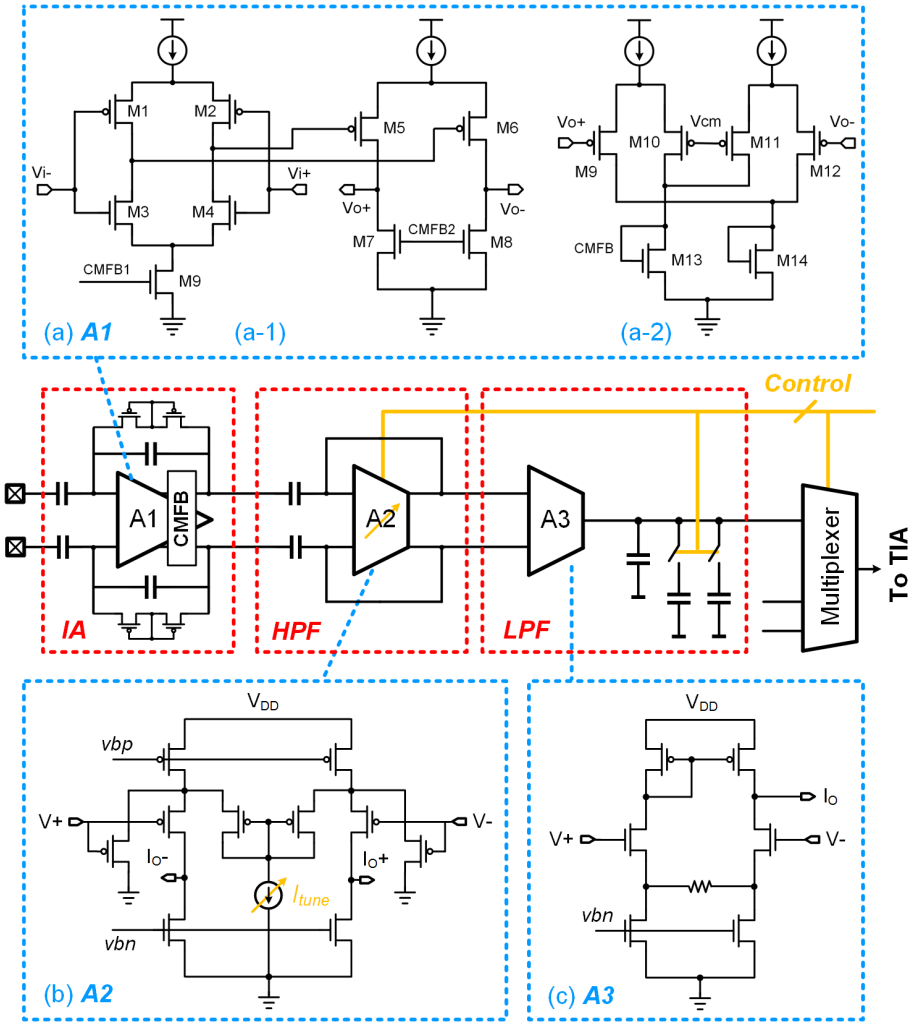


Figure 2.43: Analog front-end of the proposed system (Part I). The signal chain includes the signal amplification, filtering, voltage-to-current conversion and multiplexed to a shared ADC. The circuit schematic of the (a) low noise amplifier, (b) OTA with extended linear range, and (c) OTA with programmable transconductance.

stage (M1-M4) is used to increase the overall transconductance without increasing the quiescent current. The complementary input amplifier suffers from PVT variations [116], thus additional common-mode feedback circuit, as shown in Fig. 2.43 (a-2), is adopted to stabilize the DC output at half supply voltage. All of the input transistors are biased in the sub-threshold region to achieve a high energy efficiency. Since the

complementary stage has a limited input range, a fully differential structure is chosen. The first stage dominates the noise, and the input-referred noise of the OTA can be expressed as:

$$\overline{v_{i,n,tot}^2} = \frac{1}{(g_{m1}+g_{m3})^2}[8KT\gamma(g_{m1}+g_{m3})+ \\ 2(\frac{K_N g_{m3}}{C_{ox,N} f W_N L_N} + \frac{K_P g_{m1}}{C_{ox,P} f W_P L_P})]\Delta f \tag{2.53}$$

where g_{m1} ($=g_{m2}$) are the transconductance of M1 (M2), and g_{m3} ($=g_{m4}$) are the transconductance of M3 (M4). The flicker noise can be reduced by increasing the widths and lengths of the input transistors. A biasing current of 1μA is used in the first stage as a trade-off between power and noise. A biasing current of 20nA is used in the second stage. The dominant pole is set at the second stage, and the stability is guaranteed by adding an additional capacitive load.

An ultra low-power programmable bandpass filter is integrated into each channel for selecting the frequency band of interest. The first stage is a fully-differential Gm-C highpass filter. The circuit schematic of the Gm block is shown as A2 in Fig. 2.43 (b). Current division and local feedback are used to achieve low transconductance and an extended linear input range. The cut-off frequency can be programmed by tuning the transconductance. The second stage of the filter is a single-ended Gm-C based lowpass filter. The circuit schematic of the Gm block is shown as A3 in Fig. 2.43 (c). Source degeneration is used to achieve a high linearity. The differential voltage signal is converted into a single-end current signal. Since a standard current mirror load is used, no extra power is wasted for this conversion, but the single-ended operation reduces the capacitor array size by half, which is important for this design to be implemented at the channel level. The lowpass frequency can be programmed

by selecting the load capacitor.

The shared part of the analog recording front-end is shown in Fig. 2.44. The

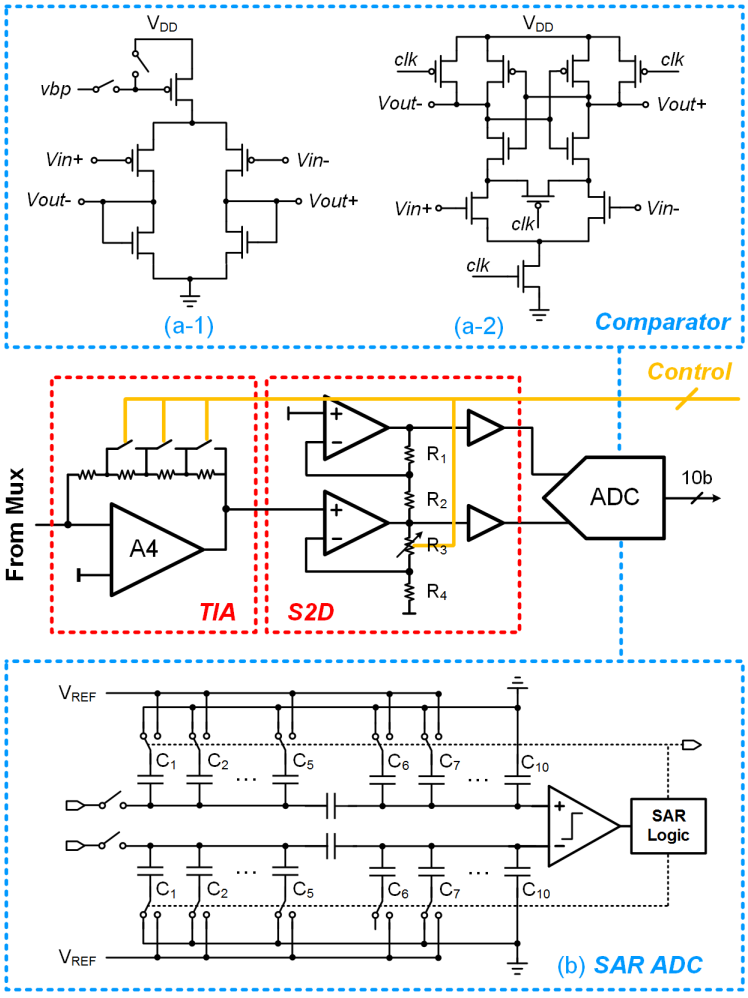


Figure 2.44: Analog front-end of the proposed system (Part II). A current-to-voltage conversion with programmable gain and a 10-bit SAR ADC is used to digitize the signal. The boxed windows show the circuit schematic of the (a) comparator and the (b) SAR ADC.

current output from each channel is multiplexed and then converted to a voltage using a transimpedance amplifier (TIA) with a programmable gain. A single-to-differential

(S2D) converter is used to drive the differential input ADC with an additional programmable transimpedance. A 10-bit SAR ADC digitizes the signal. The design details of the ADC have been presented in Section 2.4.

The single-ended current output from the 16 channel is selected by a multiplexer. The single-ended signal reduces the effort in routing, and the R-I drop in the long routing line doesn't corrupt the current signal, thus making it less susceptible to noise. The following TIA is used to convert the current signal back to a voltage in a programmable gain, as shown in Fig. 2.44. The gain can be set to be 5x, 6x, 7x, 8x by the compressed sensing digital processor. The gain of 2x, 4x, can be easily achieved in the binary digital processor, and the 3x can be achieved from shifting the 6x signal by 1 bit.

2.5.4.2 On-chip Wireless Power and Data Link

A low-power backscatter based wireless transmitter communicates with the external transceiver [11]. The backscatter transmitter consists of a PWM encoder and a buffered transistor for the antenna impedance modulation.

An active rectifier is used to achieve a higher power efficiency [157]. Coupling coils are implemented off-chip. The system clock is recovered from the power waveform [100]. The circuitry of the clock recovery and division module is shown in Fig. 2.45. The module consists of a Schmitt trigger and several D flip-flops. The Schmitt trigger makes the circuit more resistent to the noise of the power waveform. The circuit schematic of the Schmitt trigger is shown in Fig. 2.46. The D flip-flop guarantees the clock has a 50% duty cycle. Several different clocks can be divided from the following D flip-flops. The clock frequency selection can be configured by a shift register.

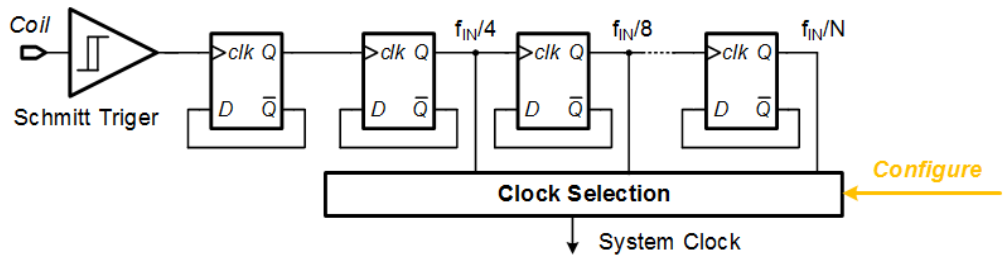


Figure 2.45: The circuit schematic of the clock recovery and clock division module.

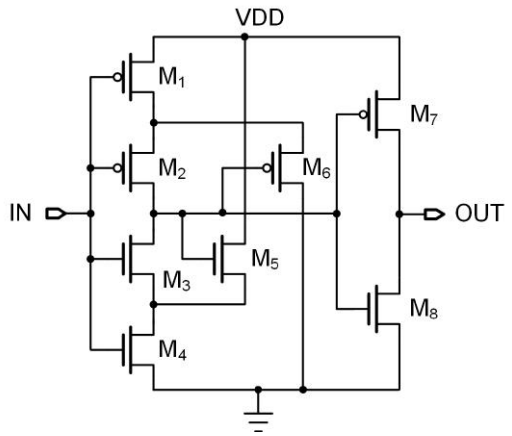


Figure 2.46: The circuit schematic of the CMOS Schmitt trigger [19].

Standard bandgap reference and low drop-out (LDO) circuits are used in the power management unit. The block diagram and the circuit schematics of the power management module are shown in Fig. 2.47. A push-pull comparator with source input is used to drive the active diodes. The design details of the active rectifier can be found in references [157–159].

2.5.4.3 External Wireless Relay Board

An external wireless relay board has also been designed to demonstrate the proposed paradigm. The external subsystem consists of a microcontroller with an integrated

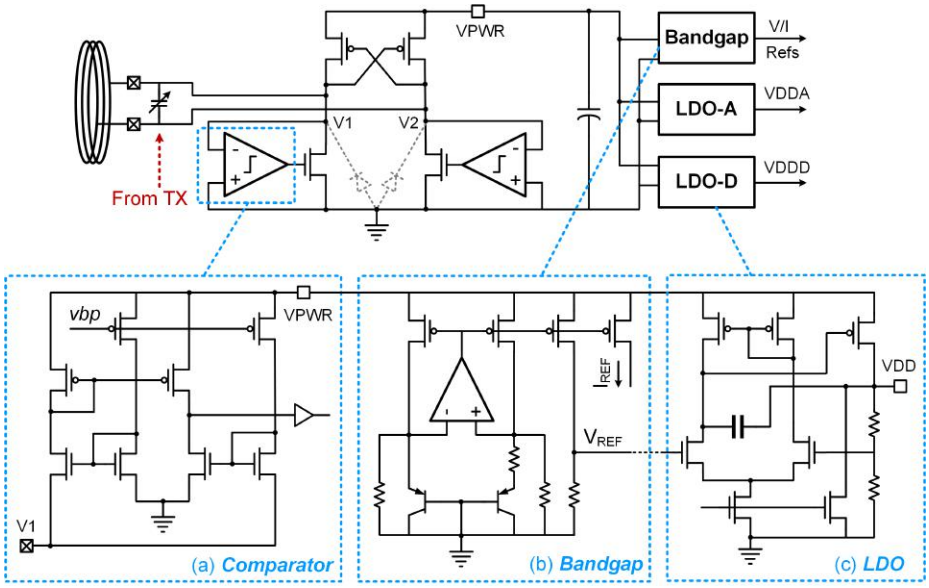


Figure 2.47: Inductive power management module, including active rectifier and LDOs for analog and digital power supplies. (a) circuit schematic of the comparator, (b) bandgap reference, and (c) LDO (start-up circuits are not shown).

wireless transceiver, envelope detection circuits for reading the backscattered signal, power transmitter circuits, and a battery management system.

A 32-bit ARM Cortex-M0 based wireless transceiver (Nordic Semiconductor nRF51822) is used as the central processor and wireless transceiver. It features a 2.4GHz transceiver, and supports Bluetooth 4.0 low-energy protocol, which provides an easy interface to the computer or mobile devices. A reliable wireless communication up to 5m was measured in the normal indoor environment. A Serial Peripheral Interface (SPI) based microSD card interface is optional in the system to allow long-term wireless recording without limited receiver range.

A computer user interface has been developed in Matlab to configure the device and read back the data. The signal conditioning and off-line analyses are also performed in the user interface.

2.5.5 Measurement Results

The proposed SoC design has been fabricated in an IBM 180nm standard CMOS technology, occupying a silicon area of 2.1mm×0.8mm, excluding the IO pads. A microphotograph of the fabricated chip is shown in Fig. 2.48, with major building blocks highlighted.

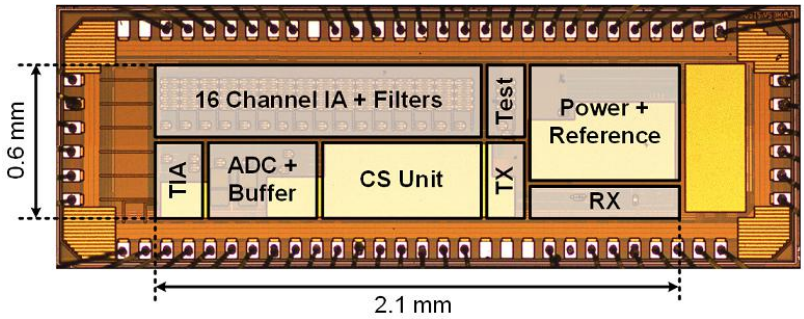


Figure 2.48: The micrograph of the fabricated fully integrated compressed sensing neural recording front-end chip.

Bench testing was conducted to verify the functions of the chip and the system. The measured frequency response of the low-noise amplifier is shown in Fig. 2.49. The frequency response was measured point by point using a function generator 33521A and an oscilloscope MSO7034B from Agilent. The phase shift was calculated in the oscilloscope. The measured midband gain is 34.1dB. The measured CMRR and PSRR of the analog front-end in the frequency range of 0.5Hz to 7kHz are >80dB, and >67dB, respectively.

The input-referred noise spectrum is shown in Fig. 2.50. The noise was measured with the inputs shorted by an internal switch. The noise spectrum density is $200nV/\sqrt{Hz}$ at 10Hz, $49.1nV/\sqrt{Hz}$ at 100Hz, and $23nV/\sqrt{Hz}$ at 1kHz. An integration under this curve from 1Hz to 7kHz yields a rms noise floor of 2.85μV. The

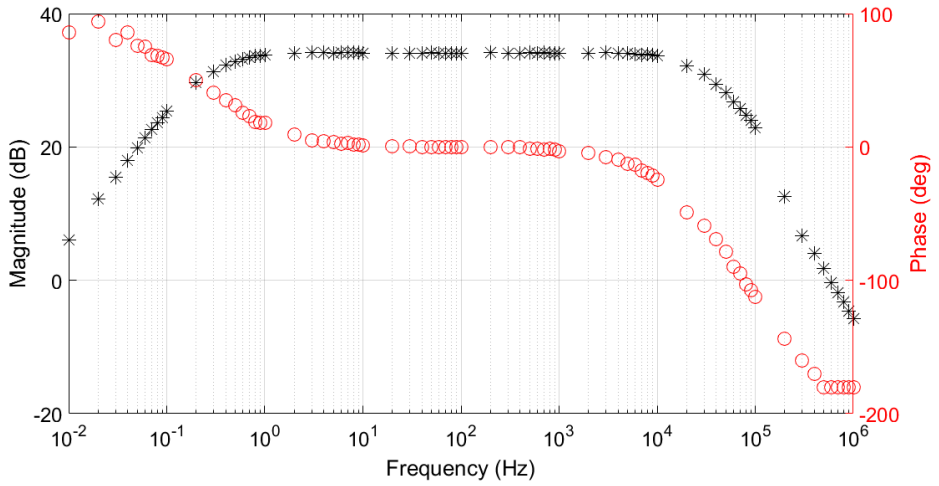


Figure 2.49: The measured frequency response of the low noise amplifier (without filtering stages).

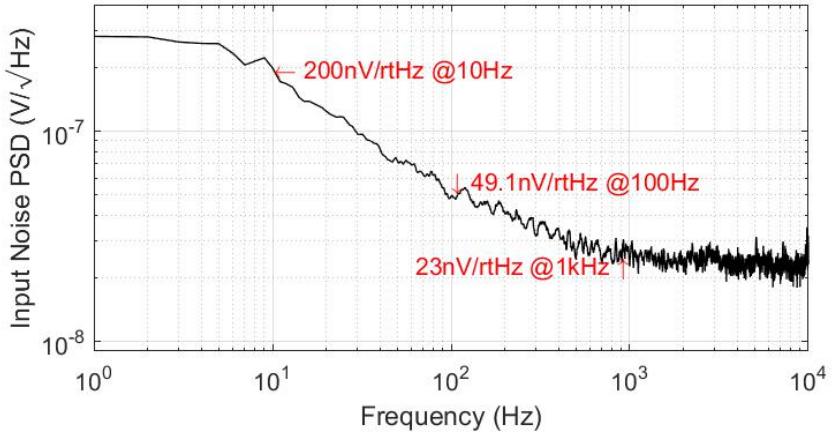


Figure 2.50: The measured input-referred voltage noise spectrum.

total harmonic distortion of the amplifier was measured to be -63dB, with an input amplitude of 1mV.

An invasive neural recording was performed in an anesthetized rat with a tungsten microelectrode placed in its motor cortex. Action potential data is extracted by configuring the filter with a passband of 300Hz to 7kHz. Different compression ratios

from 2, to 4, to 8 and to 16 have been applied, respectively. Dual-threshold level crossing spike detection has been used for both the uncompressed data and the restored data. Signal-to-noise distortion ratio (SNDR) of 3.60dB, 9.78dB, 30.60dB and 52.99dB are achieved for compression ratios 16, 8, 4 and 2, respectively. Near-lossless spike detection can be achieved while a compression ratio lower than 8 is applied.

Fig. 2.51 compares the time-domain waveform of the uncompressed and restored local field potential (LFP) sampling data sets. And Fig. 2.52 shows the comparison of

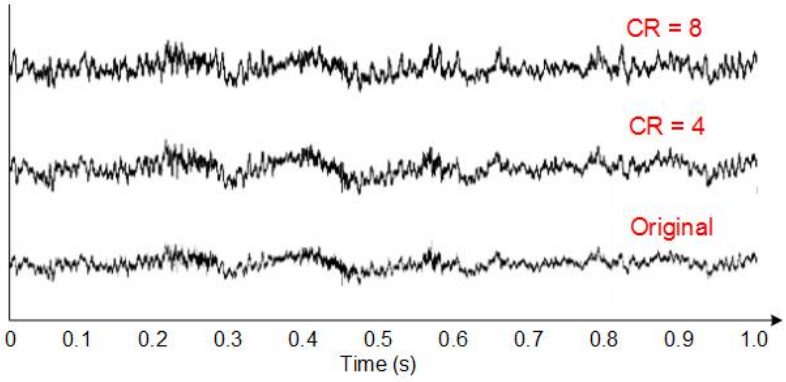


Figure 2.51: Comparison between the uncompressed sampling results and data restored from different compression ratio (CR).

the spectrum of the original uncompressed and restored LFP sampling data sets. The LFP exhibited rhythmic bouts of broadband power interleaved with low power epochs. According to Fig. 2.52, the time-frequency content of the restored signal was very similar to the uncompressed LFP. Signal-to-noise distortion ratio (SNDR) of 9.04dB, 4.85dB and 3.78dB are achieved for compression ratios 4, 8 and 16, respectively.

A demonstration system was developed to show the proposed concept, as shown in Fig. 2.53. An open cavity plastic package was used for packaging the chip, thus the size of the demonstration implantable system was limited by the package. Commercial coils were used for power and data transfer. An additional ceramic capacitor was used

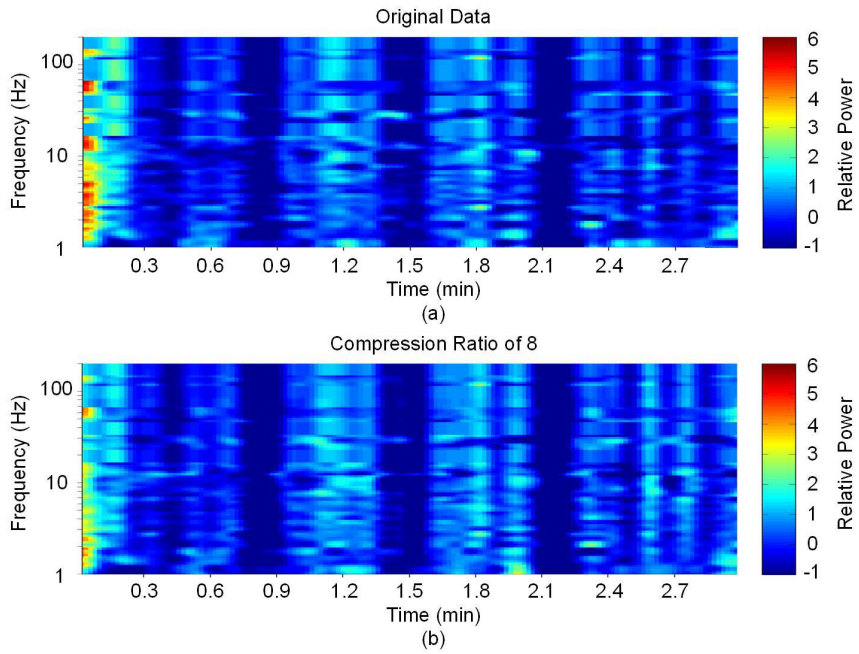


Figure 2.52: Comparison of the spectrograms of (a) the uncompressed sampling results and (b) the data restored with a CR of 8.

to improve impedance matching. Two LEDs were used only for debugging purpose. A couple of programming and debugging pads are left. No other off-chip components were required.

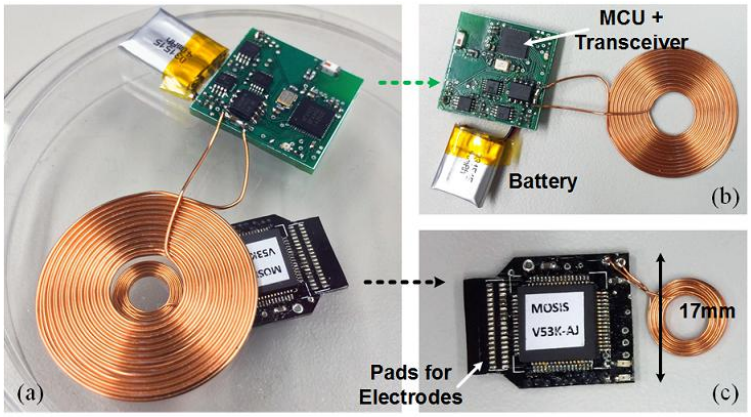


Figure 2.53: Photography of an assembled demonstration system. (a) Power and data transmission testing setup across 5mm plastic cap, (b) external transceiver board, (c) implantable device.

In-Vivo evaluation of the device for the long-term operation was conducted in a rhesus macaque. An electrode was chronically implanted in the hippocampus. The recording device, including an external transceiver, was housed in a small chamber that was fixed to the skull. Fig. 2.54 shows the spectrogram of a 24-hour continuous recording while the monkey was freely behaving in his home cage. The recording

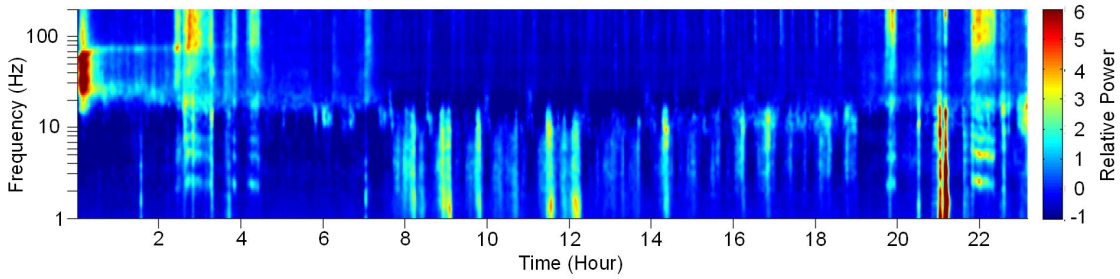


Figure 2.54: A 24-hour continuous recording in the hippocampus of a rhesus macaque during free behavior.

shows the states of hippocampal activity throughout the day. Greater power at higher frequencies (>20 Hz) was associated with periods in which the animal was awake and freely moving about his home cage (hours 0-7.5 and 19-24). Greater power at low frequencies (<20 Hz) was associated with sleeping (hours 7.5-19). Individual sleep cycles can be seen. Some broadband chewing artifacts were also present (around hours 3-4.5 and 20-22) corresponding to the times when the animal was fed. The overall activity pattern matches previous observations of sleep-wake changes in neural activity. The measured performance of the chip is summarized in Table 2.8.

In this work, a fully integrated wireless neural signal acquisition system is presented. A high efficiency wireless neural signal recording SoC with integrated compressed sensing processor was designed and fabricated in 180nm CMOS technology. An external wireless relay was used to power the implantable SoC, read back the data through backscattering, and transmit the data through a universal wireless link.

Table 2.8: Chip Specifications Summary

Neural Amplifier		**CS Processor**	
Midband Gain	34.1dB	Input Channel	up to 16
Bandwidth	0.5Hz - 7kHz	CS Ratio	up to 8x
LNA Noise	2.85μVrms	Clock freq.	4MHz
THD (1mV)	-63 dB	**Wireless Power and Data**	
NEF/PEF	1.58/4.5	Carrier freq.	13.56 MHz
CMRR	>80dB	Power efficiency	up to 73%
PSRR	>67dB	Distance	up to 10mm
SAR ADC		**Power**	
ENOB	9.1	Analog Front-end	2.5μW (per ch.)
Sampling Rate	1MSps	ADC	35μW(@1MSps)
INL (LSB)	+0.62/-0.85	CS Processor	77μW
DNL (LSB)	+0.69/-0.92	TX transmitter	27μW
FoM(fJ/step)	34.2	Total (avg.)	254μW

The system features high energy efficiency, high flexibility, compatibility, upgradability without compromising the signal recording quality. By performing the on-chip compressive sampling, the data rate is significantly reduced, which allows the system to support more recording channels without a power penalty. According to the experimental results, a compression ratio up to 8x will cause negligible loss of the data quality and/or information contained in the raw data. A pre-implantable system was assembled and successfully demonstrated the proposed paradigm. Bench testing and *In-Vivo* experimental results are presented. Table 2.9 compares the performance of the proposed work with prior published compressed neural signal recording front-end designs. The system shows a promising chronic neural signal recording paradigm for neuroscience research and BMI applications.

Table 2.9: Comparison with State-of-the-art Works

Reference	Deepu [112]	Gangopadhyay [160]	Zhang [161]	Biederman [72]	This work
Publication	2014 JSSC	2014 JSSC	2015 JNE	2015 JSSC	-
CMOS technology	0.35um	130nm	180nm	65nm	180nm
No. of channels	4	64	4	64	16
Signal type	ECG	ECG	Extracellular	Extracellular	LFP/Extracellular
Input-referred noise	1.46uV	<2uV	3.1uV	7.5uV	2.8uV
Sampling rate/ch	256/512Hz	2kHz	20kHz	20kHz	20kHz
Front-end NEF/PEF	3.31/26.3	-	-	3.6/12.9	1.58/4.5
ADC ENOB	9.3	6.5	-	8.2	9.1
AFE+ADC power/ch	0.54uW (512Hz)	28nW (2kSps)	15uW (20kSps)	1.84uW (20kSps)	3.2uW(20kSps)
Compression method	Lin slope predict.	CS	CS	Spike Dect.	CS
Compression ratio	2.55x	up to 6x	8x - 16x	8.3x (epochs)	8x - 16x
Reconstruction SNR	-	-	<9dB (16x)	-	9.78dB (8x)
Wireless	-	-	-	-	Backscattering
In-vivo experiment	Yes	-	Yes	Yes	Yes
System integration	-	-	-	Complete	Complete

Chapter 3

Neural Feature Extraction

3.1 Introduction

Feature extraction, or feature learning, is an important technique to transform the raw data input to a representation that can be effectively understood [162]. Neural feature extraction allows one to acquire the qualitative and quantitative information from the neural signal. It has been widely used in the neuroscience and neuroprosthetic research for pattern recognition, numerical or symbolic regression, probability estimation, and dynamical system modeling [163]. Moreover, neural feature extraction provides the inputs to the decision support system, like the machine learning core in a brain-machine interface (BMI) device. The real-time neural feature extraction implemented on-chip is especially important for the operation of the closed-loop BMI devices. By applying feature extraction and machine learning techniques, the BMI devices have been successfully used in decoding motor function [164, 165], detecting epilepsy [113, 166], Parkinson's disease [167], depression [168], and so on.

The feature extraction can be performed in different domains, including: i) time domain, ii) frequency domain, iii) wavelet domain, iv) statistics process, v) information theory (eg. entropy, mutual information), and vi) fractal geometry [163]. However, the implementation of the real-time feature extraction in the BMI devices is limited by the hardware resource including the computation ability, the memory size, and the power consumption. Thus, an energy efficient implementation is especially important. Some energy-efficient neural feature extraction techniques have been reported in the literature [113, 144, 169–172]. It should be noticed that the choice of a suitable set of features is also a challenging task. The brain signal contains a large number of simultaneous sources, and the information of interest might be overlapped with other sources time and frequency. There are many existing methods for feature selection, including principal component analysis (PCA) [173], independent component analysis (ICA) [174], genetic algorithm (GA) [175], sequential forward/backward selection (SFS) [176], and so on.

Although the use of neural features varies significantly for different applications, it is very helpful to identify the most commonly used features for the BMI devices. Both field potential and action potential features have been used in real-time, closed-loop BMI devices, as briefly summarized below:

- **Field potential features**
  - **Energy in multiple frequency bands:** spectral characteristics of neural field potentials have been used to identify, classify, and analyze brain activities [169, 177];
  - **Features of different brain states:** field potentials can be used to define different brain states [178, 179];

  - **Synchronization between electrodes:** synchronization of oscillations between different brain areas can be used to define different brain activities [180, 181].

- **Action potential features**
  - **Action potential detection:** action potentials are activities of individual neurons [182, 183];
  - **Action potential alignment and sorting:** on-line action potential sorting can be used to identify the signal from different neurons presented on the same electrode [182, 184];
  - **Action potential firing rate:** the action potential fire-rate presents the active level of an individual neuron [185, 186].

This chapter presents the analysis and design of the neural feature extraction from three different perspectives, the energy extraction, the action potential detection, and the matched filtering. Several novel techniques in the circuitry, algorithm, and system levels are proposed, with a focus on the energy efficient implementation for closed-loop BMI devices. The chapter is organized as follows. Section 3.1 introduces the neural features, and summarizes the commonly used features for real-time BMI devices. Section 3.2 describes the energy features in the LFP, and proposes a novel extraction circuit with frequency tuning in the natural logarithmic domain. Section 3.3 analyzes the real-time action potential detection and classification units, followed by the design of a low-power current mode action potential unit. Section 3.4 analyzes the matched filter with pre-whitening and its application to the phase-amplitude coupled neural feature extraction.

3.2 Natural Logarithmic Domain Neural Energy Extraction

3.2.1 Introduction

A substantial amount of information regarding motor intent can be inferred from the field potential recordings [187, 188]. The field potentials, either recorded with electrodes penetrating the brain or on the brain surface, reflect the summed activity of thousands to millions of neurons. Oscillations are particularly prominent in field potential recordings and reflect synchronous, rhythmic changes in the activity across the network. The recorded oscillations contain information correlated with a number of different behavioral processes, i.e., motor planning [189]. While decoding the intent from the field potentials for a neuroprosthetic application, it is typical to extract energy from several discrete frequency bands [169].

A variety of distinct brain oscillations exist, with center frequencies spaced logarithmically [190], as illustrated in Fig. 3.1. Commonly used frequency bands include: delta band (1-4Hz), theta band (4-10Hz), beta band (10-30Hz), gamma band (30-80Hz), and fast band (30-80Hz). Neural oscillations associated with a certain cognitive state can be in a very narrow frequency band, especially in the low-frequency range. For example, the first discovered and the best-known frequency band is the alpha activity which is 7.5-12.5Hz. This places a big design challenge in the neural energy extraction unit. If the frequency programming uses a linear step, a high-frequency resolution will have to be realized. Similarly, if a Fast Fourier transform (FFT) analysis is used, that requires a large number of FFT points and large memory size for buffering the data. In order to address this problem, a natural logarithmic

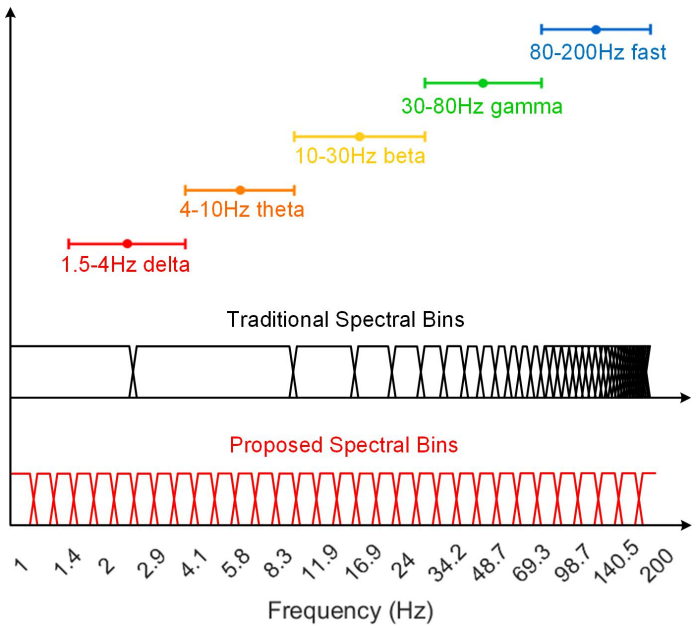


Figure 3.1: Brain oscillation bands are shown together with the traditional frequency tuning bins in linear steps, and the proposed tuning bins in the natural logarithmic domain. A total of 32 steps in a frequency range from 1Hz to 200Hz is shown for illustration.

domain tuning is proposed in this work, which provides sufficient resolution for extracting the low-frequency brain oscillations, without increasing the number of tuning steps. Fig. 3.1 compares the frequency bins for the conventional linear step filter and the proposed natural logarithmic domain filter when they have the same number of bins.

3.2.2 System and Circuits Implementation

The processing flow of the LFP energy extraction is shown in Fig. 3.2. A lowpass filter with a frequency corner of 300Hz is first used to remove the high-frequency components in the signal. Then, two 2nd-order stagger-tuned biquad filters are cascaded

Figure 3.2: The processing flow of the LFP energy extraction.

to bandpass the neural signal with a programmable center frequency and quality factor [170]. The filtered signal is then squared in a Gilbert multiplier to calculate the energy. Finally, the energy integral is produced by a leaky integrator with a tunable time-constant [191].

A prototype system that consists of 16 neural feature extraction channels is designed in this work. Each channel can be programmed independently. The feature extraction module in each channel can also be combined as a filter bank to perform spectrum analysis for one channel. Fig. 3.3 illustrates the configuration of the system.

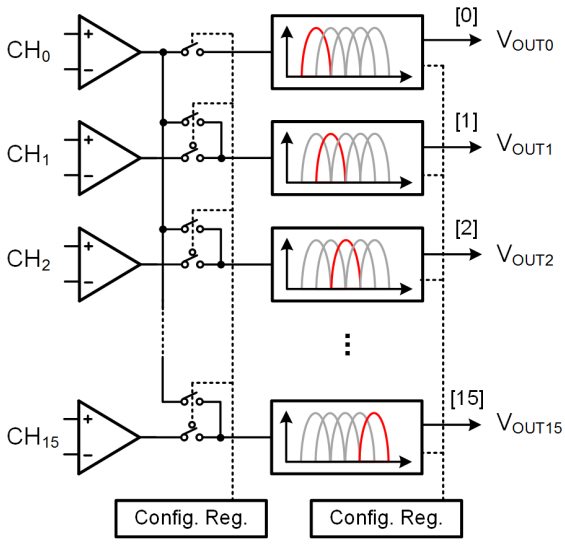


Figure 3.3: The diagram of the 16-channel energy extraction module.

3.2.2.1 Design of the GmC Filter

Given the low frequency nature of the neural signal, filters with very large time-constant have to be integrated on-chip. There are several methods to implement filters in CMOS circuits: i) The op-amp based filters can achieve a high linearity and a good signal-to-noise ratio (SNR), but suffers from a high power consumption, large passive components (non-linear if MOS resistors are used), and difficulties in tuning; ii) The switched capacitor filter can achieve a high linearity and a high frequency accuracy with a good tunability [169], but it has limitations and designs challenges in the tunable range, the capacitor size, the non-idealities from the clocks, and requires additional clock generation circuits; iii) The Gm-C filters can realize large time-constant in an ultra-low power consumption and in a very compact layout. As a result, these filters have been widely used in biomedical applications [170]. But Gm-C filters also have limitations in terms of linearity and the frequency corner accuracy. iv) The digital filters can achieve good filter characteristics, but requires a pre-digitization of the signal, a memory for buffering the data, and a dedicated DSP core [113]. A high-order digital filter requires accurate coefficients and a sufficient number of bits during the computation to prevent overflow. However, with the development of advanced CMOS technology, the digital filters may surpass the analog filters in both accuracy and power consumption. In this work, a Gm-C based filter is designed with a tunable transconductance in a range of two decades, with an extended linear range.

The circuit schematic of the first implemented Gm block is shown in Fig. 3.4. The input transistors are biased in the sub-threshold region [192]. The transconductance features a linear relation with the biasing current in the sub-threshold region [193],

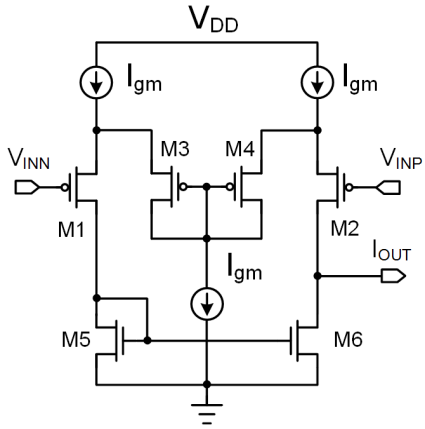


Figure 3.4: The circuit schematic of the transconductance (Gm) block with source degeneration.

as expressed:

$$g_m = \frac{I_{DS}}{\zeta U_T} \tag{3.1}$$

where ζ is a parameter that depends on the process, and $U_T = kT/q$. The transconductance of the Gm block can be directly tuned by the biasing current I_{DS}. A local feedback is used to reduce the distortion [121]. Instead of using resistors, diodded-connected transistors M_3 and M_4 are connected to the sources of the input transistors M_1 and M_2. The total resistance between the sources of the input transistors becomes $2/g_{m3}$. A reduction factor n is defined as:

$$n = 1 + \frac{g_{m1}}{g_{m3}} \tag{3.2}$$

The third-order harmonic distortion HD_3 is reduced by a factor of n^2 [194]. In this work, M_1 and M_3 are set to have the same dimension with the same biasing current to maximize the common mode input range [195].

Thick oxide transistors are used in this circuit for a lower transconductance and to reduce the leakage current. The current mirrors are biased in the strong inversion

region for a better matching. But there is a limitation in the minimum current that can be reliably copied in the circuit. To reduce the capacitor size in the Gm-C filter, the transconductance needs to be further reduced. So, a modified version of the Gm Block is implemented. The circuit schematic is shown in Fig. 3.5. The input

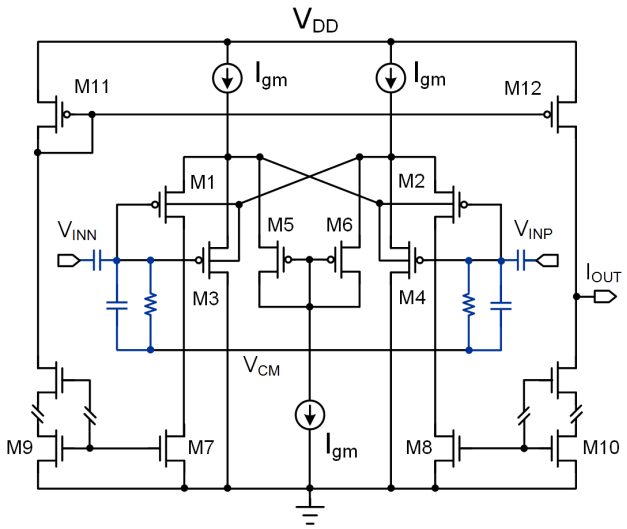


Figure 3.5: The circuit schematic of the transconductance (Gm) block with extended linear range. The biasing current is used to tune the transconductance.

transistors ($M1-6$) are biased in the sub-threshold region, and current division is used at the input differential pair to reduce the transconductance. Besides the local feedback as in Fig. 3.4, a bulk degeneration [196] is used to further enhance the linear input range. In a testing version, capacitor attenuation [197] is used to reduce the input signal swing and lower the overall transconductance.

3.2.2.2 Biasing Current Generation

In the first version, a linear 6-bit current-mode DAC is integrated for generating the biasing current for the Gm block. The circuit schematic is shown in Fig. 3.6. Thick

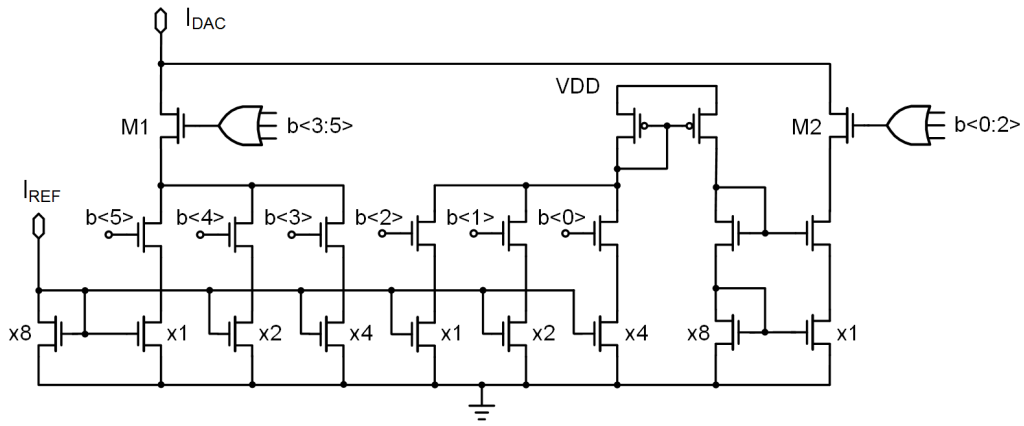


Figure 3.6: The circuit schematic of the 6-bit current-mode DAC for generating the biasing current for the Gm block. Thick oxide transistors are used to reduce the leakage current.

oxide transistors are used to reduce the leakage current. The DAC has two segments with binary weighted transistors. Two gating transistors $M1$ and $M2$ are used to further reduce the current leakage.

The second version of the biasing current generation module is designed for the proposed natural logarithmic domain tuning. Fig. 3.7 shows the programmable biasing current generation module. A two-step 6-bit resistor ladder based DAC is used to generate the 64-step linear tuning voltage between V_{cm} and V_{ref}. A custom designed current generation module ($M1$ to $M6$) converts the linear voltage to a natural exponential current. In the current generation module, all transistors $M1$ to $M6$ are biased in the sub-threshold region. When V_{DS} is higher than three or four times of the thermal voltage U_T, the sub-threshold transistor is in the saturation region [118]. The equation of the sub-threshold current can be simplified to:

$$I_D = I_o \exp \frac{V_{GS}}{\zeta U_T} \tag{3.3}$$

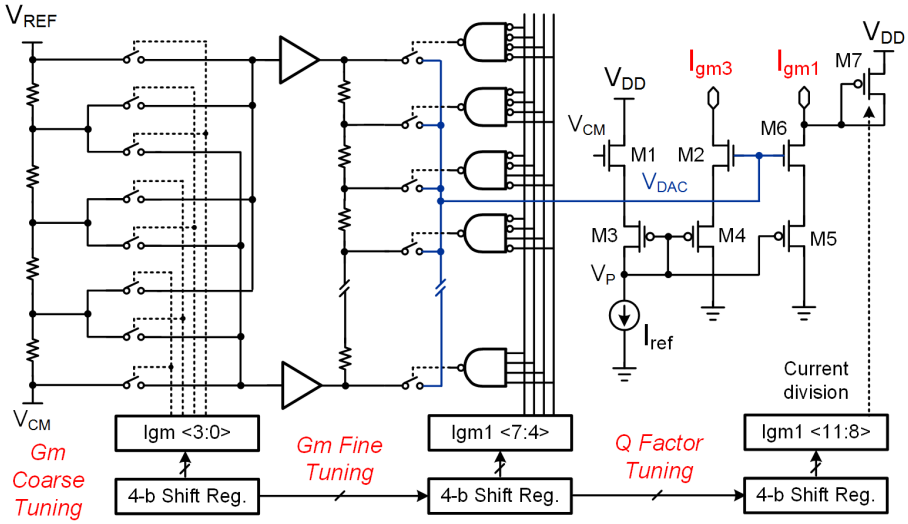


Figure 3.7: The circuit schematic of the biasing current generation module. a 64-step natural exponentially spaced biasing current can be generated.

In the current generation circuit, consider the transistors $M1$ to $M4$. The equations for the currents can be written as:

$$I_{ref} = I_{on} \exp \frac{V_{GS1}}{\zeta U_T} = I_{op} \exp \frac{V_{GS3}}{\zeta U_T} \tag{3.4}$$

$$I_{gm} = I_{on} \exp \frac{V_{GS2}}{\zeta U_T} = I_{op} \exp \frac{V_{GS4}}{\zeta U_T} \tag{3.5}$$

where I_{ref} is generated by an on-chip bandgap reference, and is independent from the temperature and the supply voltage. Eq. 3.4 and 3.5 lead to:

$$V_{GS1} - V_{GS3} = V_{GS2} - V_{GS4} = \zeta U_T \ln(\frac{I_{op}}{I_{on}}) \tag{3.6}$$

Also, from the circuit:

$$V_{CM} = V_{GS1} + V_{GS3} + V_P \tag{3.7}$$

$$V_{DAC} = V_{GS2} + V_{GS4} + V_P \tag{3.8}$$

Substitute Eq. 3.4 and 3.5 into Eq. 3.7 and 3.8 gives:

$$V_{GS1} = \frac{1}{2}(V_{CM} - V_P + \zeta U_T \ln(\frac{I_{op}}{I_{on}})) \tag{3.9}$$

$$V_{GS2} = \frac{1}{2}(V_{DAC} - V_P + \zeta U_T \ln(\frac{I_{op}}{I_{on}})) \tag{3.10}$$

The generated biasing current can be expressed as:

$$I_{gm} = I_{ref} \exp \frac{V_{DAC} - V_{CM}}{2\zeta U_T} \tag{3.11}$$

Thus, the linear voltage from the DAC is converted to an exponentially increasing biasing current. According to Eq. 3.1:

$$g_m \propto I_D \propto e^{(V_{DAC} - V_{CM})} \propto e^{code} \tag{3.12}$$

where *code* is the digital input of the DAC. Large gate area transistors are used in the current generation module to minimize the mismatch. The process variation can be further calibrated by tuning the reference voltage V_{ref}.

The $M7$ in Fig. 3.7 is a diode-connected transistor used to divide the generated current reference. The ratio of I_{gm1} and I_{gm3} can be programmed to tune the quality factor of the filter, which will be explained in the next section.

3.2.2.3 Design of the Biquad filter

A staggered tuned 4th-order band-pass filter is implemented by cascading two biquad filters [198]. A biquad filter is a 2nd order recursive linear filter containing two poles and two zeros [199]. Fig. 3.8 shows the circuit schematic of the biquad filter

implemented in this work. The biquad filter consists of four Gm blocks. The center

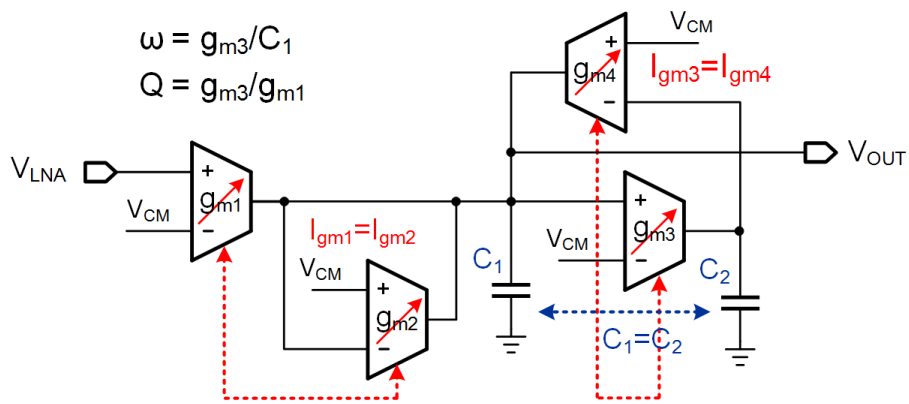


Figure 3.8: The circuit schematic of the designed biquad filter. The center frequency and the quality factor of the filter can be tuned independently.

frequency and the quality factor of each biquad are independently tunable. Only two capacitors with one terminal grounded are used in each biquad, resulting in a very compact layout. The transfer function is given by:

$$H(s) = \frac{s\dfrac{g_{m1}}{C_1}}{s^2 + s\dfrac{g_{m2}}{C_1} + \dfrac{g_{m3}g_{m4}}{C_1C_2}} \tag{3.13}$$

The biasing currents for the Gm blocks are designed to be I_{gm1}=I_{gm2}, and I_{gm3}=I_{gm4}, so that the transconductance of the Gm blocks are g_{m1}=g_{m2}, and g_{m3}=g_{m4}. The capacitors are set to be C_1=C_2. Thus,

$$\omega_C = \sqrt{\frac{g_{m1}g_{m2}}{C_1C_2}} = \frac{g_{m1}}{C_1} \tag{3.14}$$

If the C_1 is fixed, the center frequency is a function of g_{m1}. Also from Eq. 3.12 and Eq. 3.14:

$$\omega_C \propto g_m \propto e^{code} \tag{3.15}$$

Thus the center frequency of the biquad can be exponentially tuned by the digital code.

Also, the quality factor Q can be expressed as:

$$Q = \sqrt{\frac{C_1 g_{m3} g_{m4}}{C_2 g_{m2}^2}} = \frac{g_{m3}}{g_{m1}} \tag{3.16}$$

So the quality factor can be tuned by changing the ratio of I_{gm1} and I_{gm3}. As explained in section 3.2.2.2, the M7 in Fig. 3.7 is a diode-connected transistor with the same length as the current mirrors for copying I_{gm} used in the Gm block. The width of the M7 can be programmed to divide the generated current reference, so the ratio of I_{gm1} and I_{gm3} can be programmed to tune the quality factor.

3.2.2.4 Multiplier and Integrator

A Gilbert multiplier is used to square the band-passed signal. The Gilbert multiplier, or Gilbert cell, is commonly used analog multiplier circuit introduced by B. Gilbert in 1963 [200]. A tutorial of analog multipliers design can be found in [201]. The circuit schematic of the Gilbert multiplier implemented in this work is shown in Fig. 3.9. The output current of the multiplier is determined by [202]:

$$I_O = \sqrt{2K_\alpha K_\beta}(V_{IN} - V_{REF})^2 \tag{3.17}$$

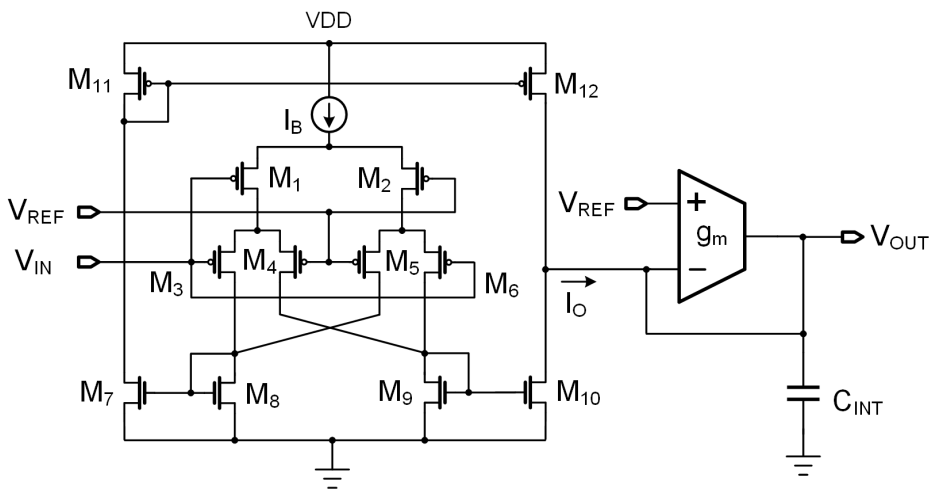


Figure 3.9: The circuit schematic of the Gilbert multiplier and the integrator.

where K is the transconductance parameter, K_α=K_1=K_2, and K_β=K_4=K_5=K_6. So the gain can be tuned by the biasing current I_B.

The integral of the output current of the multiplier is computed in the leaky Gm-C integrator [170]. The circuit schematic is shown in Fig. 3.9. The moving window length can be tuned by programming the time-constant of the integrator.

3.2.3 Measurement Results

The design has been fabricated in IBM 180nm CMOS technology. Bench testing was conducted to verify the function and performance of the fabricated design. The microphotography and layout of the LFP energy extraction channel are shown in Fig. 3.10. The occupied silicon area of the design is 850μm×115μm.

Bench testing was conducted to verify the function and performance of the fabricated chip. The frequency response of the natural logarithmic tuning neural energy extraction module was measured. The measurement was conducted point by point

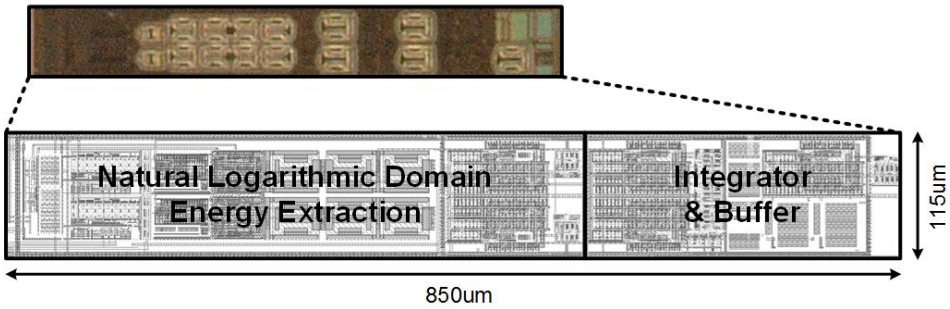


Figure 3.10: The microphotography and layout of the LFP energy extraction channel. The major building blocks are highlighted in the layout.

using a function generator 33521A and an oscilloscope MSO7034B from Agilent. The reference voltage was calibrated to set the center frequency of the unit programming step. Fig. 3.11 shows the measurements of every four steps out of the 64 possible steps, with a frequency ranging from 1Hz to 200Hz. Notice that the measurement

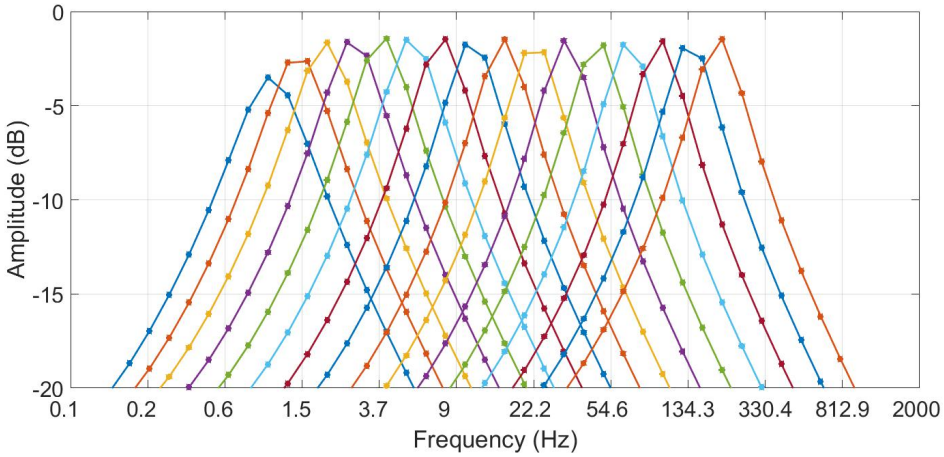


Figure 3.11: The measured frequency response of the biquad filter tuning in the proposed natural logarithmic steps. A total of 16 steps were measured.

and x-axis in the figure are both in the natural logarithmic domain.

Similarly, the tuning of the quality factor was measured. Fig. 3.12 shows the measurement result with the center frequency of the filter configured at 10Hz. The

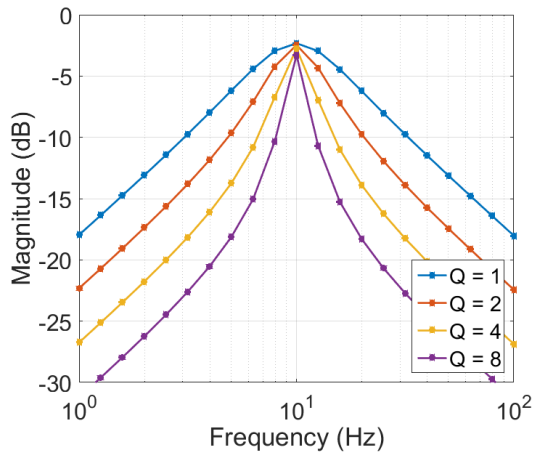


Figure 3.12: The measured the frequency response of the biquad filter with different quality factors. The center frequency is configured at 10Hz.

quality factor can be configured at 1, 2, 4 and 8.

The biquad filter was also tested with a synthetic sine wave generated from the function generator. The sine wave has a constant amplitude, and the frequency was swept from 0.1Hz to 1kHz logarithmically. The synthetic waveform and the output of the biquad filter are shown in Fig. 3.13. The center frequency was configured at 18Hz. The frequency sweeping measurement verifies the response of the biquad filter in a straight forward manner.

The Gilbert multiplier was tested with an amplitude modulated 10Hz sine wave generated from the function generator. The measurement result is shown in Fig. 3.14. The measured output of the Gilbert multiplier was plotted in comparison with the simulation result after a gain calibration. The measurement matches the simulation well.

Fig. 3.15 shows the output of the leaky integrator with an amplitude modulated 40Hz sine as the input. The measured output is compared with the theoretical

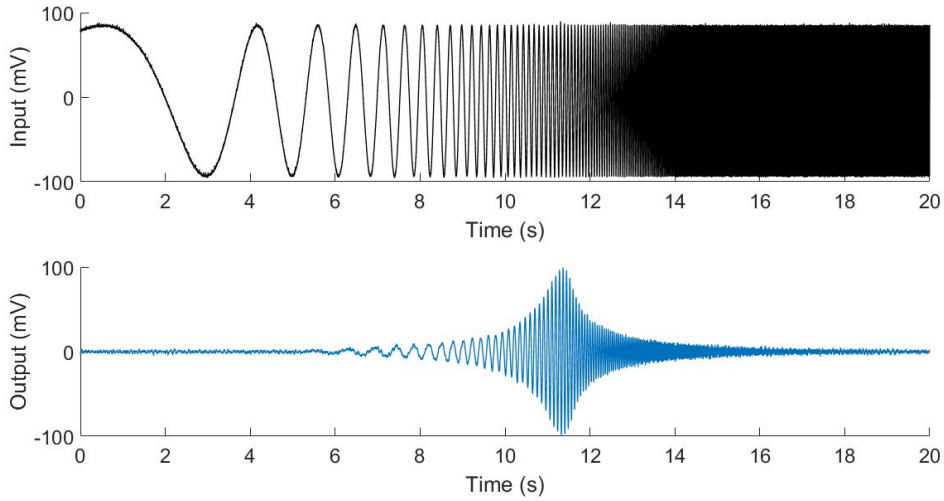


Figure 3.13: The measured response of one biquad filter with a synthetic sine wave with frequency sweeping from 0.1Hz to 1kHz.

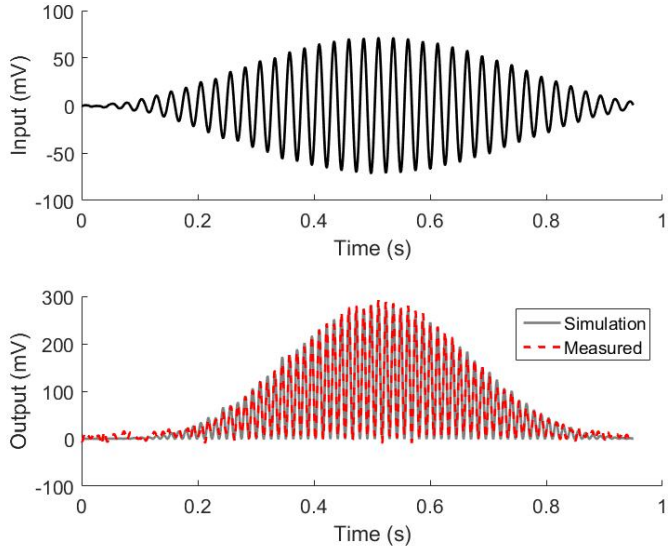


Figure 3.14: The measured output of the Gilbert multiplier with an amplitude modulated 10Hz signal. The measurement result is plotted in comparison with the simulation (after a gain calibration).

computation of the square of the input signal. The measurement matches the computation.

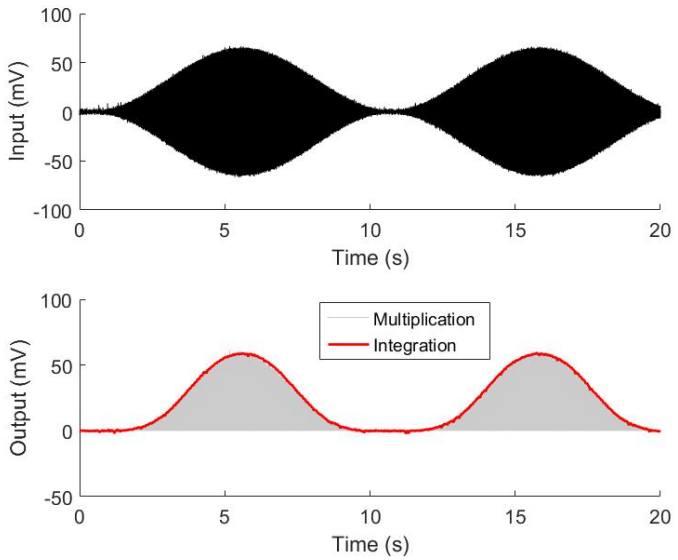


Figure 3.15: The measured outputs of the multiplier and the LFP energy integrator (phase shift have been corrected).

Fig. 3.16 shows the power spectrum of a 6-hour from a male rhesus macaque (Macaca mulatta) with electrodes implanted chronically in the left hippocampus. The recording presents the awake and asleep transition. The activities from different

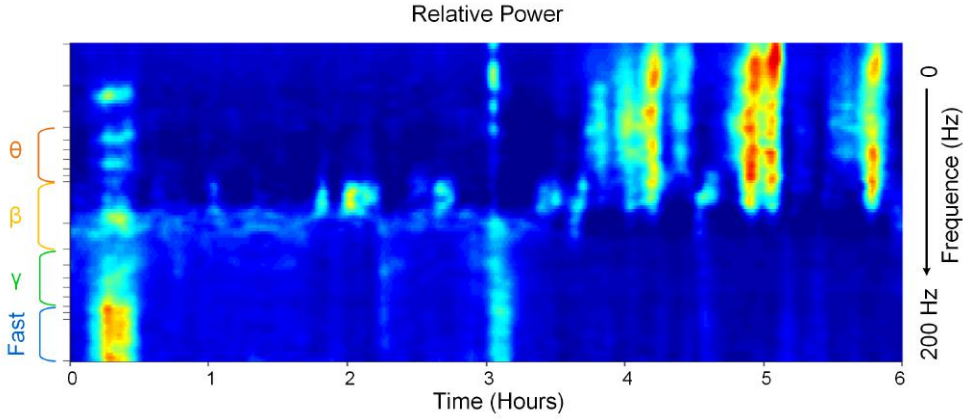


Figure 3.16: The spectrum of a 6-hour continuous recording using the prototype device. The animal was from awake (high-frequency oscillation more active) to sleep (low-frequency oscillation more active).

frequency bands are clearly visible in the figure.

Fig. 3.17 shows a 20-seconds segment of the LFP recording. The original record-

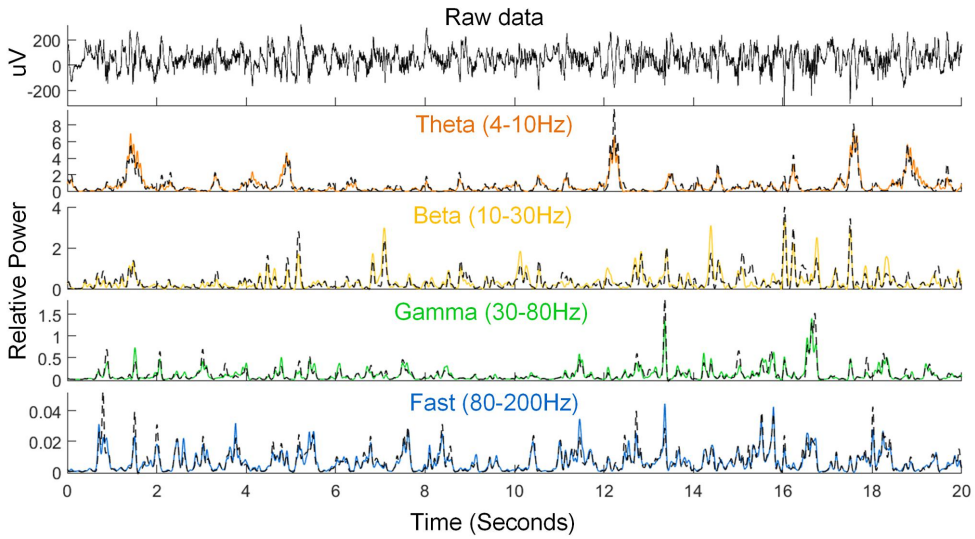


Figure 3.17: In-vivo recording in a Rhesus macaque using the designed chip. The extracted energy in four brain oscillation bands (Theta, Beta, Gamma, and Fast) compared with the theoretical computations (dashed lines).

ed signal is shown in the top line. The energy in four commonly used frequency bands (solid lines) was extracted using the designed chip, including θ band (4-10Hz), β band (10-30Hz), γ band (30-80Hz), and Fast band (80-200Hz). The measured output is compared with the theoretical computation plotted in the dashed lines after a gain normalization. A close matching between the waveforms can be observed.

3.3 Action Potential Detection

3.3.1 Introduction

Monitoring the activity of a single neuron is the basis of understanding the brain mechanisms [203]. When multiple neurons are close to one recording electrode, it is important to extract the identities of the spikes corresponding to different neurons. Given the distance and orientation relative to the recording electrode, different neuron presents different action potential waveforms. The action potentials can be then classified into different clusters. The process is known as *spike sorting* [204]. Even nearby neurons have similar responses, it is important to distinguish them and observe their individual characteristics [205].

In this section, the neuron model is briefly presented, followed by a review of action potential detection and classification methods. The design of an energy efficient continuous-time current-mode action potential detection unit is described. Circuit implementations and experimental results are presented.

3.3.1.1 Integrate and Fire Model

A good understanding of the principle of an action potential is the basis for designing a good detection circuit. The Hodgkin-Huxley Model (HHM) is a well-known model which can approximate the generation of the action potential accurately [206]. The goal of this section is to implement HHM Model for a better understanding of the integrate and fire process of a single neuron. The HHM is constructed by membrane current as the sum of a leakage current, a delayed-rectified K+ current, and a transient

Na+ current:

$$i_m = \bar{g_L}(V - E_L) + \bar{g_K} n^4 (V - E_K) + g_{Na} m^3 h(\bar{V} - E_{Na}) \tag{3.18}$$

where n, m and h are the gating variables. A channel acts as if it has gates. The opening of the gate is called activation of the conductance, and gate closing is called deactivation. The probability that the gate is open increases when the neuron is depolarized and decreases when it is hyperpolarized. In general, n, m, h are introduced as variables of the voltage and time. They are within (0,1), and can be estimated by the gating equations:

$$\tau_n(V)\frac{dn}{dt} = n_\infty(V) - n \tag{3.19}$$

$$\tau_n(V) = \frac{1}{(\alpha_n V + \beta_n V)} \tag{3.20}$$

$$n_\infty(V) = \frac{\alpha_n(V)}{\alpha_n(V) + \beta_n(V)} \tag{3.21}$$

where α_n and β_n can be found by:

$$\alpha_n = \frac{0.01(V + 55)}{(1 - e^{-0.1(V+55)})} \tag{3.22}$$

$$\beta_n = 0.125 e^{-0.0125(V+65)} \tag{3.23}$$

where m and h can be calculated in the same formula shown above:

$$\alpha_m = \frac{0.1(V + 40)}{(1 - e^{-0.1(V+40)})} \tag{3.24}$$

$$\beta_m = \frac{1}{1 + e^{(} - 0.1(V + 35))} \tag{3.25}$$

$$\alpha_h = 0.07 e^{-0.05(V+65)} \tag{3.26}$$

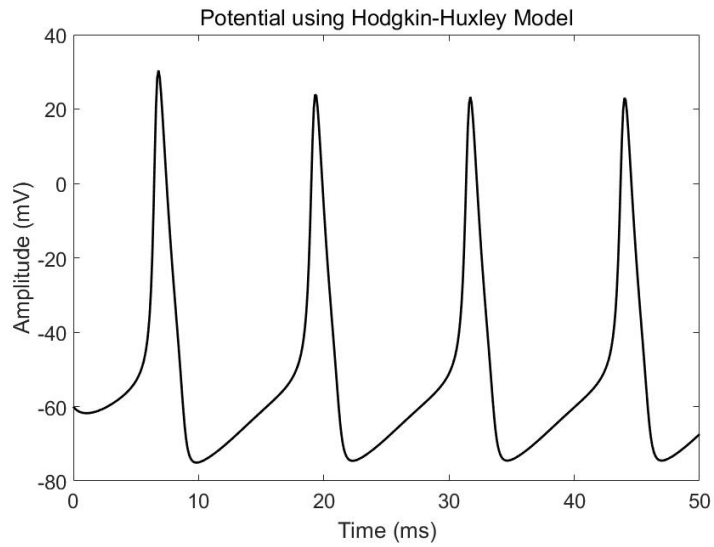


Figure 3.18: Matlab simulation of the implemented HHM model.

$$\beta_h = \frac{1}{(1 + e^{(} - 0.1(V + 35)))} \tag{3.27}$$

The HHM model was simulated in Matlab. Runge-Kutta method was used to find the arithmetic solution of the differential equations. The method requires initial conditions, which was taken from the reference [206]. The membrane potential simulation using the HHM model is shown in Fig. 3.18.

3.3.1.2 Review of Action Potential Detection Methods

The real-time spike detection and classification methods have been widely reported in the literature since the pioneering work in the 1920s [182]. Comprehensive reviews of spike detection algorithms can be found in papers [182, 203, 205, 207]. In summary, an effective method relies on a good signal-to-noise ratio and a robust detection and

classification algorithm. The general steps for action potential sorting are:

i) **Filtering**: filter the raw data between 300Hz to 6kHz for the following processing;

ii) **Detection**: detect the spike, e.g. by applying an amplitude threshold on the filtered signal. Artifacts or noise might be detected as spikes in this step;

iii) **Extraction**: extract the relevant features of the spike waveform;

iv) **Classification**: apply classifier on the extracted features for spike sorting.

Both action potential detection and classification have been implemented on-chip [144, 171, 172]. Action potential detection can be easily performed in real-time. The detection results can be used to reduce data transmission rate [13] or trigger predefined stimulation [6, 9]. Commonly used action potential detection methods are summarized here:

i) **Absolute threshold detection**, which uses a predefined threshold for threshold detection [207]. The threshold can be manually set or using several times (3x - 5x) of the root mean square value of the signal;

ii) **Non-linear energy operator (NEO)**, which extracts the energy from the action potential signal to improve the detection integrity [143]. A modification of the NEO, called the multiresolution Teager energy operator (METO) combines the results of NEO in different resolution scales, also shows good performance;

iii) **Wavelet analysis**, which projects the signal to certain wavelets domain [172]. The wavelet transform can be seen as a bank of matched filters.

The performance of different spike detection algorithms has been compared in [171, 207]. The conclusion is that for systems with limited computational resources, applying an absolute threshold on the signal is just as effective for detecting spikes as applying more elaborate energy-based nonlinear operators.

Commonly used features of action potentials include: i) the maximum spike amplitude, ii) the minimum spike amplitude, iii) the spike width, and so on. Intuitively, more features of the action potentials give better performance in distinguishing the clusters. However, manually choosing the features sometimes yields a poor separation. One method for choosing the features automatically is principal component analysis (PCA) [182]. PCA can find an ordered set of orthogonal basis vectors that capture the directions in the data of the largest variation.

There are many methods for clustering [182], including K-means clustering, Bayesian clustering, support vector machine (SVM), and so on. For example, K-means clustering, or nearest-neighbor clustering is a hardware friendly classifier. K-means clustering defines the cluster location as the mean of the data within that cluster. A spike is classified to the cluster with minimum the Euclidean distance. The performance of different classifiers has been compared in [182, 203]. To be noticed that there are other issues affecting the spike sorting algorithms, including electrode drifting, spike overlapping, neuron bursting and so on.

3.3.2 Circuit Implementation

The action potential detection can be performed in the analog domain [143, 208] or the digital domain [209]. The duration of an action potential is less than 2.5ms [205]. For the accuracy of the classification, the sampling rate should be at least 10kSps. Analog spike detection can achieve an ultra-low power consumption, while digital domain processing can achieve superior performance and classification accuracy.

In this work, a current-mode continuous time action potential discrimination unit has been designed. The current-mode circuits present the signal as a current instead

of a voltage, thus the dynamic range of the signal is not limited by the supply voltage. This can be very useful for implementing a large dynamic range signal processing in advanced CMOS technology where a low supply voltage is often used. The block diagram of the current-mode action potential detection module is shown in Fig. 3.19. The overall system consists of a low-noise amplifier, a bandpass filter, and the action

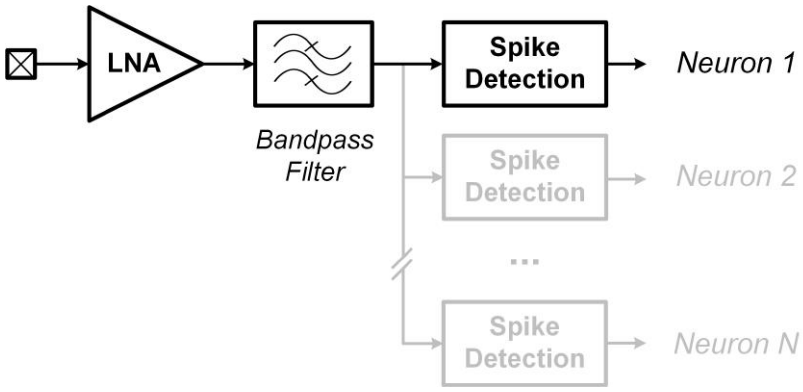


Figure 3.19: The block diagram of the action potential detection module.

potential detection unit. The bandpass filter is usually configured with a passband from 300Hz to 6kHz in a 2nd order or higher. Multiple spike detection units can be connected to discriminate more than one neuron per channel. In a multiple channel recording system, this can be achieved by designing a multiplexing module among channels.

The working principle and the block diagram of the action potential detection unit are shown in Fig. 3.20. Two amplitude thresholds and time windows are used to discriminate the APs from different neurons [67]. After a bandpass filter, the neural signal is first compared with a depolarization threshold $TH1$. If the signal exceeds $TH1$, the comparator is then disabled for a period of $\Phi 1$. Then the signal is compared with a second repolarization threshold $TH2$ for a period of $\Phi 2$. If the signal crosses

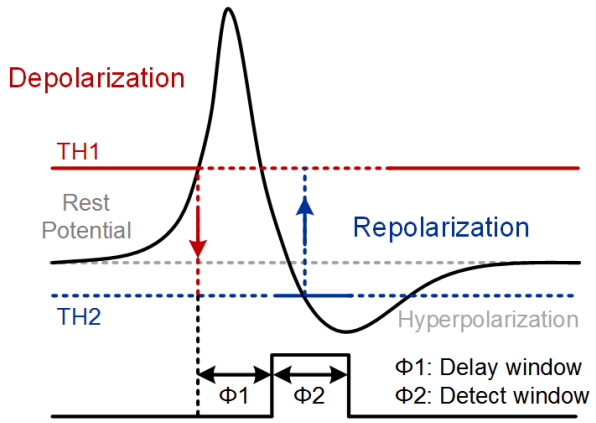


Figure 3.20: Illustration of the window discriminator for action potential detection principle.

$TH2$ during $\Phi 2$, an action potential is detected. The $TH1$ and $\Phi 1$, $TH2$ and $\Phi 2$ form two discrimination windows, which can be programmed.

The circuit block diagram of the implemented action detection unit is shown in Fig. 3.21. The designed action potential detection unit consists of a transconductance

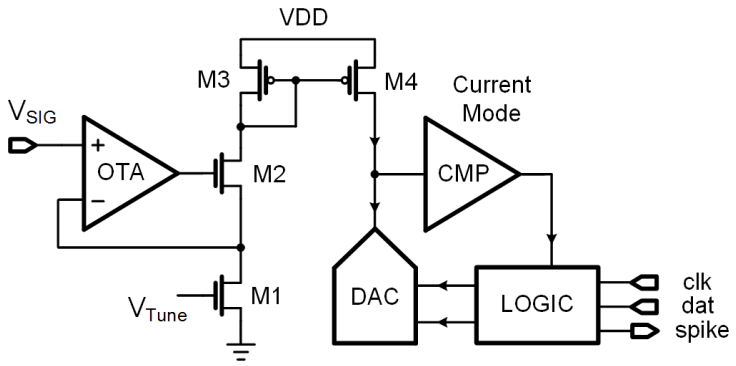


Figure 3.21: The block diagram of the programmable current mode spike detection unit with integrated programmable amplitude-window discriminator. The filtering state is not shown in this diagram.

amplifier, a current-mode DAC, a current-mode comparator, and a digital timing and logic module. The transconductance is set by the biasing voltage V_{Tune}, while M1 is in the deep triode region. V_{Tune} can also compensate the threshold variation of M1.

The threshold currents are generated by a 6-bit current DAC. The circuit schematic of the current-mode DAC is shown in Fig. 3.22. The DAC uses binary weighted current mirrors. The current steering can be disabled by shorting I_{OUTP} to ground. Disabling the current steering lowers the settling speed, which may limit the performance of the DAC. No additional calibration is implemented in this work. Notice that 5-bit resolution is usually more than sufficient for the window discrimination algorithm. A finer tuning of the threshold values won't give a better discrimination accuracy.

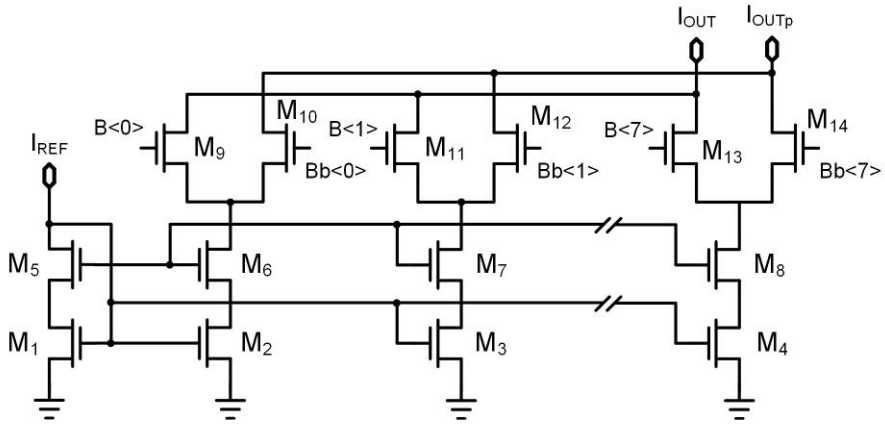


Figure 3.22: The circuit schematic of the current-mode DAC.

The comparison with a threshold current is performed in a current-mode comparator. The circuit schematic of the current-mode comparator is shown in Fig. 3.23. An ideal current comparator has a low input impedance, and the input node voltage should be fixed [210]. However, a capacitive input stage can detect a low current with a much faster response and with a low power consumption [211]. But the input node voltage of a capacitive input stage cannot be well controlled. As a result, a combination of both capacitive input stage and resistive feedback is implemented. The transistors $N3$ and $P3$ work as non-linear feedback resistors to set the input voltage. When the input signal is small, the feedback loop is disabled and the comparator

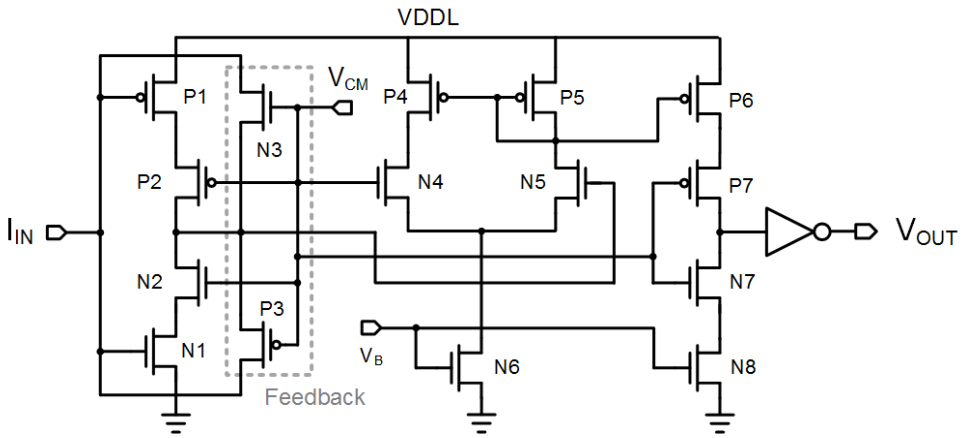


Figure 3.23: The circuit schematic of the current-mode comparator.

appears capacitive characteristic, which ensures a high resolution and speed. It is important to minimize the input capacitance, especially when the designed current is low. A differential pair is used as the second stage, followed by a current starved output buffer.

The digital timing and logic module are designed with custom two wire interface. There are four registers for the two threshold amplitudes and time windows. The output of the digital module is a flag for the detection of an action potential. The flag signal is synchronized with the clock externally. Like many real-time spike detection algorithms, one drawback of this design is the false negative detection of an action potential signal occurring right after an artifact. In this case, the spike detection won't be able to recover in time. Data buffer can be used to address this problem by re-alignment of the input data. However, an analog buffer is difficult to implement. An alternative solution will be to use detection units in parallel, similar to Fig. 3.19. A digital logic will need to be designed to properly address the conflict by allowing one unit to process one action potential once a time.

It should be noticed that with the current-mode comparator and DAC, if a SAR logic is added, the action potential detector can be extended to a current-mode ADC. In future work, it might be beneficial to implement the action potential detection in the analog domain and follow it by digitization for further processing in the digital domain.

3.3.3 Experimental Results

The designed action potential detection unit has been fabricated in IBM 180nm C-MOS technology. The design has a dimension of 125μm×25μm. The microphotograph of the fabricated chip is shown in Fig. 3.24, with major building blocks highlighted.

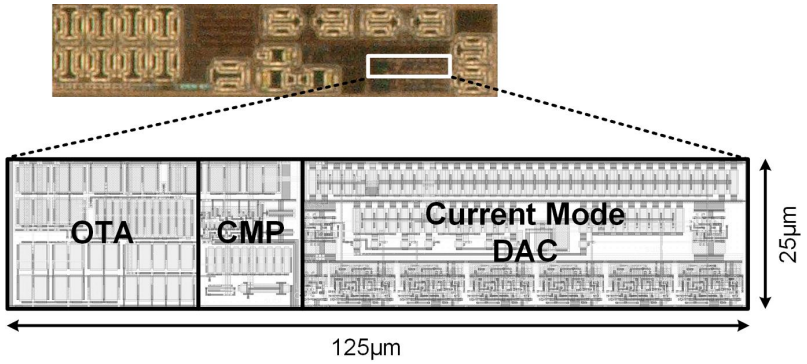


Figure 3.24: The micrograph and layout of the designed analog action potential detection module.

A couple of bench tests were conducted to verify the functions and evaluate the performance of the unit. The experimental results are presented as follows. The DAC was measured with a worst INL and DNL less than 1LSB. The ENOB is 5.6-bit. The supply voltage is from 1 to 1.8V. The average power consumption of the module is

4μW with a supply voltage of 1V. The clock frequency for configuration and output synchronization is set to be 100kHz, which gives a maximum latency of 10μs.

Synthetic neuronal signals with different SNRs were generated using an arbitrary function generator 33521A from Agilent to test the action potential detector. A 2-min recording segment was used for testing. The first 10-seconds signal is shown in Fig. 3.25 for illustration. Fig. 3.25 (a) shows the original signal recorded by the PennBMBI recorder from the whisker motor cortex in an anesthetized rat. The sampling rate was 21kSps. The recorded data has an SNR of 25.7dB. The real action potentials are marked by triangle markers. Fig. 3.25 (b) and (c) show the recorded signal with additional white noise and artifacts resulting in a SNR of 20dB and 15dB, respectively. The artifacts are designed to mimic the motion or chewing effects which commonly occurred during the neural recording in a freely behaving animal. There are N real action potentials in the recording, where N=117 in this case. The true positive T_P is defined as the correct recognitions. The false negative F_N is defined as the wrong recognitions. The false positive F_P is the missed action potentials. The evaluated performance of the designed module is listed in Table 3.1. The ratio

Table 3.1: The Measured AP Detection Accuracy of Signal With Different SNR

SNR	Algorithm	TP	FN	FP
25dB	Win Discrim	95.8%	4.2%	0%
	Threshold	97.4%	2.6%	0%
20dB	Win Discrim	93.2%	6.8%	0%
	Threshold	77.6%	2.0%	20.4%
15dB	Win Discrim	83.8%	6.2%	10.0%
	Threshold	50.5%	8.0%	41.5%

was calculated over the total number of T_P $+F_N$ $+F_P$. The performance of the two-window discrimination is also compared the detection using only a threshold.

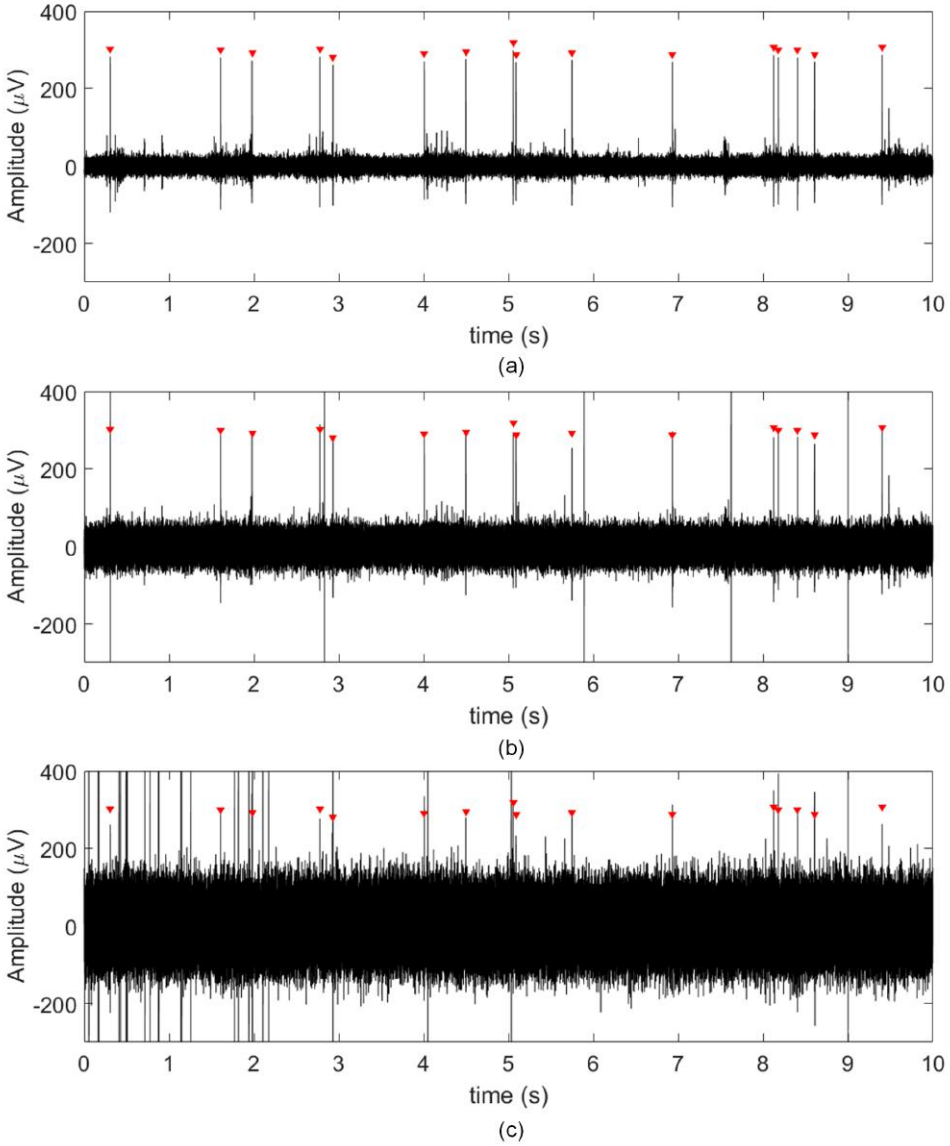


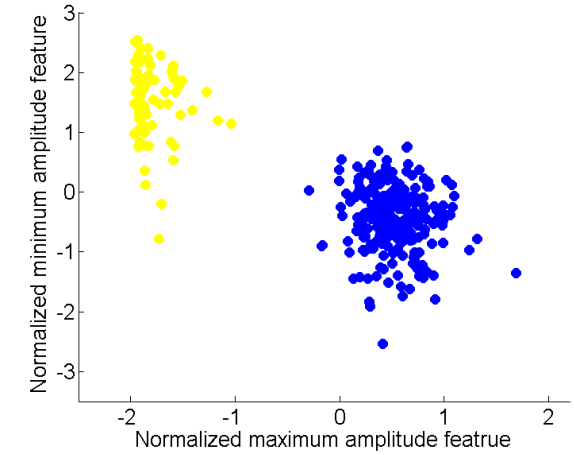
Figure 3.25: The action potential signal used for testing the designed detection module. The real action potentials are marked by triangle markers. (a) The original signal with an SNR of 25.7dB. (b) Testing signal adding white noise and artifacts with an SNR of 20dB. (c) Testing signal adding white noise and artifacts with an SNR of 15dB.

The experimental results suggest that with a good SNR, both a simple threshold and the window discrimination give excellent detection results. The simple threshold gives slightly better result than the window discrimination, mainly because several

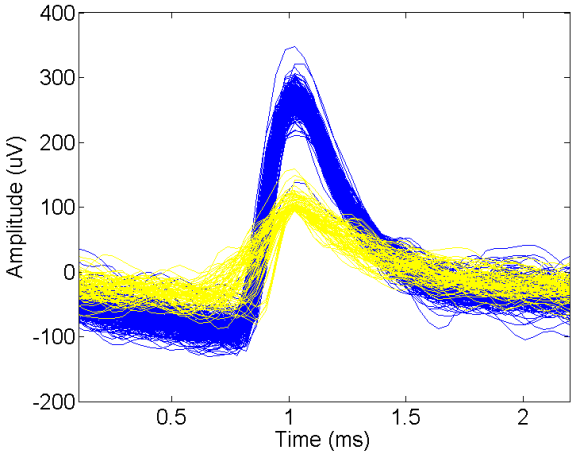
action potentials fail to pass the second window. The performance of using the simple threshold drops significantly after adding artifacts. With the increased noise, it is hard to set the threshold voltage, and the window discrimination clearly rejects more false detections than the simple threshold. But the window discrimination also makes more mistakes in the higher noise environment. Some artifacts are mistaken as the real action potentials when the noises pass the second window.

A cluster analysis was performed in the microcontroller. The algorithm was programmed in the C language. During this experiment, two neurons were captured in the same recording electrode. The normalized maximum and minimum amplitudes were calculated and used as two features for the clustering analysis. The K-means clustering was used to separate the two neurons. Fig. 3.26 (a) illustrates the analysis result in the feature domain, which clearly shows the two clusters well separated. The action potentials are plotted with color coding based on the classification results, as shown Fig. 3.26 (b).

In this section, the design of a real-time current-mode action potential detection and discrimination unit is presented. The design features low power, robust detection, and small silicon area, which is suitable for an integration into a high channel count neural recording front-end or a BMI device. A classification was implemented in the microcontroller. In the future, an on-chip classification can be integrated to further support the closed-loop BMI applications.



(a)



(b)

Figure 3.26: A cluster analysis of the action potentials from two neurons. (a) Normalized maximum and minimum amplitudes are calculated and used as two features for the analysis. (b) The action potentials are labeled with different colors according to the classification results.

3.4 Matched Filter for Neural Feature Extraction

3.4.1 Introduction

A matched filter is the optimal linear filter for maximizing the signal-to-noise ratio in the presence of additive stochastic noise [212]. The matched filter is obtained by correlating a known template with the unknown input signal to detect the presence of the desired signal [213]. Matched filters are commonly used in wireless communications [214], radar and sonar [215], gravitational-wave astronomy [216], medical applications [217], and so on [212].

A number of studies propose to implement matched filters for action potential detection [218–220]. In addition, the matched filter can be applied to detect phase-amplitude coupled low-frequency neural rhythm [221]. For example, the cortical μ rhythm is an event-related desynchronization commonly used for BMI control. However, the μ rhythm's typical frequency band is 8-12 Hz, which is overlapping with the virtual α rhythm (section 3.2). Thus, an energy based feature extraction method often has difficulties in discriminating them. In this case, matched filters have the advantage in accurately modeling the phase-coupled rhythm.

Moreover, the performance of the matched filters can be optimized by pre-whitening the signal. This process can be achieved by implementing the pre-whitening filter proposed in Chapter 2. By combining the phase correction filter and matched filter together, a very energy efficient hardware implementation can be achieved. By programming the coefficients of the filter, it can be used in a wide range of applications, and is very suitable for an integration on a neural interface.

This section presents the analysis, design and testing of a matched filter with pre-whitening for the neural signal extraction. In the end of the section, a compressed sparse matched filter is explored to reduce the requested computation and hardware cost.

3.4.2 Matched Filter and Pre-whitening for Optimum Correlation Detection

The process of the matched filter is illustrated in Fig. 3.27. The input $x(t)$ consists

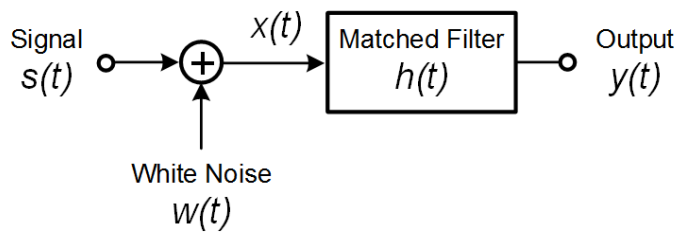


Figure 3.27: The block diagram of the matched filter.

of the signal $s(t)$ corrupted by a white noise $w(t)$, as shown by:

$$x(t) = s(t) + w(t) \tag{3.28}$$

where $w(t)$ has a zero mean and power spectral density of $N_o/2$. If the filter is linear, the output is:

$$y(t) = s_o(t) + w_o(t) \tag{3.29}$$

The signal to noise ratio (SNR) is defined as:

$$SNR = \frac{|s_o(T)|^2}{w_o^2(t)} \tag{3.30}$$

Assume the noise spectral density is $N_0/2$, and the transfer function is $H(f)$, Eq. 3.30 can be written as:

$$SNR = \frac{|\int_{-\infty}^{\infty} H(f)S(f)e^{j2\pi f t_d} df|^2}{\frac{N_0}{2}\int_{\infty}^{\infty} |H(f)|^2 df} \tag{3.31}$$

To find the maximum SNR, use the conclusion of Schwarz inequality [222]:

$$|\int_{-\infty}^{\infty} f_1(x)f_2(x)dx|^2 \leq \int_{-\infty}^{\infty} |f_1(x)|^2 df \int_{-\infty}^{\infty} |f_2(x)|^2 dx \tag{3.32}$$

only if

$$f_1(x) = k f_2^*(x) \tag{3.33}$$

Now set

$$f_1(x) = H(f) \quad \text{and} \quad f_2(x) = S(f)e^{j2\pi f t_d} \tag{3.34}$$

So the Eq. 3.32 can be rewritten as:

$$|\int_{-\infty}^{\infty} H(f)S(f)e^{j2\pi f t_d} df|^2 \leq \int_{-\infty}^{\infty} |H(f)|^2 df \int_{-\infty}^{\infty} |S(f)|^2 df \tag{3.35}$$

And the Eq. 3.31 can be rewritten as:

$$\begin{aligned} SNR &= \frac{|\int_{-\infty}^{\infty} H(f)S(f)e^{j2\pi f t_d} df|^2}{\frac{N_0}{2}\int_{\infty}^{\infty} |H(f)|^2 df} \\ &\leq \frac{\int_{-\infty}^{\infty} |H(f)|^2 df \int_{-\infty}^{\infty} |S(f)|^2 df}{\frac{N_0}{2}\int_{-\infty}^{\infty} |H(f)|^2 df} = \frac{2}{N_0}\int_{-\infty}^{\infty} |S(f)|^2 df \end{aligned} \tag{3.36}$$

Thus the maximum SNR can be found from Eq. 3.36, when:

$$H(f) = kS^*(f) \tag{3.37}$$

In the time domain,

$$h(t) = ks^*(t_d - t) \tag{3.38}$$

where k is an arbitrary constant. The matched filter $h(t)$ is just a time-reversed version of the signal with a gain factor.

The above analysis assumes that the noise has a white spectral density. However, the neural signal and electronics noise both have a frequency-dependent spectral density, as analyzed in Chapter 2. If the noise and background signal can be pre-whitened, the correlation detection can still be optimized by the matched filtering [223]. The process is shown in Fig. 3.28. The generation of the pre-whitening filter

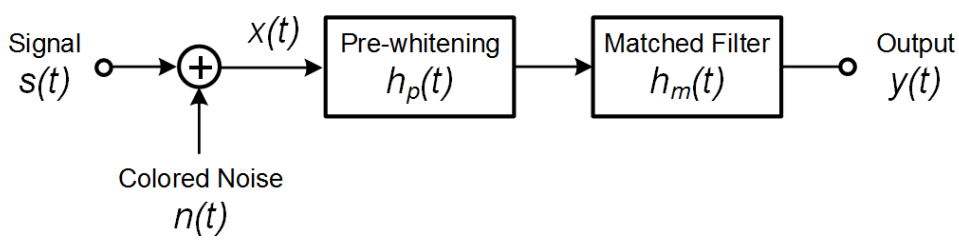


Figure 3.28: The block diagram of the matched filter in combination with the pre-whitening filter for the correlation optimization.

requires a prior knowledge of the noise spectrum. However, this is usually not feasible for a real-time implementation. The pre-whitening filter proposed in Chapter 2 is a low-cost hardware solution to improve the performance of the matched filter. The following analysis and test results verify the hypothesis.

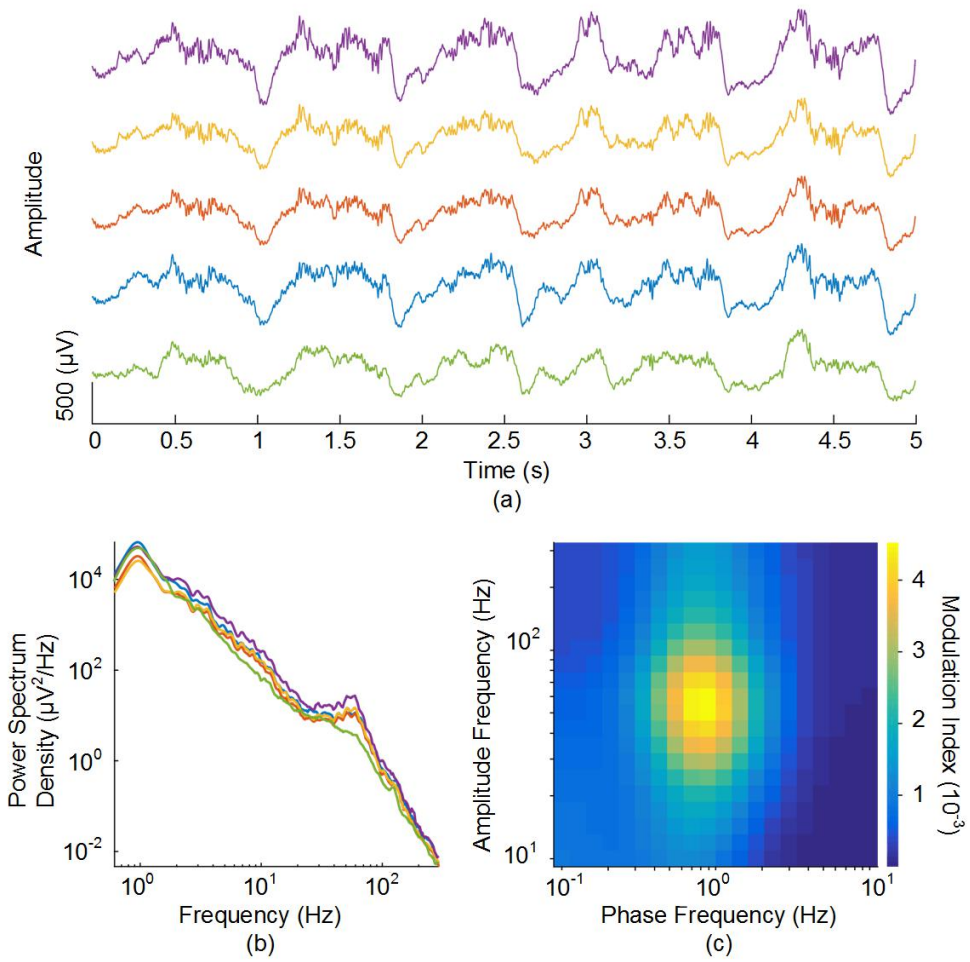


Figure 3.29: (a) A 5-sec recording from an anesthetized rat. (b) Power spectrum density of the recording. (c) Phase-amplitude coupling analysis of the signal.

3.4.3 Methodologies

A. Dataset

The neural dataset used in this study is a 5-min recording from an anesthetized rat. The sampling rate was 24.41 kHz and was down-sampled to 2441 Hz before the processing. Fig. 3.29 (a) shows a 5-sec recording from five different channels. Lots of 1-Hz oscillations (typically called an up-down state) can be observed in the recording.

An energy peak at 1Hz can also be seen from the power spectrum, as shown in Fig. 3.29 (b). The oscillation has a strong amplitude-phase coupling, as shown in Fig. 3.29 (c) [224].

A cycle-triggered average analysis was applied to find out these 1-Hz oscillations: 1) a total of 310 segments were detected in the recording, which were used as the data bank in the following study; 2) these segments were all aligned on the down-state peaks; 3) the average of these segments was used as the target neural feature waveform, and is referred as the template in the following study. Fig. 3.30 (a) shows the neural feature waveform (template). The template is assumed to be noiseless.

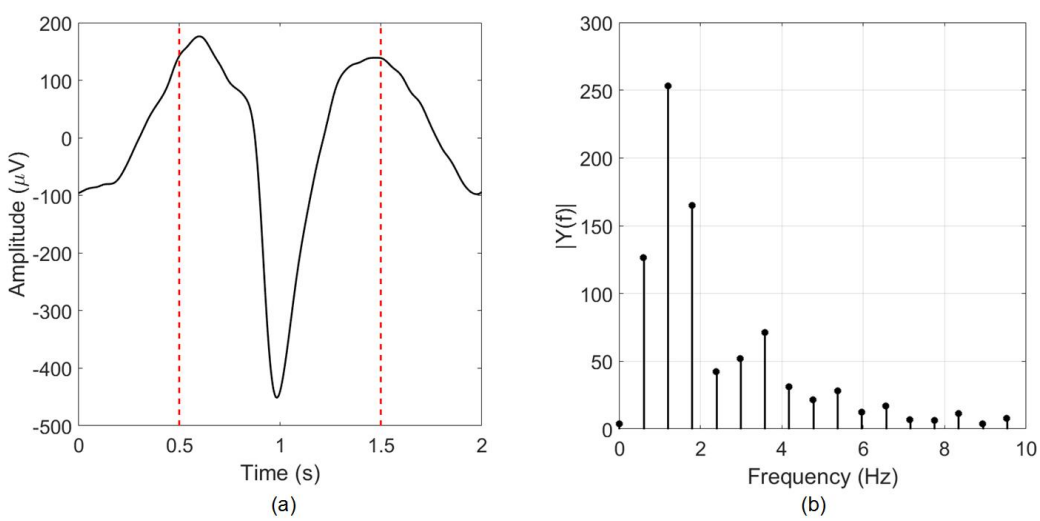


Figure 3.30: (a) The neural feature waveform (template). The waveform has more time in the "up-state" than the "down-state", so it is not an ideal sinusoid wave. (b) Frequency analysis of the neural feature.

Clearly, the waveform has more time in the "up-state" than the "down-state", so it is not an ideal sinusoid wave. Fig. 3.30 (b) shows the frequency analysis of the neural feature. The dominant frequency components are from 0.6 Hz to 1.5Hz.

B. Bandpass Filter

Bandpass filters are used for comparing the detection performance with the proposed matched filter. Several bandpass filters have been implemented, including Butterworth filters, Chebyshev filters, Biquad filters, and different types of FIR filters. Fig. 3.31 shows a comparison of the frequency responses of these filters. The cost for

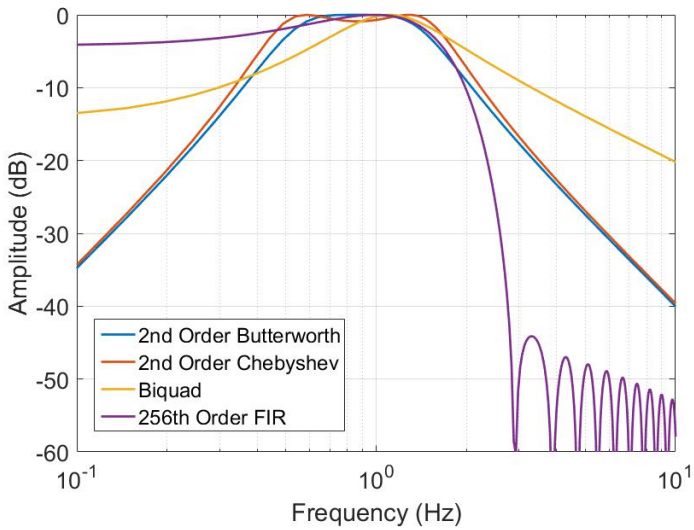


Figure 3.31: A comparison of the frequency response of different filters for extracting the slow oscillation.

a FIR filter to achieve such a narrow frequency band is significantly higher than the IIR filters, especially when the sampling rate is high. For simplification, a 2nd order Butterworth filter is used as the bandpass filter in the following study for the comparison purpose. The cut-off frequencies were chosen to be 0.6 Hz and 1.5Hz. Notice that the Butterworth filter can be implemented in the hardware in either analog or digital circuits.

C. Matched Filter with Pre-whitening

As derived in Eq. 3.38, the matched filtering was performed by the convolution of the signal and the time reversed version of the template:

$$y(t) = x(t) * h_m(t) \tag{3.39}$$

where $x(t)$ is the input signal, $h_m(t)$ is the matched filter, and $y(t)$ is the output signal. The pre-whitening filtering was implemented by a 1st order highpass filter, with a corner frequency of 100Hz. This mimics of the actual hardware circuit implementation of the pre-whitening filter proposed in Chapter 2. The template was also pre-whitened by the same filter to compensate the phase distortion.

$$y(t) = (x(t) * h_w(t)) * (h_m(t) * h_w(t)) \tag{3.40}$$

where $h_w(t)$ is the pre-whitening filter. It should be noticed that the $h_m(t) * h_w(t)$ can be pre-computed to save the cost in the hardware implementation.

After the filtering, the output signal was squared to find the energy. A moving average filter with a window size of 1-sec is the then applied to find out the envelope, and a threshold is used to detect the event.

3.4.4 Experimental Results

3.4.4.1 Detection of Synthesized Signal

This section describes the detection results of the synthesized signal with different SNR. Take the 2-sec template signal (noiseless) $s(t)$ and add random generated pink noise $n(t)$ to get the test signal $x(t) = s(t) + n(t)$. Since the signal power is known,

so the SNR can be controlled by changing the energy of the pink noise.

$$SNR = \frac{P_{Signal}}{P_{Noise}} \tag{3.41}$$

Fig. 3.32 shows an example of 16 synthesized testing signals with SNR ranging from 0.2 to 6. Then each of the 2-sec synthesized signals was superimposed on a 100-sec

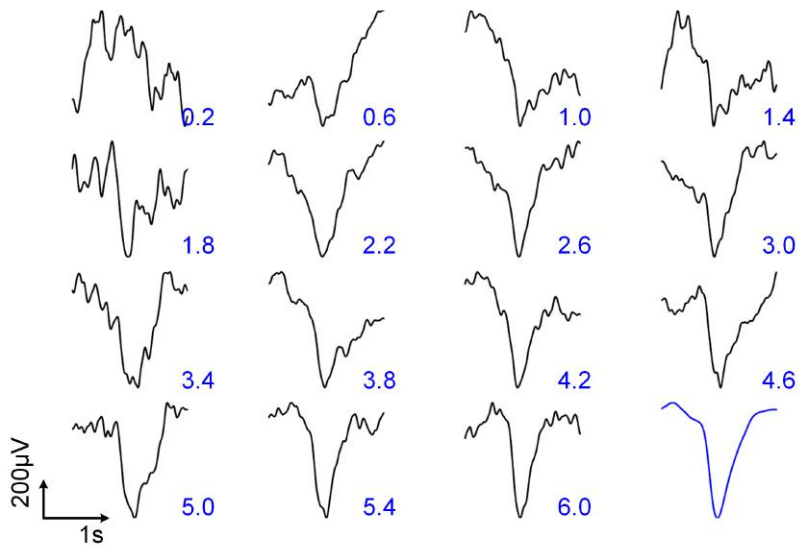


Figure 3.32: Examples of the synthesized test signals with SNR ranging from 0.2 to 6. The last one is the template.

pink noise signal. Different filters were applied to the 100-sec data for the neural feature detection. The detection accuracy is defined as:

$$Accuracy = \frac{T_P}{T_P + F_N + F_P} \tag{3.42}$$

where the truth positive T_P is the correct detection, the truth negative F_N is the wrong detection, the false positive F_P is missed detection.

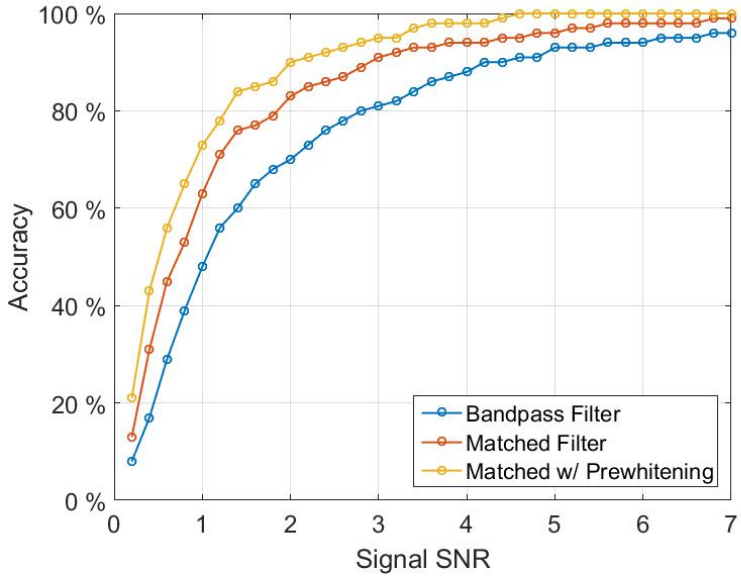


Figure 3.33: Detection accuracy of different filters for signals with different SNR (0.2 to 7 in a step of 0.2). The detection accuracy for each SNR step was an average of 100 trials with random pink noise. A total of 3,500 trials were tested for each filter in this experiment.

The test signals with SNR ranging from 0.2 to 7 with a step of 0.2 were generated. 100 trials with random pink noise were generated for each SNR step. So a total of 3,500 trials of 100-sec testing signals were used for testing the performance of each filter. Fig. 3.33 shows the testing results. The result shows that the matched filter has a better accuracy than the bandpass filter for detecting this neural feature. And the pre-whitening filter further improves the detection accuracy of the matched filter. The experimental results verify the hypothesis.

3.4.4.2 Detection of Recorded Neural Signal

This section evaluates the detection accuracy of the real neural signals from the data bank. 100 real neural signal segments were randomly selected from the data bank.

Fig. 3.34 shows 15 examples of the segments, in comparison with the template. 100 segments of 100-sec pink noise were generated. The 100 neural signal segments

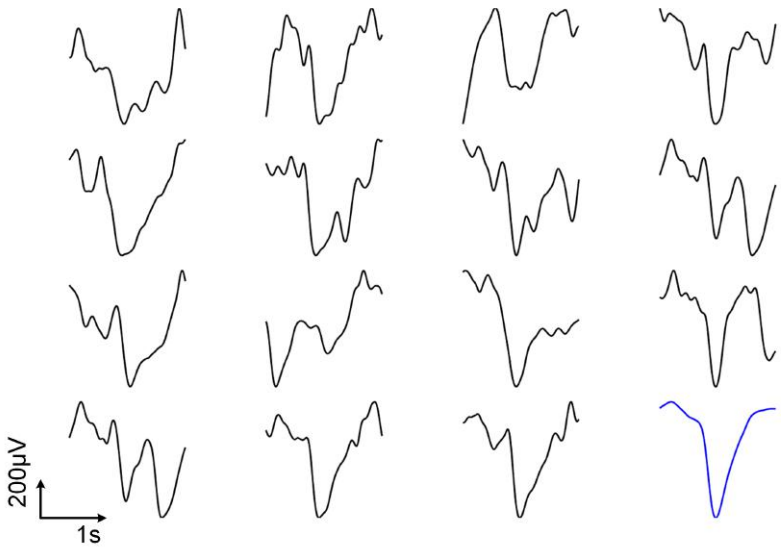


Figure 3.34: Examples of the randomly selected real neural signal segment from the data bank. The last one is the template.

were superimposed on these pink noise segments, so a total of 10,000 trials were generated. The SNR for each signal segment was calculated. The detection results using the bandpass filter, the matched filter, and the matched filter with pre-whitening are plotted in Fig. 3.35. The detection result shows that the pre-whitening filter improves the matched filters performance, especially in the low SNR cases. The average detection accuracy is lower than the first experiment, which may due to the existence of more than one oscillation in the 2-sec data segment (only 1 true feature is assumed in each trial). This experiment verifies the hypothesis that matched filter with pre-whitening can achieve superior performance in detecting phase-amplitude coupled neural rhythm.

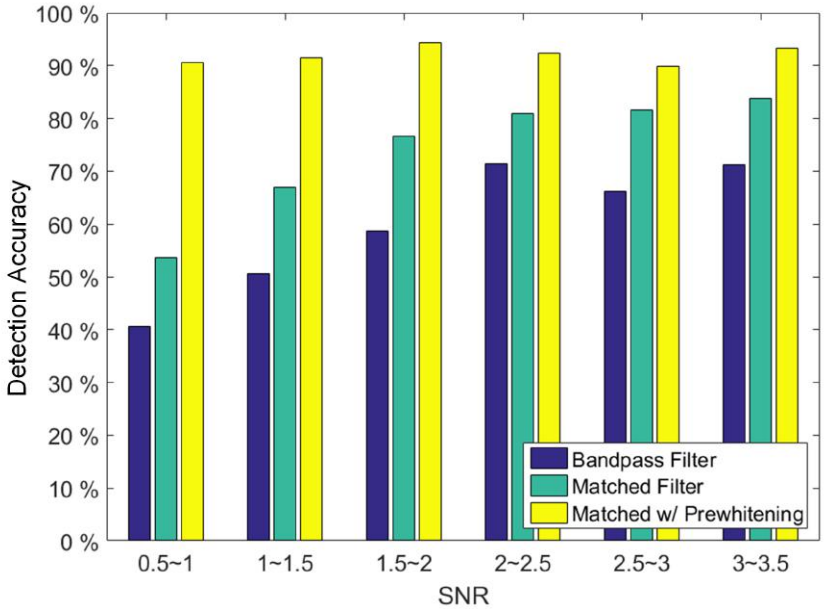


Figure 3.35: Detection accuracy of different filters for 100 signal segments randomly selected from the data bank. Each of the signal segments is inserted into 100 random pink noise trials. A total of 10,000 trials were tested for each filter in this experiment. The detection result shows that the pre-whitening filter improves the matched filter's performance in low SNR.

3.4.4.3 Application of Compressed Sampling

This section shows the detection results by applying compressed sampling techniques to the matched filter. The experiment setup is the same as in the previous sections. The template used for matched filter is randomly compressively sampled, and the incoming data is sampled in the same way.

$$y(t) = ((x(t) * h_w(t)) * V) * ((h_m(t) * h_w(t)) * V) \tag{3.43}$$

where V is a sparse vector contains only 0 and 1. The number of ones over the total length of the vector is the compression ratio. $(h_m(t) * h_w(t)) * V$ can be pre-computed

to save the computational cost.

The experimental results are shown in Fig. 3.36. The detection accuracies of

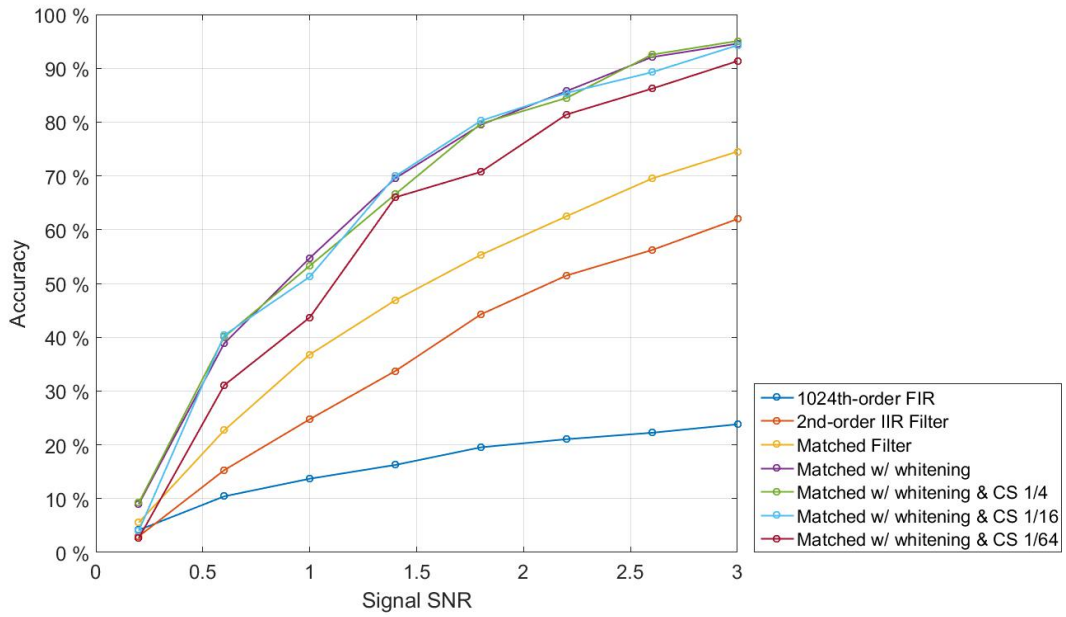


Figure 3.36: Comparison of the detection accuracy of different filters. The experimental results show that the matched filter with pre-whitening has the best performance. The matched filter with pre-whitening and a compression ratio of 64x still has a better performance than the matched filter without pre-whitening.

different filters with SNR ranging from 0.2 to 3 in a step of 0.4 are compared. The results show that the matched filter with pre-whitening has the best performance, while the conventional bandpass filter is the worst. Applying compressed sampling of the matched filter doesn't compromise the detection accuracy up to a compression ratio of 16x. However, even with a compression ratio of 64x, the matched filter with pre-whitening still has a better detection accuracy than the matched filters without pre-whitening. Finally, the 64x compressive sampled matched filter with pre-whitening achieves a detection accuracy over 90% given an SNR of 3dB, and over 98% given an SNR of 6dB. It should be noticed that the pre-whitening performed in

these experiments is simply highpass filtering as proposed in Chapter 2, thus can be easily implemented in hardware.

In summary, this section has described the design and testing of the matched filter for neural feature extraction. The pre-whitening filter is used to further improves the detection accuracy. In addition, compressed sampling has been used to reduce the computational cost. The experiment was based on a dataset of recordings in an anesthetized rat. A 1-Hz up-down state oscillation was used as the target neural feature. The experimental results suggest that: 1) The performance of the matched filter is better than the conventional bandpass filter in detecting this feature; 2) The pre-whitening processing further improves the performance of the matched filter, especially in low SNR cases; 3) Compressively sensing up to 64x can be applied to the matched filter with pre-whitening, achieving a similar performance to the matched filter without pre-whitening, which is still much better than conventional bandpass filter. This proves that the matched filters with pre-whitening are promising for closed-loop BMI integration for extracting a wide-range of amplitude-phase coupled neural features.

Chapter 4

Neural Stimulator Design

4.1 Introduction

Electrical stimulation of excitable neurons is one of the most prevalent functions performed in biomedical implantable devices [225]. The first electrical brain stimulation was pioneered by researchers Luigi Rolando and Pierre Flourens in the early 19th-century [226], and the development of the medical stimulator began with the early pacemaker design in the 1930s [227]. The development of electronics, especially integrated circuit technologies, enables the design of accurate, reliable, and miniature stimulators for neuroscience research and clinical treatment. Nowadays, electrical stimulators have been widely used for deep brain stimulation (DBS), functional electrical stimulation (FES), spinal cord stimulation (SCS), visual and auditory neural stimulation, brain-machine interface (BMI), neuroprosthetics and many other clinical therapeutic treatments [225, 228].

The clinical adoption of this technology requires the neural stimulator to be designed with a high level of safety, reliability, programmability, and minimum power consumption. In addition, a sufficient number of channels, stimulation bandwidth, flexible configuration, implantable device dimension, and wireless communication capability are also essential features. Lots of circuit techniques have been developed to address the challenges of neural stimulator development.

This chapter presents the analysis and design of high efficiency electrical neural stimulators. The design of a general-purpose neural stimulator is reviewed and summarized, and a novel stimulation strategy is proposed to address a practical problem based on the understanding of the electrode-electrolyte interface. The chapter is organized as follows. Section 4.1 introduces the background of neural stimulation and the physicochemical properties of the electrode-electrolyte interface. Section 4.2 gives an overview of the stimulator design. The key stimulator design requirements are summarized, and previous state-of-the-art techniques are reviewed. Section 4.3 presents a general-purpose programmable neural stimulator design. Section 4.4 describes the novel net-zero charge neural stimulator design. The circuit implementation and the experimental results are presented. Finally, section 4.5 concludes this chapter.

4.1.1 Background of Neurostimulation

Neurostimulation is a method for modulating the nervous system's activity using non-invasive or invasive means [229]. The controlled electrical, magnetic [230, 231], chemical [232, 233], or optical stimulation (optogenetic modulation [234, 235]) of the central or peripheral nervous systems is usually referred to as *neuromodulation* in the medical literature [236]. The main focus of this work is the electrical stimulation.

The mechanism of electrical neurostimulation is a consequence of the depolarization and hyper-polarization of excitable cell membranes from the applied electrical currents. However, other mechanisms including thermal and neurohumoral effects may also involve with the process. The neuron membrane acts as a capacitor by separating the charges lying along its interior and exterior surfaces. The membrane conductance depends on the densities and types of the ion channels. The channels are highly selective, allowing only a single type of ion to pass. The membrane also contains selective pumps that expend energy to maintain the differences in the concentration of ions inside and outside the cell. A neuron will typically fire an action potential when its membrane potential reaches a threshold voltage. A simplified integrate-and-fire model can mimic this mechanism. The entire membrane conductance is modeled as a single term:

$$i_m = g_L(V - E_L) \tag{4.1}$$

where g_L is the transconductance, V is the membrane potential, E_L is the equilibrium potential. The membrane potential is determined by:

$$c_m \frac{dV}{dt} = -(V - E_L) + \frac{I_e}{A} \tag{4.2}$$

where c_m is the membrane capacitance, I_e is the injected current, and A is the surface area. Eq. 4.2 can be rewritten using the membrane time constant τ_m by:

$$\tau_m \frac{dV}{dt} = E_L - V + R_m I_e \tag{4.3}$$

After firing the action potential, the membrane potential is reset to E_L.

To better understand the process, this model is simulated in Matlab with a dynamically injected current. In this simulation, assume τ_m=10ms, R_m=$10^7\Omega$, the membrane potential threshold for firing the action potential is -50mV, and the time resolution is 10μs. The simulation result is shown in Fig. 4.1. The plot illustrates the relation between a dynamic stimulation current and the evoked action potentials. The model can also be modified for discrete simulation current pulses, and will be revisited in the study of the closed-loop control of neuromodulation in Chapter 5.

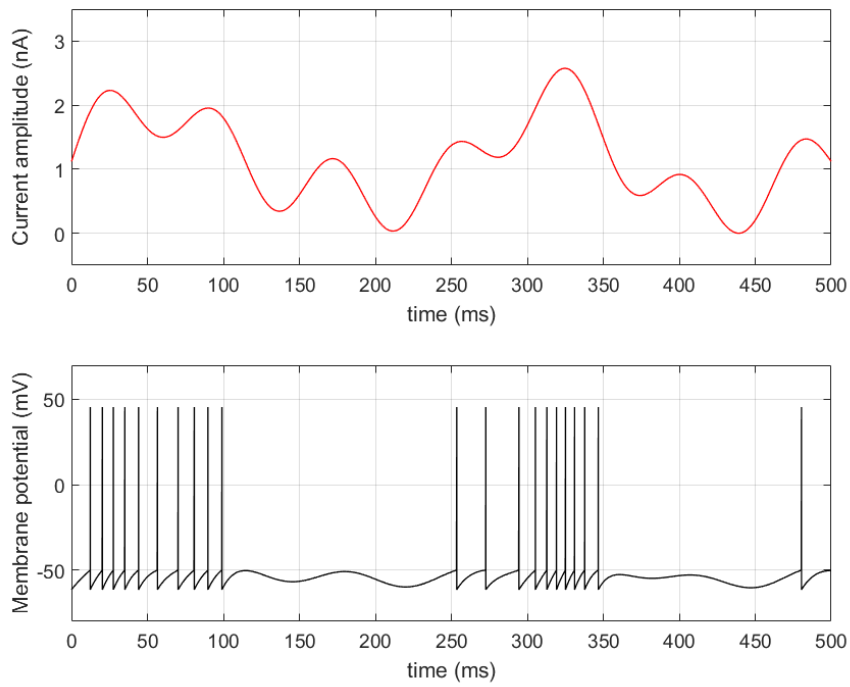


Figure 4.1: Matlab simulation of the membrane potential with dynamic simulation current based on the integrate-and-fire model.

4.1.2 Electrode and Electrolyte Interface

The essential process during an electrical stimulation is the charge transfer and redistribution across the electrode and electrolyte interface. It should be noticed that in the metal electrode and the electrical circuits, the charges are carried by the electrons; while in the physiological medium, the charges are carried by the ions, mainly including sodium, potassium and chloride. Fig. 4.2 (modified from the reference [20]) (a) illustrates the two primary mechanisms:

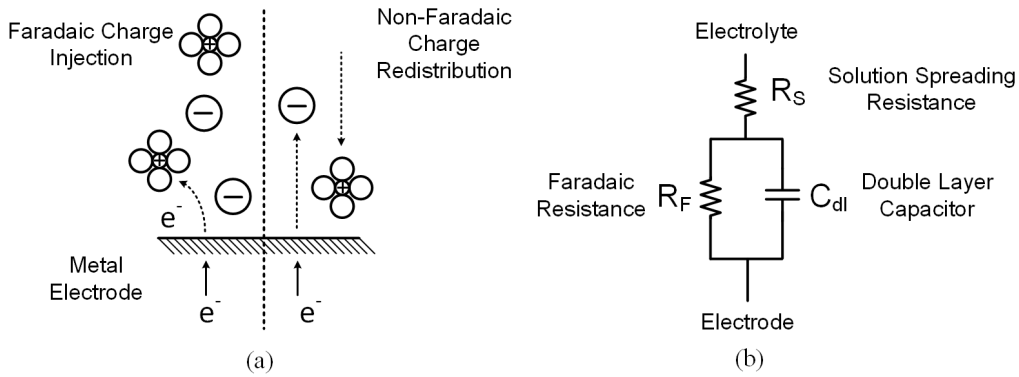


Figure 4.2: Illustration of the electrode and electrolyte interface, modified from [20]. (a) The physical representation, and (b) a simplified electrical circuit model.

- **Faradaic charge transfer**, or non-polarizable mechanism, where electrons transferred between the electrode and electrolyte interface causes reduction and oxidation reactions. The Faradaic reaction may be reversible or irreversible.

- **Non-Faradaic charge redistribution**, or polarizable mechanism, where a double layer capacitor C_{dl} is formed on the surface of the electrode, and the stimulation process involves charging and discharging the C_{dl} without direct electrons transfer [20, 237].

The accurate modeling of the electrode impedance, however, is a rather complicated task [238]. A simplified linear model, modified from [20], is adopted in the analysis of this work. The model has been widely used in neural interface research, and proves to be sufficient in estimating of the properties of the electrodes used for neural recording and stimulation [92, 239, 240].

To verify the model, the impedance of two types of low-cost tungsten electrodes commonly used in this research were measured. Ten electrodes of each type were measured in 0.9g/100mil Sodium Chloride. The measurement results are shown in Fig. 4.3. The linear model is used to fit the measurement results. The electrode

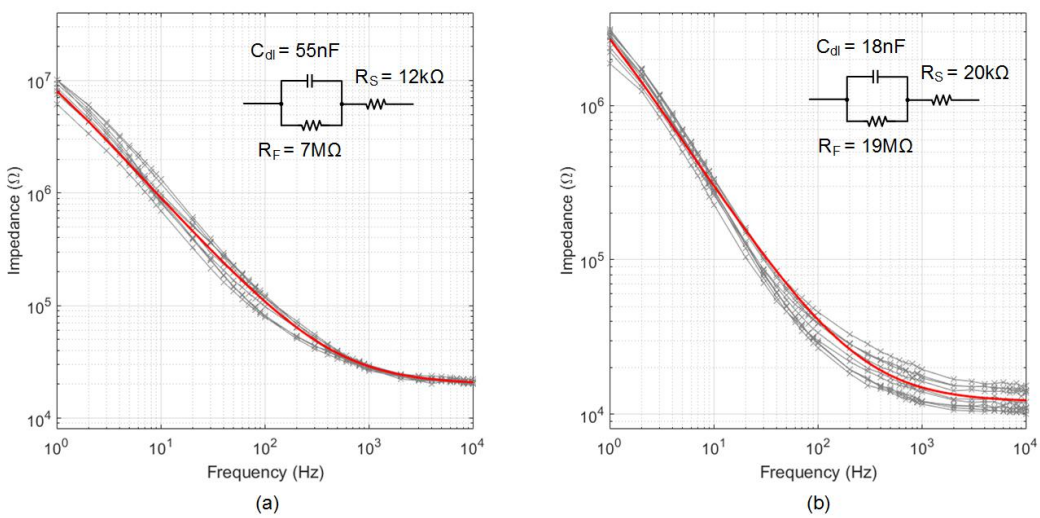


Figure 4.3: The measured impedance of two types of custom made tungsten electrodes used in this research. (a) and (b) shows the electrode with a diameter of 75μm and 50μm, respectively. Each figure shows an overlay of the measurements of 10 electrodes, and a fitting curve in red. The parameters of the fitting models are labeled in the figures.

with a diameter of 75μm has an average C_{dl} of 55nF, R_F of 7MΩ, and a spreading resistance of 12kΩ. The electrode with a diameter of 50μm has an average C_{dl} of

18nF, R_F of 19MΩ, and a spreading resistance of 20kΩ. The electrode with the smaller contact area gives a higher impedance in general. This measurement result gives a good insight of the electrode characteristics, and is used in several of the following studies.

4.2 Overview of Electrical Stimulator Design

Lots of electrical simulation techniques have been developed to produce the charges needed to recruit a neural response. An ideal stimulator triggers the desired neural response with minimum injected charges, and leaves no residue charge. However, the ideal stimulation is not always achievable. When it comes to the electronics design, the safety, the power efficiency, and the circuit performance all need to be taken into consideration. The design trade-off becomes more difficult for an implantable device which requires high channel count, minimum chip area, and low power density. Fig. 4.4 highlights the trade-offs in the neural stimulator design.

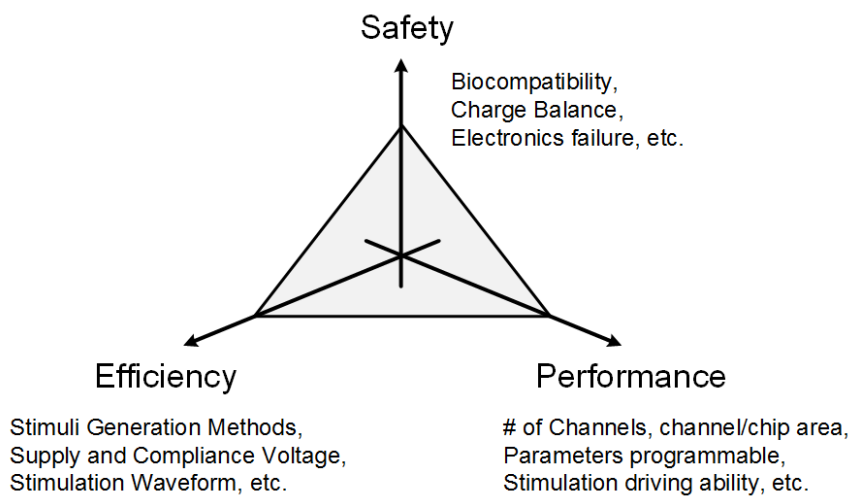


Figure 4.4: The design considerations and trade-offs of an electrical neural stimulator.

Primarily, a stimulator design should take the safety as the top priority. A safe long-term stimulation requires the stimulator to give a charge balanced stimulation without dc current injection. The prior study shows a current leakage of 100nA will cause a permanent damage to the tissue [118]. Secondly, the performance requirements of the stimulator mainly include the number of channel count, the occupied

silicon area, the programmability of the stimulation parameters, the stimulation current driving ability, and so on. The third dimension is the efficiency of the stimulator. The overall efficiency should consider both the power efficiency of the stimulator for generating the stimuli, and the efficiency of the stimuli for triggering the desired neural response. However, the latter is much harder to be quantized. Among all methods for generating the stimuli, voltage, current and charge regulation all have pros and cons, and different stimulation waveforms will also lead to different efficiencies. A very high power efficiency neural stimulator design may not give the best charge-balance performance, a sophisticated charge cancellation technique may not be suitable for a high channel-count integration, and a high channel-count design may not allow all parameters to be programmable. An aggressive optimization on one dimension might cause drawbacks in the other two dimensions, and eventually, makes the overall system design non-practical. There is no best universal stimulator design, but good designs for certain applications.

4.2.1 Methods of Stimuli Generation

In general, the neural stimuli is generated from the electronics by regulating the voltage, the current, or the total amount of charges. Essentially, it is the charges that disturb the membrane equilibrium and evoke the neural response [20]. However, different generating methods give different levels of control of the charges. Of course, a high controllability usually comes with costs in circuit complexity and power consumption. The pros and cons of each method are summarized as below.

A. Voltage-Regulated Stimulation

In the voltage-regulated stimulation, a certain stimulus voltage is applied between two electrodes (or between one electrode and the common tissue ground). Current passes through the electrodes depending on the tissue and electrode impedance. Since the circuitry has no control over the total amount of injected charges, it is difficult to achieve a charge-balanced stimulation. In the clinical use, the tissue impedance is usually well documented across patients, and the physicians will assign a proper stimulating voltage to achieve the desired neurophysiologic response.

A voltage controlled stimulator usually has a high overall efficiency and simple circuitry. However, it is poor in the controllability of charge injection and thus lacking of safety. It has been used in high-density applications like retinal implants, and power hungry clinical use including pacemaker and deep brain stimulator. The circuitry implementation of the voltage-regulated stimulation, and techniques for improving its safety has been reported in literature [52, 241, 242].

B. Current-Regulated Stimulation

In current-regulated stimulation, a certain stimulus current is passed between two electrodes (or between one electrode and the common tissue ground). The compliance voltage between the two electrodes depends on the tissue and the electrode impedance, and is limited by the supply voltage of the system and the circuitry implementation. The total amount of injected charges can be controlled by the stimulus current and the stimulating time, and charge balanced stimulation can be well achieved. Monophasic current stimulation [56] and biphasic current stimulation [9, 243, 244] have been reported in literature.

The current-regulated stimulation is the most widely used topology in electrical stimulator designs, given its high controllability of charge injection and the high safety. However, current-regulated stimulator has relatively poor efficiency. The circuitry implementation of the current-regulated stimulator, and techniques for improving its efficiency has been reported in literature [9].

C. Charges-Regulated Stimulation

In a charge-regulated stimulation, a capacitor tank is connected to one electrode and discharged to a reference electrode. The discharging current is used to excite the tissue. The circuitry implementation of charges-regulated stimulator has been reported in literature [245, 246].

The charge-regulated stimulation potentially can achieve both a high power-efficiency and a good controllability of the total amount of injected charges. However, the discharge time constant is still not well controlled, and the implementation of storage capacitors takes large silicon area or has to be implemented off-chip.

4.2.2 Stimulation Waveform and Electrodes Configuration

Various stimulation waveforms have been used in research and clinical treatment. Among them, biphasic stimulation is the most commonly used method. A typical biphasic stimulation mainly consists of a cathodic (stimulation) phase and an anodic (reversal) phase. The cathodic phase is to elicit the desired physiological effect such as initiation of an action potential, and the anodic phase is used to reverse the electrochemical processes occurring during the cathodic phase. It should be noticed that the

cathodic-first topology is chosen because the electrons move in the opposite direction of the current. Thus pulling a cathodic current is, in fact, pushing the electrons into the tissue.

The complete waveform with key parameters marked is shown in Fig. 4.5. A

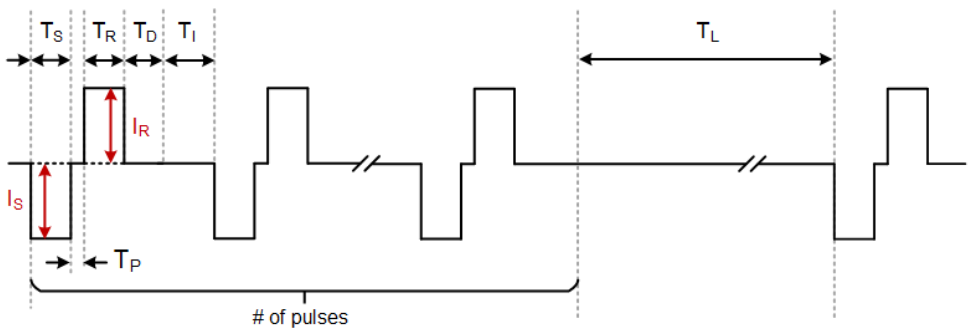


Figure 4.5: Illustration of a typical biphasic stimulation waveform with the parameters marked. I_S: stimulation current, I_R: reversal current, T_S: stimulation phase time, T_R: reversal phase time, T_D: discharging phase time, T_P: phase interval, T_I: pulse interval, T_L: pulse group interval.

constant current-regulated method is used here for illustration, but all the other methods share similar parameters. I_S and I_R are the amplitudes for the stimulation and reversal phase, respectively. If the same amplitude is used for both phases, the waveform is referred to as symmetrical biphasic stimulation. In some cases, a lower amplitude is preferred in the reversal phase to reduce the damage to the tissue. Both symmetrical [243, 244] and asymmetrical [9] biphasic current-regulated stimulation have been reported in the literature. T_S and T_R are the times for the stimulation and reversal phase, respectively. T_P is the interphasic delay between the stimulation and reversal phases. The interphasic delay is added for better stimulation effect. A discharging phase T_D is added following the reversal phase for removing the residue charges. In some cases, the anodic phase is replaced by the discharging phase. T_I is

the pulse interval, and T_L is the interval between the pulse groups. The terminologies are used consistently in the following study.

Monophasic/biphasic stimulation can be confused with monopolar/bipolar stimulation. However, they are different terminologies and are not directly related. A monopolar stimulation means that both cathodic and anodic phases are generated from a single electrode, while a bipolar stimulation means that the cathodic and anodic phases are generated from a pair of electrodes. Both electrode configuration can be used to perform monophasic and biphasic stimulation. Fig. 4.6 illustrates the typical electrode waveforms for the two-electrode configuration in generating a biphasic stimulation. In the monopolar stimulating method, one working electrode is used to

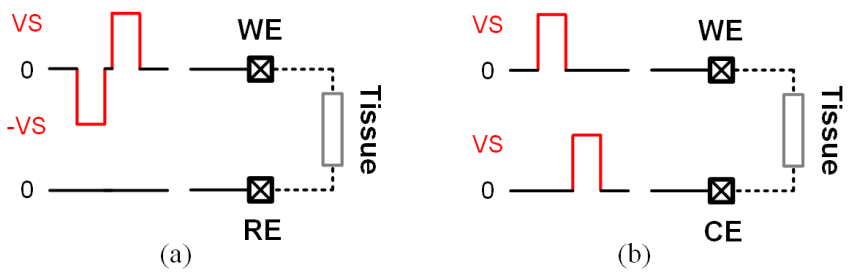


Figure 4.6: Illustration of (a) monopolar and (b) bipolar stimulation methods. A voltage-regulated stimulation is used for illustration, but a current-regulated stimulation can be applied in the same way.

generate the stimulus voltage or current with respect to a reference electrode. In the bipolar stimulating method, two electrodes are chosen as the working and counter electrodes, the generated stimulus voltage or current are to be applied between these two electrodes.

The monopolar stimulation is widely used in a high-density electrode array, in which case it delivers stimulus with respect to a common reference electrode. Bipolar stimulation has a better guided stimulus orientation than the monopolar stimulation,

at the cost of a more complicated channel selection. The bipolar stimulation also favors the single supply system, and potentially doubles the compliance voltage range. Ideally, both monopolar and bipolar configurations can achieve the charge-balanced stimulation.

It should be noticed that even though we have been discussing rectangular stimulation waveforms, non-rectangular waveforms have also been proposed in literature [22, 246, 247]. By careful design, the non-rectangular waveform, like exponential current stimuli, may give benefits in the stimulation effects and the power efficiency. Of course, these designs usually come at the cost of the circuit and control complexity. More importantly, it may cause difficulties in achieving the charge balance, which will be discussed in the following section.

4.2.3 Methods for achieving Charge Balance

The importance of the charge balance cannot be overemphasized. The building-up of the excess charges, even slowly, might cause toxic effects and lead to permanent damage. The traditional method is to place a blocking capacitor in series with the stimulating electrode. The blocking capacitor limits the total charges. However, the capacitor cannot be too small which limits the output compliance voltage range. Typically for functional electrical stimulation, these capacitors are in the order of tens of nanofarads to a few microfarads [22, 248–250]. The physical dimension of these capacitors is usually prohibitively big to be integrated on a silicon chip, especially for the high channel-count design. Various techniques have been developed to achieve the charge balance with and without the blocking capacitors. This section reviews the pros and cons of these techniques.

4.2.3.1 Matching and Calibration

Ideally, a charge balance can be achieved if the total amount of charges of the cathodic and the anodic phases are the same. The clock can achieve a high accuracy, but the mismatch between the anodic and cathodic currents can be about 2% without calibration even with careful matching in the design and layout [251]. This current mismatch might lead to a significant charge error in a heavy-duty stimulation. For that reason, a lot of research has been conducted with a focus on matching the stimulation and reversal currents, in order to achieve the net-zero change. Several important matching techniques are reviewed as follows.

J. Sit *et al.* from Massachusetts Institute of Technology proposed blocking capacitor free change-balanced stimulator design in 2007 [240]. The design uses a dynamic current balancing method to achieve current balance. The work pays special attention to switch leakage and loop stability in the dynamic current mirror. The reported DC current error is 6nA.

K. Song *et al.* from Korea Advanced Institute of Science and Technology proposed a DC-balanced adaptive stimulator in 2012 [252]. The design uses a current sources mismatch compensation method. Precise current balance is achieved by sampling the mismatch current, and making a compensation accordingly. The challenge in this work is the requirement of a large time constant sample hold circuit. The S/H circuit would need to hold the mismatch current ΔI for up to 0.5 seconds without variation. The reported current mismatch is less than 10nA.

M. Monge *et al.* from the California Institute of Technology proposed a high-density self-calibration epiretinal prosthesis in 2013 [253]. This work uses a fully digital calibration technique to match the biphasic currents during the stimulation.

A multi-point calibration scheme is proposed, which includes 5-point calibration for each driving site. With the help of the full range calibration, the stimulation can perform arbitrary waveform stimulation. The reported mismatch is 2.24%.

In summary, the matching technique can be implemented in either analog or digital domains. In the analog domain, it requires feedback or a large time-constant storage unit. In the digital domain, it requires an on-chip memory. The analog matching is attractive if only a single-point matching is needed. If a full-scale calibration is desirable, the digital calibration is more suitable.

4.2.3.2 Passive and Active Discharge

If the current matching is not sufficient for achieving the net-zero charge requirement, an additional discharge phase is commonly used to remove the residue charges. The discharge can be as simple as shorting the stimulation electrode to a common or reference electrode, which is referred to as passive discharge. However, the disadvantage of passive discharge is that the discharge current depends on the load and electrode impedance, which cannot be well controlled. If the impedance is too low, the discharge current might be too large, it might damage the tissue, thus additional current limit circuit is required [244]; if the impedance is too high, the time for discharging might be too long, so the residue charges might not be able to clear before the next stimulus, and the residue charges will accumulate.

K. Sooksood *et al.* from the University of Ulm proposed an active charge balance method in 2010 [254]. In the active approaches, the residue charges or net potential is monitored by the active circuits, and additional discharge circuits are added to maintain the net potential in a safe range in a closed-loop manner. The residue

charges can be canceled by using pulse insertion [254, 255]. In addition, E. Noorsal *et al.* from the same group proposed to regulate the residue charges by using DC biasing current in 2012 [256]. In this work, a safe window is defined approximately as 100mV (for a Pt black electrode). The net potential is compared with the safe windows right after the stimulation, and the biasing current sources can be adjusted accordingly.

In summary, passive and active discharge or charge cancellation can be used to further remove the residue charges after the stimulation. Passive discharge is simple but has no control of the discharging current and time. Active discharge by monitoring the residue charges directly can be more effective, but takes dedicated circuit design and silicon area.

4.3 Design of A General-Purpose Stimulator

This section describes the design of a 16-channel general-purpose neural stimulator. The motivation of this work is to have a highly programmable stimulator for various applications. The stimulator can perform monopolar or bipolar, monophasic or biphasic, symmetrical or asymmetrical constant-current, charge balanced stimulation. All of the parameters for the stimulator are programmable. The output current is from 0 to $\pm 255\mu$A in the low-current mode, and 0 to $\pm$2mA in the high-current mode. The design has been fabricated in IBM 180nm technology, and occupies a silicon area of $810\mu\text{m}\times 290\mu\text{m}$, excluding the IO pads.

4.3.1 Architecture of the Stimulator

The overall architecture of the stimulator is shown in Fig. 4.7. The stimulator includes a digital part designed in 1.8V, and an analog part designed in 1.8/5V. The stimulator integrates four independent driving sites. Each site includes: i) a DAC to generate a reference for the output current, ii) a current driver consisting of current sink and source output stages with high output impedance, iii) a 1:4 demultiplexer to support 4 channels and provide near-simultaneous stimulation, and iv) level-shifters to interface the low-voltage digital control signal with the high-voltage switches.

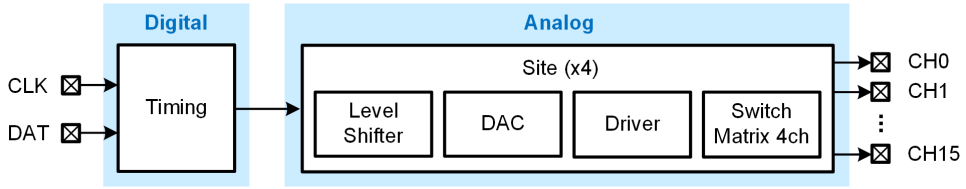


Figure 4.7: The block diagram of the neural stimulator. The stimulator includes a digital part designed in 1.8V, and an analog part designed in 1.8/5V.

The digital part can be configured via a custom designed two wire protocol. The parameters of the timing generation module in each stimulator site can be programmed individually. In addition to the regular operating modes, the stimulator can be configured to output continuous current in order to test the DAC and the output stage.

4.3.2 Circuit Implementation

The circuit schematic of the stimulator site is shown in Fig. 4.8. A 6-bit current

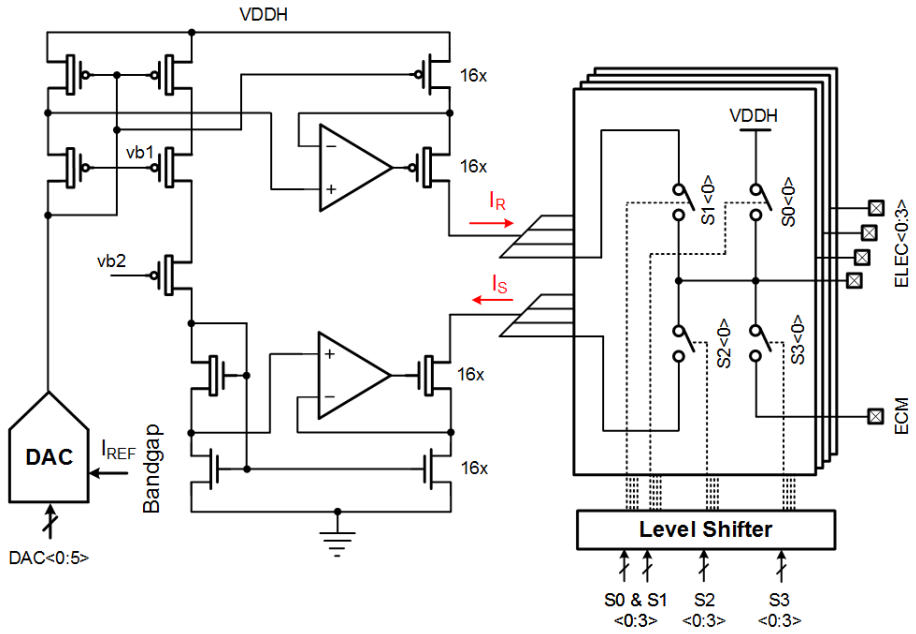


Figure 4.8: The circuit schematic of the proposed multi-mode stimulator site. Each site consists of: i) a current-mode DAC which generates a reference for the output current, ii) a current driver including current sink and source output stages with high output impedance, and iii) high voltage switches with level-shifters. Each site demultiplexes to 4 channels, and provides near-simultaneous stimulation.

mode DAC is used to generate the reference for the stimulation current. A typical binary weighted current source array is used in the DAC [257]. The transistors are

sized for a 6-bit accuracy [194]. Common-gate transistors are used to increase output impedance. The DAC is designed in thin-oxide devices and powered at 1.8V. A thick oxide transistor is cascaded in the output current path to reduce the overdrive voltage stress from the following stage. The thick oxide transistor also has low leakage current which allows a complete shutdown of the DAC.

The output stage is designed with thick-oxide devices and a supply voltage of 5V. Regulating amplifiers are used to achieve a high output impedance. A PMOS input folded-cascode amplifier is used in the current sink, and an NMOS input folded-cascode amplifier is used in the current source. These amplifiers are disabled when the stimulator is in the idle mode to reduce the power dissipation.

The circuit schematic of the level shifter is shown in Fig. 4.9. It should be noticed

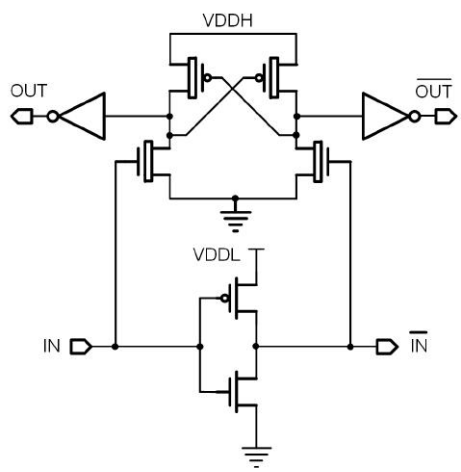


Figure 4.9: The circuit schematic of the level shifter.

that even some dynamic level shifter can achieve high switching frequency and lower power consumption, the risk of undetermined state may cause direct stimulation current leakage to the tissue, thus is not used in this design.

To perform the monopolar stimulation, one electrode is activated at a time, and the stimulation current is passed between the selected electrode and the ground electrode. The timing of the monopolar stimulation and the control of corresponding switches are shown in Fig. 4.10. The timing parameters is defined as in Fig. 4.5, and

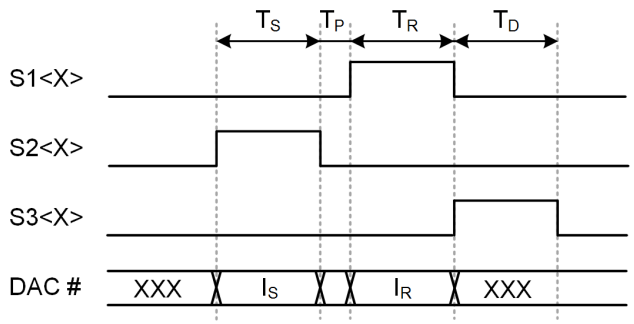


Figure 4.10: The timing for generating a monopolar stimulation at the electrode X. The DAC # belongs to the site where the electrode X locates. (I_S: stimulation current, I_R: reversal current, T_S: stimulation phase time, T_R: reversal phase time, T_D: discharging phase time, T_P: phase interval. XXX means the DAC can be in any value.)

the switching signals are defined as in the circuit schematic Fig. 4.8. XXX means the DAC can be of any value. This allows the DAC to generate the reference value for another channel.

To perform the bipolar stimulation, two electrodes can be arbitrarily selected from the 16 channels to work as the cathodic and anodic electrodes, and the stimulation current is passed between them. The timing of the bipolar stimulation and the control of corresponding switches are shown in Fig. 4.11. Notice that the electrode X and Y can be in the same stimulation site, or in two different stimulation sites. The DAC # and DAC $ will be the same DAC if X and Y are on the same site.

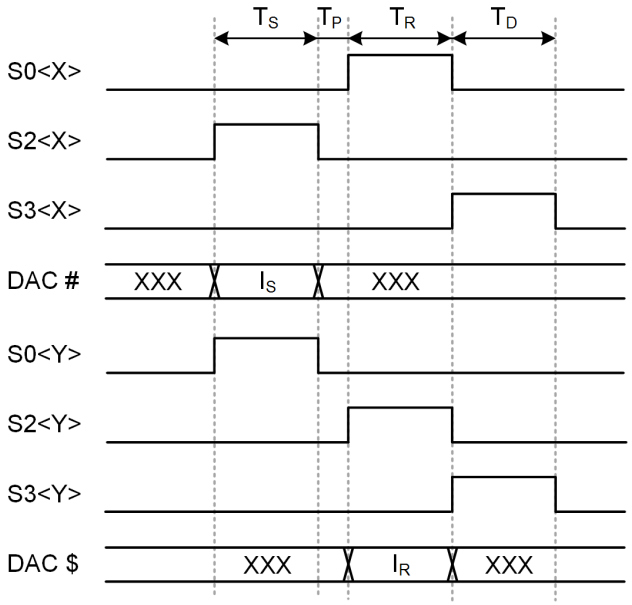


Figure 4.11: The timing for generating a bipolar stimulation between the electrode X and Y. The DAC # and DAC $ belong to the site where the electrode X and Y locates, respectively. Notice that they can be in a same site. Parameter definitions are the same as in Fig. 4.10.

4.3.3 Measurement Results

The design has been fabricated in IBM 180nm CMOS technology. The occupied silicon area is 810μm×290μm. The layout of the 16-channel stimulator is shown in Fig. 4.12. The major building blocks are highlighted in the figure.

Several bench tests have been conducted to fully evaluate the function and performance of the designed stimulator. Fig. 4.13 shows the measured output currents from the current sink and current source, with several digital input codes. The results show a large compliance voltage range with an overhead less than 264mV, corresponding to 5.28% of the supply voltage.

Fig. 4.14 shows the measured output currents from the current source and sink of the stimulator output stage. The non-linearity of the source and sink current is

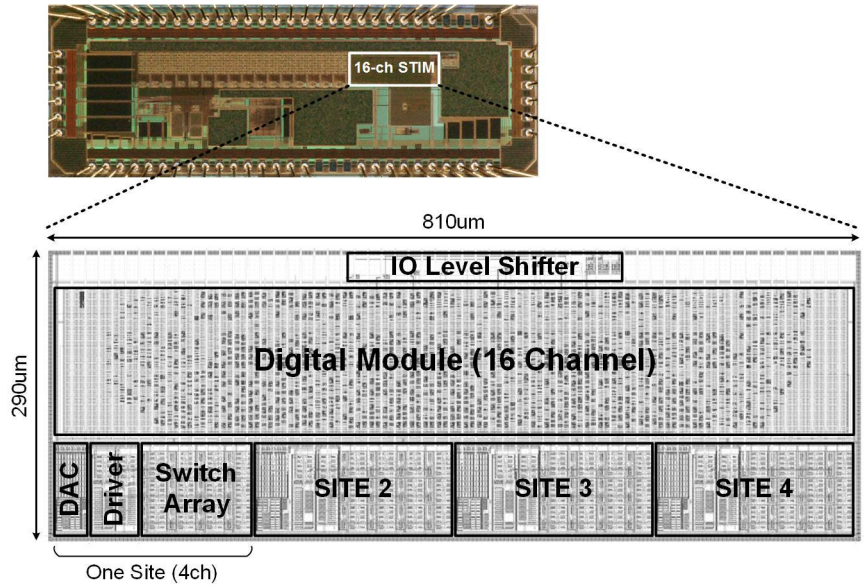


Figure 4.12: The microphotography and the layout of the neural stimulator. The major building blocks are highlighted in the layout.

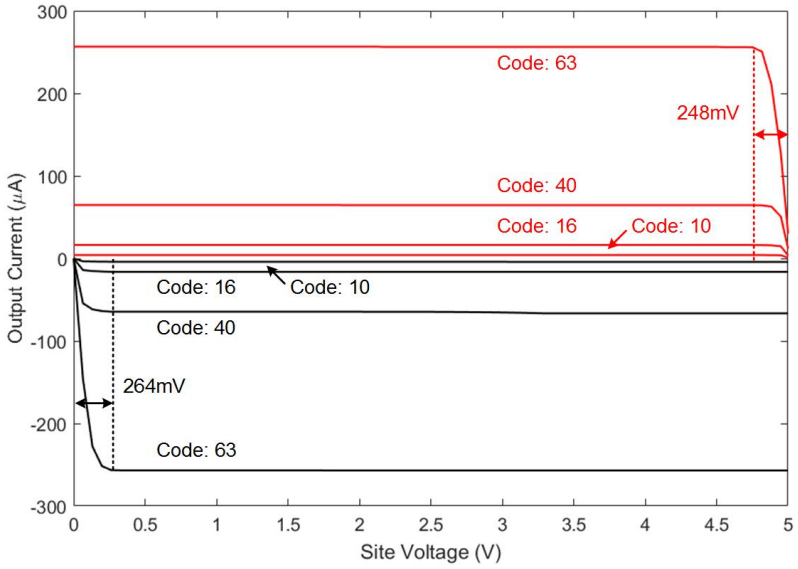


Figure 4.13: The measured stimulator output current versus output voltage.

0.31% and 0.37%, respectively. The result shows a good matching between the source and sink is 1.29% without calibration. No additional analog calibration is used in

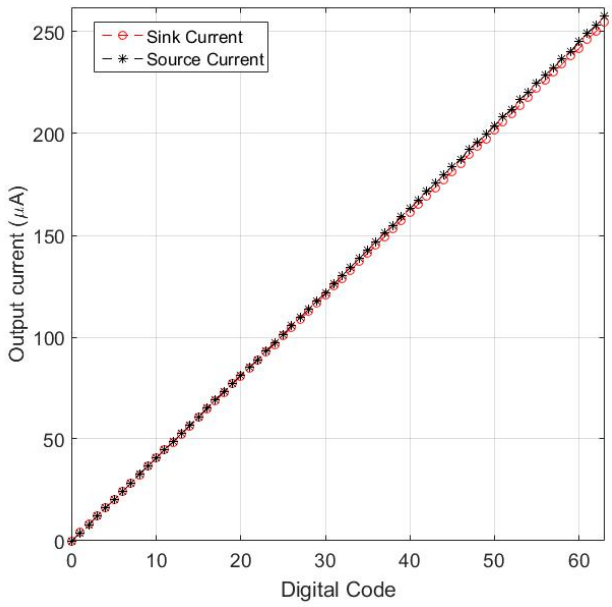


Figure 4.14: The measured output currents from the current source and current sink of the stimulator output stage. The non-linearity of the source and sink current is 0.31% and 0.37%, respectively.

this work, but a digital calibration in the digital code can be implemented for better matching. The discharging phase should always be used to avoid charge accumulation in this case.

The stimulation was measured in 0.9g/100mil Sodium Chloride. The measured simultaneous stimulation output from four independent channels is shown in Fig. 4.15. Different pulse train interval times were intentionally used for each individual channel, which showed the ability for this chip to drive simultaneous stimulation in different parameters.

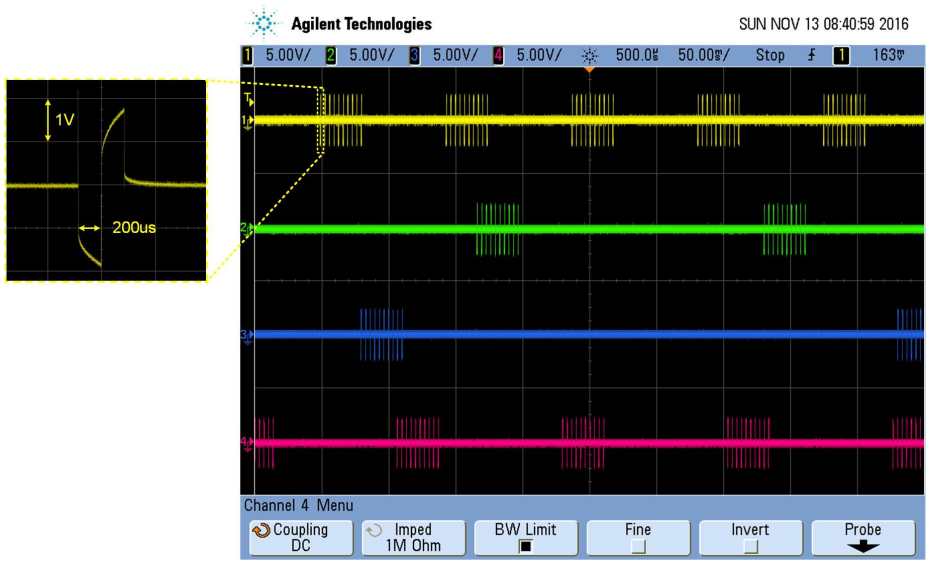


Figure 4.15: The measured simultaneous stimulation output from four independent channels. The boxed window shows the measurement of a single pulse in high resolution.

4.4 An Energy Efficient Net-Zero Charge Neural Stimulator

4.4.1 Introduction

As discussed in section 4.1.2, two primary mechanisms occur at the interface between the electrode and the physiological medium during an electrical stimulation: the direct Faradaic charge transfer and capacitive charge redistribution [20]. The Faradaic charge transfer usually involves reduction and oxidation processes, which may create damaging chemical species and dissolve the electrodes. So it is critical to avoid the onset of these reactions. A reversal phase is commonly used after the stimulation phase to reverse the electrochemical processes. However, it is not always possible to

avoid the irreversible charge injection, resulting in a certain amount of unrecoverable charges during the stimulation [20].

In order to achieve a net-zero charge, a lot of techniques have been developed, as reviewed in section 4.2.3. However, previous works have been exclusively focusing on matching the stimulation and reversal currents, ignoring the unrecoverable charge injection during the stimulating process. This work proposes a new stimulation strategy to achieve the net-zero charge by monitoring the residue charges directly on an inserted capacitor. With the proposed method, over-reversal can be avoided. Besides, a perfect matching between the current source and sink is not required, and an arbitrary stimulation waveform can be performed without calibration.

As reviewed in section 4.2.1, voltage-regulated [258], charges-regulated [259] and current-regulated [240, 252, 253] stimulation methods have been reported in literature. In summary, the voltage-regulated stimulation method has the highest efficiency, but it is difficult to control the amount of injected charges [243]. The charges-regulated stimulation limits the total amount of charges by discharging a capacitor tank, but the capacitors cost a large silicon area, and the discharging time cannot be precisely controlled. The current-regulated stimulation has a high controllability of the charge injection, thus it is the most widely used method. However, the traditional current-regulated method suffers from a low power efficiency [252]. In this work, an adaptive driving voltage is enabled by employing a feedback control scheme to improve the power efficiency. The design also enables a constant low supply voltage design for all active circuits besides the driving voltage.

The remaining of this section is organized as follows. Section 4.4.2 highlights the innovations proposed in this work. Section 4.4.3 describes the system architecture

and the circuit implementation of the building blocks, with emphasis on the large voltage compliance output stage and feed-forward compensation comparator design. Experimental results are presented in section 4.4.4.

4.4.2 Motivation and Innovation

4.4.2.1 Net-zero Charges Stimulation

As discussed in section 4.2.2, biphasic stimulation is the most commonly used stimulation waveform. During a biphasic stimulation, a stimulation phase first elicits the desired physiological effect (e.g. initiation of an action potential), and after an optional interphase delay, a reversal phase is used to reverse electrochemical processes [20]. The threshold current required to initiate the neural response decreases with an increasing stimulation pulse width. The threshold and pulse width relation, which can be experimentally quantified, is usually presented as a strength-duration curve [260]. Although it is not a physiological requirement to design the reversal current equals the simulation current, it is commonly used in the circuit design for a better matching. Although a very high current matching accuracy has been reported in literature [240, 252], these methods often ignore the fact that inevitable charge diffusion may put the matching in vain. The process is illustrated in Fig. 4.16. Take the irreversible reaction and the chemical products diffusion into account, even perfect matched cathodic and anodic currents will still leave residual charges. These residual charges will accumulate in a simulation pulse train, resulting in a more serious damage if a discharge procedure is not properly assigned. This work addresses this problem by monitoring the residue charges on an inserted blocking capacitor. The reversal phase terminates when a net-zero charge point is reached, as illustrated in

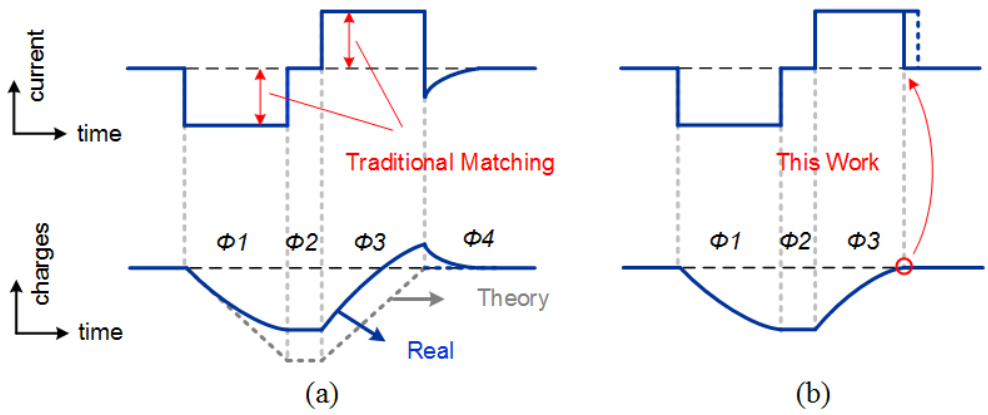


Figure 4.16: (a) The traditional charge-balancing method matches the stimulation and reversal currents. The ideal charge curve on the electrode is plotted in a dashed line. The practical charge curve varies from the ideal curve due to the irreversible reaction and chemical products diffusion. (b) this work terminates the reversal phase based on the monitoring the net-zero charge point. $\phi 1$: stimulating phase, $\phi 2$: interval phase, $\phi 3$: reversal phase, $\phi 4$: discharge phase

Fig. 4.16 (b). In this way, the systematic over-reversal in traditional methods can be avoided, and an exact matching between the stimulation and reversal currents is no longer required.

To better illustrate the effects, a simplified linear simulation model is established using ideal components, including the current sources. A typical single-supply bipolar stimulation topology is used in this simulation [21, 22]. The circuit schematic of the simulation model is shown in Fig. 4.17. A resistor R_L is used to mimic the impedance between the stimulation location and the tissue ground. The electrode parameters measured *In-Vitro* (Fig. 4.3) are used in this simulation. 500kΩ, 1MΩ, and 5MΩ resistors are used as R_L. It should be noticed that in practice, the electrochemical reaction is involved, thus the charge reduction is more complicated and is not linearly dependent on the electrical potential. Fig. 4.18 shows the simulated voltage across the blocking capacitor C_B. With R_L greater than 5MΩ, the charge diffusion due to

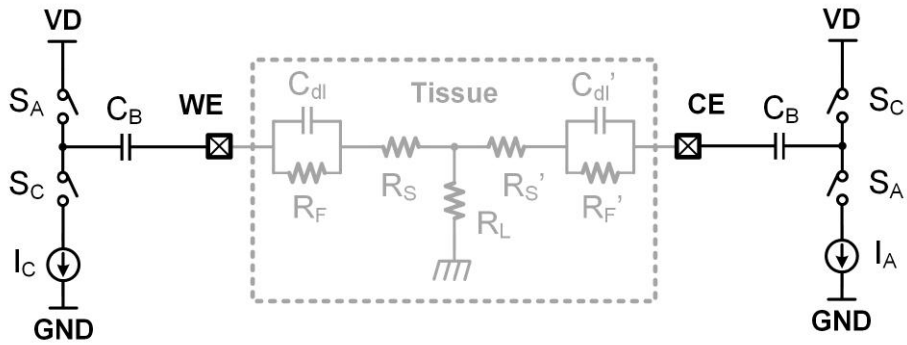


Figure 4.17: The model for simulating the effects of non-ideal charge diffusion.

the R_L is negligible. However, with a low R_L, the over-reversal is visible. Fig. 4.19

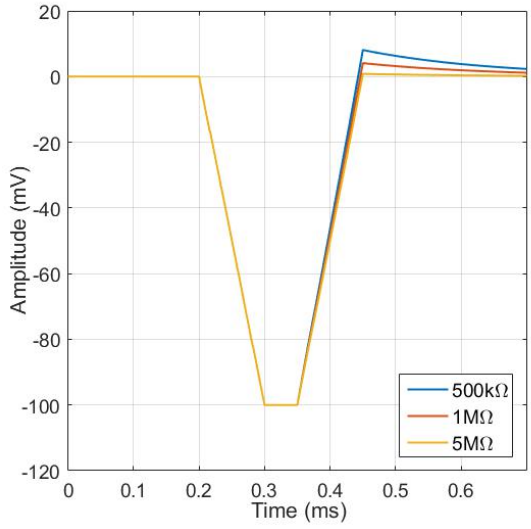


Figure 4.18: The simulation of the voltage across the blocking capacitor C_B (Fig. 4.17) during a symmetrical single pulse. Ideal current sources with equal stimulation current amplitude and time are used. 500kΩ, 1MΩ, and 5MΩ resistors are used to mimic the impedance between the stimulation location to the tissue ground.

shows the same simulation setup but in a pulse train without discharging. Even with the ideal current sources with equal amplitude and time, the charges still build up on the electrode. The building up of the charges may cause permanent damage to the tissue in a long term.

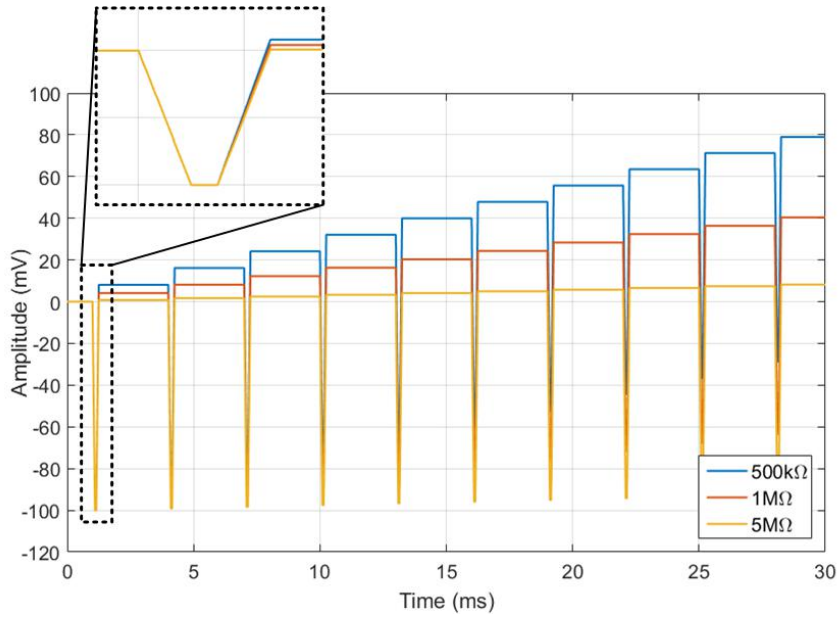


Figure 4.19: The simulation of the voltage across the blocking capacitor of a 10 symmetrical pulse train without discharging. 500kΩ, 1MΩ, and 5MΩ resistors are used to mimic the impedance between the stimulation location to the tissue ground. The charges build up even using the ideal current sources with equal amplitude and time.

Notice that the blocking capacitor is commonly used to ensure the safety by preventing the direct current injection and limiting the maximum net charges. Thus this work doesn't require a major change of the configuration in the conventional stimulator designs.

4.4.2.2 Adaptive Driving Voltage

In the simplified linear model, the compliance voltage for a charge balanced biphasic stimulation can be expressed as:

$$V_C = 2R_S I_S + \frac{I_S T_S}{C_{dl}} \tag{4.4}$$

In conventional designs, the supply voltage of the stimulator is set to be higher than the peak compliance voltage with headroom to avoid cut-off. However, the uncertainty and the drifting of the electrode impedance makes it difficult to predict the peak compliance voltage. So the supply voltage needs to be over-designed to guarantee a sufficient compliance voltage. As a result, a lot of power is wasted in the circuitry headroom instead of on the load tissue. The overall efficiency of the system is:

$$\eta = \frac{P_{load}}{P_{load} + P_{circuits}} = \frac{I_{stim}^2 Z_{Tissue}}{(I_{stim} + I_{circuits})V_{supply}} \tag{4.5}$$

In this work, an adaptive driving voltage instead of a constant high supply voltage is used for improving the power efficiency. In contrast to the conventional output stage design which includes both current sink and source, this design only uses the current sink. The working electrode (WE) sinks the current, and the counter electrode (CE) only needs to generate a potential difference with respect to the WE [21, 22]. This operation is illustrated in Fig. 4.20 (a). As a result, all circuits for the WE can be

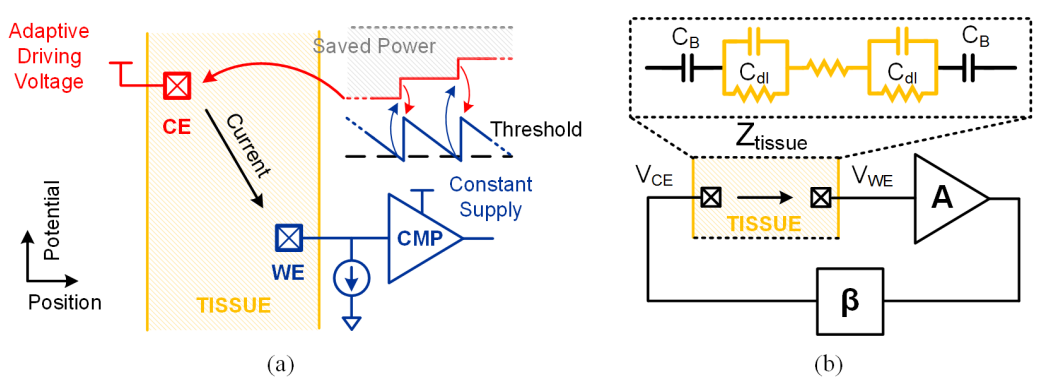


Figure 4.20: Illustration of the adaptive driving voltage stimulation.

designed in a low supply voltage, and only the driving voltage of the CE needs to be boosted. In the simplified linear model, the required counter electrode voltage V_{CE}

can be expressed as:

$$V_{CE} = V_{WE} + 2\frac{I_{stim}\Delta t}{C_{dl}} + 2\frac{I_{stim}\Delta t}{C_{bk}} + I_{stim}R_{tissue} \tag{4.6}$$

where R_{tissue} is unknown and varies from site to site. Assume the minimum biasing voltage for the current sink is V_{sink}, and the WE's potential should always be higher than V_{sink}. This gives an opportunity to design a feedback control scheme for the driving voltage by monitoring the WE voltage. A boosting converter for generating the driving voltage can be designed with a continuous tuning or several discrete output levels. When V_{WE} is lower than a pre-defined threshold V_{th} ($V_{th} > V_{sink}$), a 1-bit digital signal is generated to let the boosting converter's output increase in one step, in order to provide enough compliance voltage for the current sink. The process can be understood as a typical feedback control system, as shown in Fig. 4.20 (b). It should be noticed that the system is always stable, if the boosting converter's output is set to be the minimum value at the beginning, and only changes in one direction (increasing).

4.4.2.3 Arbitrary Channel Configuration

In order to achieve the best stimulation performance, the ability to perform stimulation in an arbitrary location and direction from the implanted electrode array is very helpful. It can fully take the advantage of the high-density electrode array. However, conventional stimulator design with the current source and sink matching technique can hardly realize an arbitrary channel configuration, since the matching is usually designed to perform between pre-defined electrode pairs. This is illustrated in Fig. 4.21 (a). With the proposed stimulation technique, a perfect matching is no longer

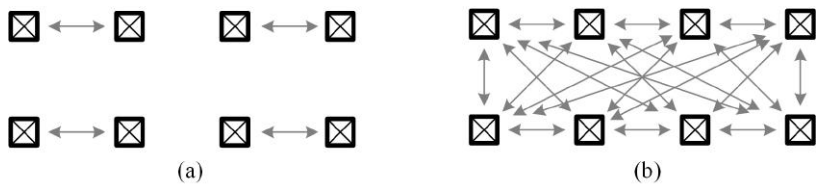


Figure 4.21: The channel configuration of (a) the traditional stimulation, and (b) the proposed work. Arbitrary channel configuration is feasible without pre-calibration.

required. Thus an arbitrary channel configuration of the working and counter electrodes is feasible, and more precision stimulation pattern can be generated from a limited number of electrodes, as illustrated in Fig. 4.21 (b).

4.4.3 Circuit Implementation

4.4.3.1 System Architecture

The architecture of the proposed stimulator system is shown in Fig. 4.22. The

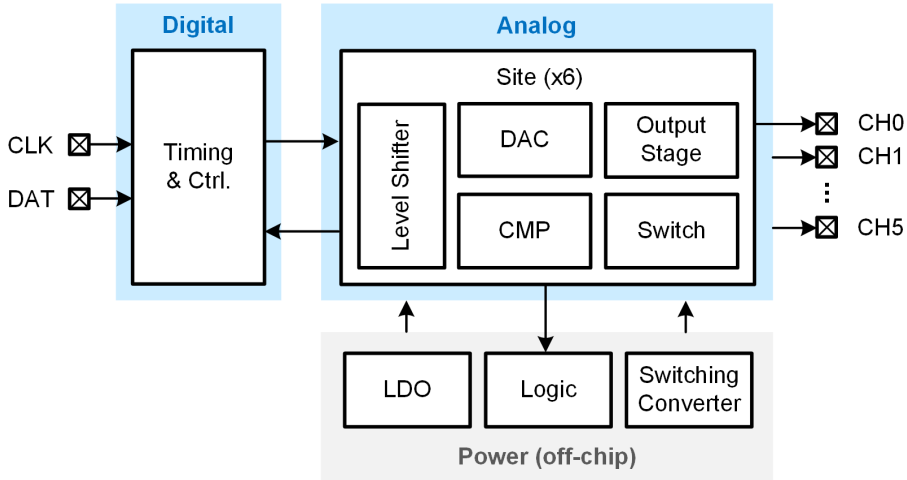


Figure 4.22: The block diagram of the net-zero charge neural stimulation system. The system consists of an analog core, a digital module, and off-chip power management units.

system contains 6 driving sites. Each site can be configured as the working electrode driver (WE mode) or the counter electrode driver (CE mode). The stimulation and reversal currents are generated by reversing the current path between the WE and CE. Each site contains a current sink with a high output impedance, with a 6-bit current mode DAC. Two comparators with different specifications are integrated into the driving site. The low-speed comparator is used to monitor the electrode voltage. The high-speed comparator is used to detect the zero-net charge crossing point, and to terminate the reversal phase. The functions of the digital module include: i) the output mode selection, ii) the output stage switch control, iii) the DAC and comparator configuration, iv) the supply voltage adjusting request generation, and v) the feed-forward comparator calibration.

The working flow of the proposed stimulation strategy is shown in Fig. 4.23. After all stimulation parameters are received, the system first selects and enables the selected WE and CE driving sites, and the DACs of the output stages are configured. The stimulation phase starts first, with a timer controls the stimulation time. The low-speed comparator monitors the compliance voltage and generates driving voltage adjustment signal accordingly. The stimulation phase terminates by the timer, and then the interphase timer starts. After the interphase, the reversal phase starts. The comparators are used to detect the net-zero charge point. The low-speed comparator is first used to perform the coarse detection, and it triggers a high-accuracy high-speed comparator when the voltage gets close. The reversal phase stops by the output of the high-speed comparator. After the reversal phase, an optional discharge phase can be used to clear the capacitor and amplify the residue voltage across the blocking capacitor. This phase is usually used in a training mode. If the residue voltage is beyond the safe range, a calibration value is adjusted and stored in the register. The

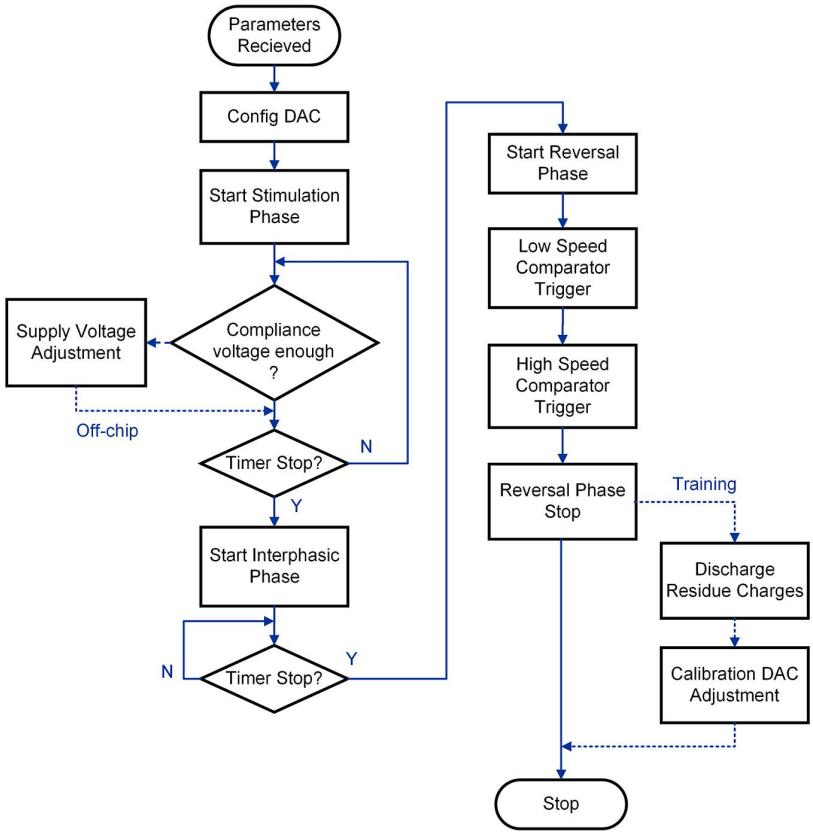


Figure 4.23: The working flow chart of the proposed stimulation strategy.

neural stimulation usually consists of a train of pulses with the same amplitude and pulse width. So the system will learn and save the calibration value for the foregoing stimulation.

4.4.3.2 Output Stage with Dynamic Element Matching

A high output impedance output stage with a high voltage compliance is critical for a neural stimulator. A transconductance amplifier (OTA) with series-series feedback can make a simple current generator with a high output impedance, as shown in Fig. 4.24 (a). The output current can be controlled by adjusting the reference voltage

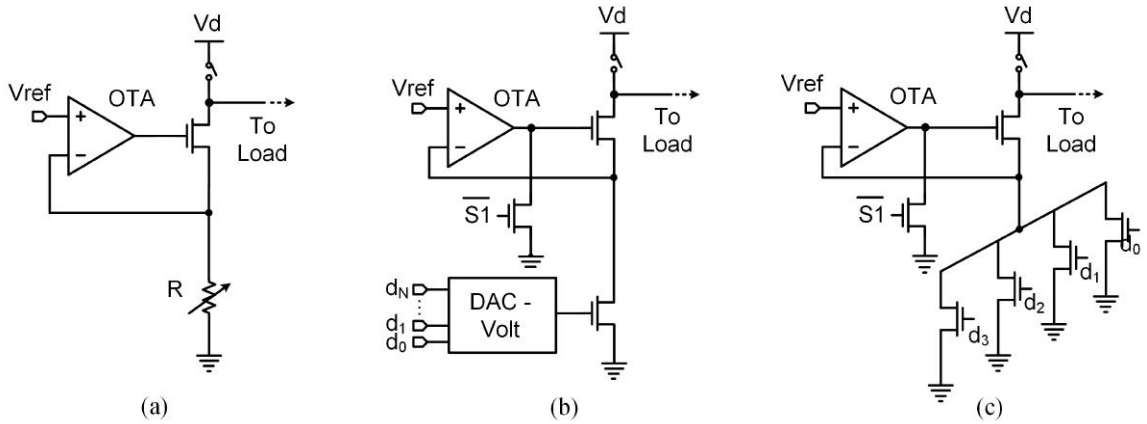


Figure 4.24: Circuit schematic of (a) basic current generator, (b) output stage with voltage-controlled transistor, modified from [21], (c) output stage with digital-set DAC, modified from [22].

or the tail resistor. The resistor can be implemented using a transistor biased in the linear region [21], or a batch of transistors biased in the deep triode region [22], as shown in Fig. 4.24 (b) and (c), respectively. A voltage mode DAC is used to bias the transistor to control the output current. However, the threshold voltage variation appears to be a problem, and it is especially important when the circuits are implemented for driving a micro-electrode array contains hundreds of channels. In this work, a current-mode DAC with dynamic element matching is used to address this problem.

Fig. 4.25 shows the simplified circuit of the proposed output stage. A 6-bit binary weighted DAC is used to generate output current. The transistor M1 is one-bit of the DAC, a dummy cell M2 is put on the side of M1. Instead of using digital signal or voltage mode DAC to bias M1, the gate voltage of M1 is generated by charging M2 using a reference current I_C. Thus the gate voltage of the M2 is given by:

$$V_C = \sqrt{\frac{2I_C L_2}{\mu C_{ox} W_2}} \tag{4.7}$$

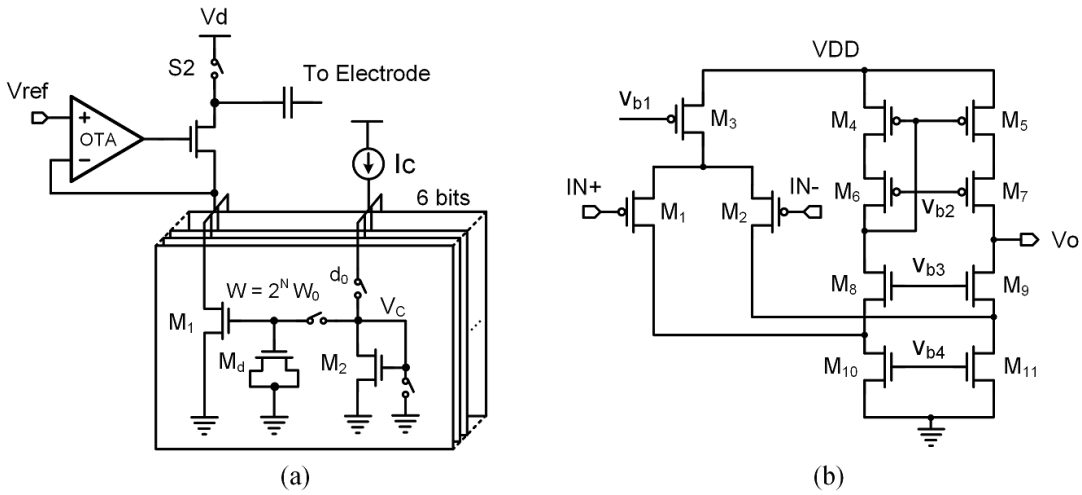


Figure 4.25: (a) The proposed output stage with current set current-mode DAC. (b) The OTA used in the work.

The drain current of M1 can thus be calculated as:

$$
\begin{aligned}
I_{d0} &= \frac{\mu_n C_{ox}}{2}\frac{W_1}{L_1}\left[2(V_C - V_{th1})V_{ref} - V_{ref}^2\right] \\
&= \frac{\mu_n C_{ox}}{2}\frac{W_1}{L_1}[2(\sqrt{\frac{2I_C L_2}{\mu C_{ox} W_2}} + (V_{th2} - V_{th1}))V_{ref} - V_{ref}^2] \\
&\approx \frac{\mu_n C_{ox}}{2}\frac{W_1}{L_1}[2\sqrt{\frac{2I_C L_2}{\mu C_{ox} W_2}}V_{ref} - V_{ref}^2]
\end{aligned}
\tag{4.8}
$$

where V_{ref} is the reference voltage set by the OTA, V_{th1} and V_{th2} are the threshold voltages of M1 and M2, respectively. Thus the output current will only depend on the threshold difference in the local area, instead of the threshold voltage itself, and the variation will be much smaller. The output current of the DAC is:

$$
\begin{aligned}
I_{dac} &= I_{d0} + I_{d1} + I_{d2} + I_{d3} + I_{d4} + I_{d5} \\
&= \sum_{n=0}^{5} 2^n d_n (A\sqrt{I_C} + B)
\end{aligned}
\tag{4.9}
$$

where d_n's are input digital codes, and A, B are constants controlled by design parameters and by V_{ref}. A 100-run monte-carlo simulation including both process corners and mismatch of the different structures of the output are shown in Fig. 4.26. The

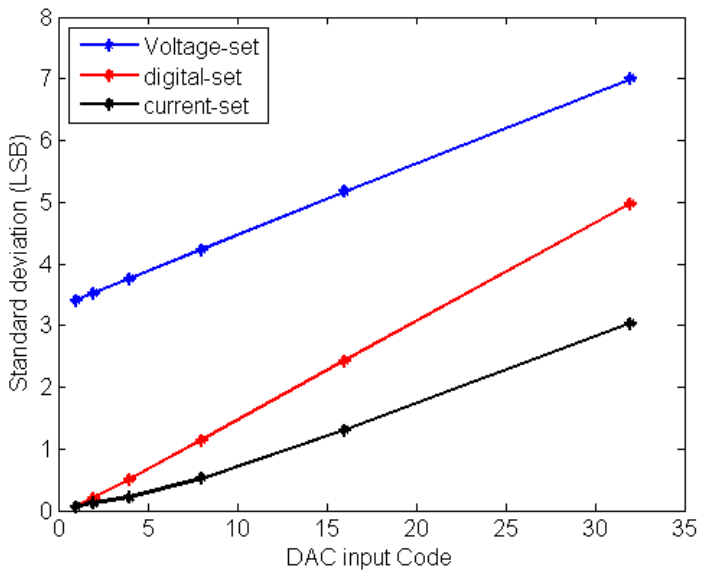


Figure 4.26: 100 runs monte-carlo simulation of the different output stage architectures with mismatch and process variation.

result shows that even with the worst variation, the proposed current-set dynamic element matching method reduces the output variation significantly.

4.4.3.3 Feed-forward Error Compensation Comparator

In the output stage, two comparators are connected to the blocking capacitor. A low-power, low-speed op-amp based continuous time comparator (LS comparator), and a high-speed high accuracy error compensation comparator (HS comparator). Both comparators have a shut-down option for saving the power consumption.

The HS comparator is designed with a 3-stage preamplifier and a dynamic latch. The offset of the comparator originates from the imperfect symmetrical layout and the variation during the fabrication. The CMOS latches implemented with small devices have larger offsets compared to the pre-amplifiers. The output offset auto-zeroing circuits are employed for the three-stage pre-amplifier to suppress the offsets in this work. The sources of the comparator delay include the charging time of the blocking capacitor, the converging of the pre-amplifier and the latch, as illustrated in Fig. 4.27.

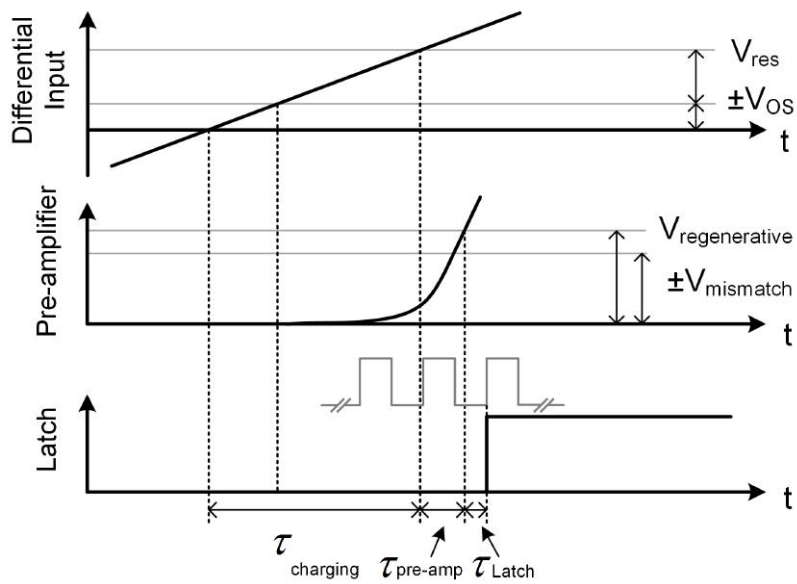


Figure 4.27: Analysis of the delay of the comparator for determinating the zero-crossing point of the blocking capacitor.

The total time error for this comparator can be expressed as:

$$\tau_{total} = \tau_{charging} + \tau_{pre-amp} + \tau_{latch} \tag{4.10}$$

where $\tau_{charging}$ is the time it takes for the differential input voltage to meet the resolving voltage of the comparator, which depends on the stimulating current I_{stim} and the size of the blocking capacitor C_B.

$$\tau_{charging} = \frac{(\pm\Delta V_{OS} + V_{res})C_B}{I_{stim}} \tag{4.11}$$

With a typical size of the blocking capacitor and the stimulation current, the $\tau_{charging}$ will be in tens of nanoseconds to several microseconds, which might dominate the total time error. This, unfortunately, causes a systematic delay, which is highly undesirable.

In order to address this problem, a feed-forward error compensation mechanism is introduced. The circuit schematic for the comparator is shown in Fig. 4.28. The comparator consists of a 3-stage preamplifier with output auto-zeroing, and a dynamic latch with a 4-bit current DAC for calibration. Neural stimulation usually consists of trains of stimulation pulses with the same amplitude and pulse width but varies in frequency (time interval between pulses). According to Eq. 4.11, the $\tau_{charging}$ will be the same for a train of pulses. So the delay of the comparator can be learned during the first few pulses and used to compensate the foregoing stimulation. Two four-bit DACs are used to calibrate the dynamic latch. The error of the comparison is learned from the residue charges after a stimulation pulse. In the first few stimulation pulses, an additional discharge phase is triggered after reversal phase to clear the charges on the capacitor. A switched capacitor amplifier is used to amplify the residue charges, and a dual threshold comparator is used to decide whether the residue charge is within the safe range or not. The schematic of the discharge and amplification circuit is shown in Fig. 4.29. The calibration DAC's value is changed according to the

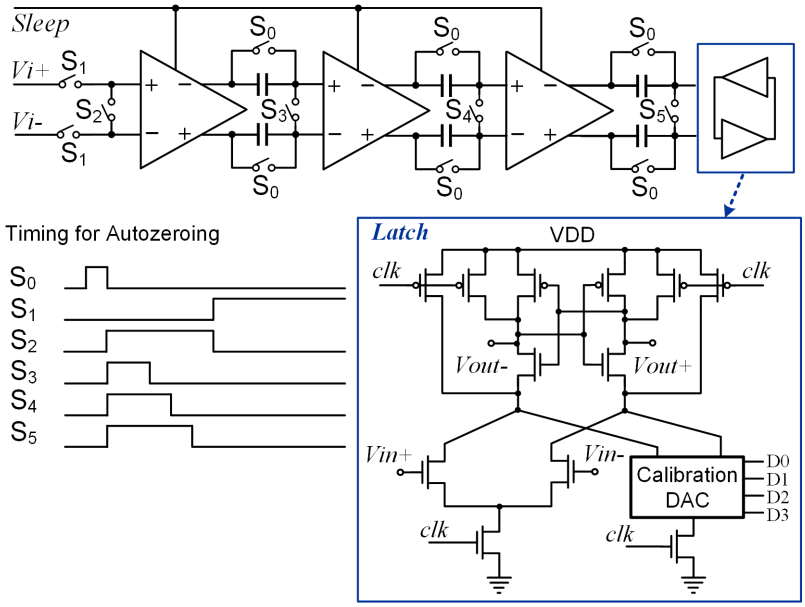


Figure 4.28: The circuit schematic of the comparator consisting of a 3-stage pre-amplifier and latch. The pre-amplifier has auto-zero calibration, and the latch has a 4-bit DAC for calibration.

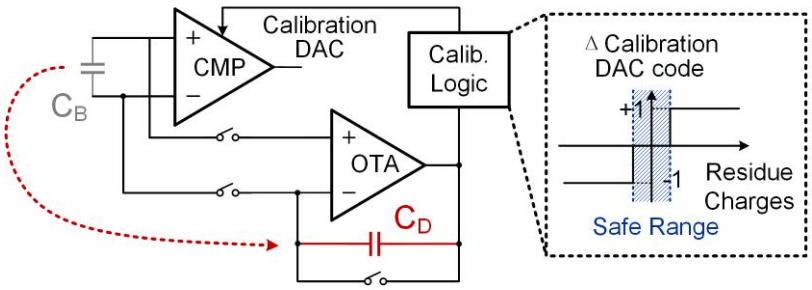


Figure 4.29: The circuit schematic of the switched capacitor circuit used to discharge the blocking capacitor. The circuit is also used to amplify the residue voltage for the calibration purpose. The amplified residue voltage is compared with two pre-defined safe voltage window. If the residue charge is out of the safe window, the calibration DAC of the comparator will be changed accordingly.

comparison result. The calibration DAC is designed to change 1 LSB each time for stability and simplicity. So in the worst case, it takes 16 cycles to change from 0 to the full range of the DAC, which will be finished in one or two pulse groups.

4.4.4 Experimental Results

The design has been fabricated in IBM 180nm CMOS technology. The occupied silicon area of the full chip is 3×1.5mm^2, including IO pads. One driving site features a dimension of 700μm × 150μm. The micrograph of the chip and the layout of one channel are shown in Fig. 4.30.

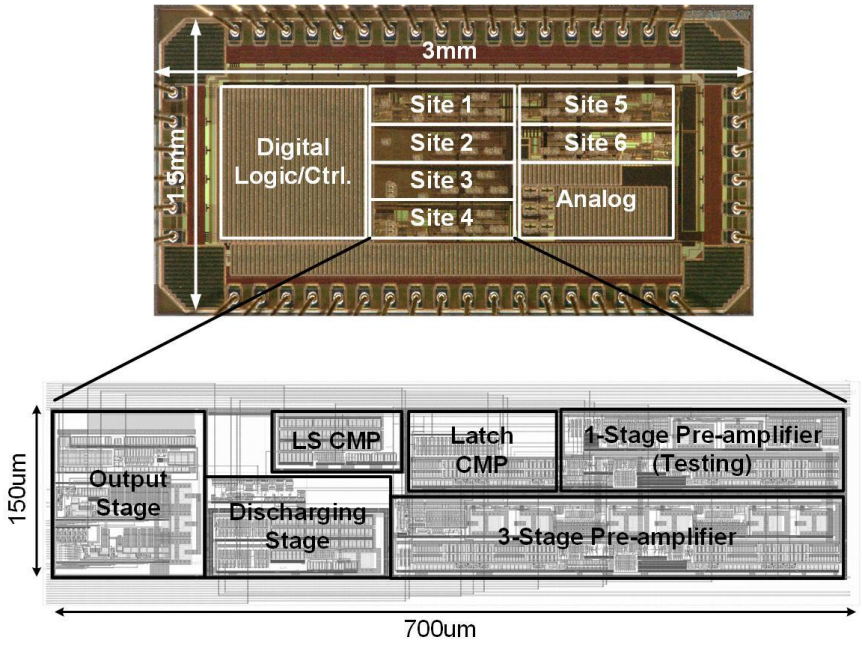


Figure 4.30: The micrograph of the fabricated stimulator chip. The occupied silicon area is 3×1.5mm^2.

Bench testing was conducted to verify the functions of the chip and the system. The measured currents from the output stage versus the output voltage are shown in Fig. 4.31. The measurement result shows a full compliance voltage range of 3.2V out of the 3.3V supply voltage at the current amplitude of 100μA, which corresponds to 97% of the full voltage range. This is much higher than the result achieved in the general-purpose design presented in section 4.3.3.

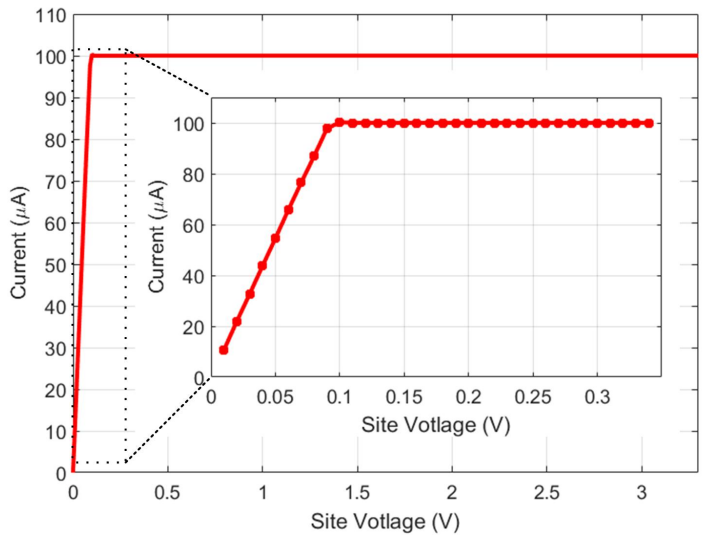


Figure 4.31: The measured current from the output stage versus the output voltage. The embedded figure shows a zoom-in plot from 0 to 0.3V.

Both traditional digital-set method and the proposed current-set method have been implemented in the chip for a comparison. The measured INL/DNL of the DAC using traditional digital-set method are 0.37/0.34 LSB, and are improved to 0.19/0.17 LSB using the proposed current-set method with the dynamic threshold variation cancellation technique. With the new technique, the charge error during a typical 100μA and 200μs is less than 0.05%.

The measured generated stimulation waveform and states of the finite state machine are shown in Fig. 4.32. The stimulation was measured with a high resistor load. The driving site was disconnected from the electrode when not activated to prevent leakage.

Another bench test is used to verify the function of the adaptive driving. Fig. 4.33 (a) shows a measurement of the electrode voltages during the driving voltage adjustment. Stimulation currents were measured under a load of two 10nF capacitors

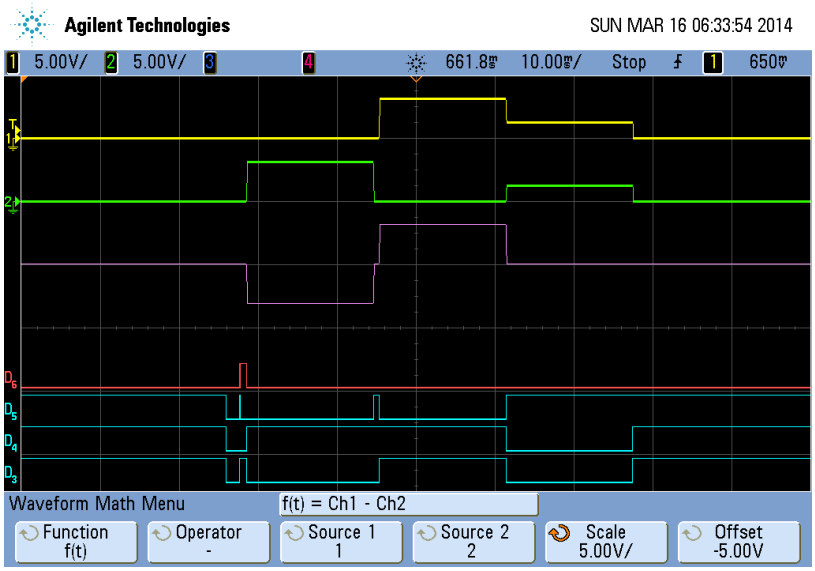


Figure 4.32: The measured generated stimulation waveforms with a high impedance load. Channel 1 and channel 2 measures the output of the WE and CE electrode, respectively. The Math channel calculates the difference between the two channels. Channel D3 to D6 show the states of the digital module.

and a 10kΩ resistor in series. A blocking capacitor of 100nF was applied. The probe placement is highlighted in the boxed figure. The boosting converter was implemented off-chip. Fig. 4.33 (b) shows the measured current across the load. The current maintains constant during the stimulation phase with the driving voltage adjustment.

In-vitro tests are conducted using a 75μm tungsten electrode in 0.9g/100mil Sodium Chloride. Fig. 4.34 shows a comparison of the measured voltages over a 5-min continuous stimulation using a traditional method and the proposed method. It was measured at the same driving site under a different configuration. Given the same mismatch in the current sources, a drifting of the electrode voltage when using the traditional digitally set DAC without discharge is shown in Fig. 4.34 (a), while the proposed method successfully resolves this problem. The measurement of the

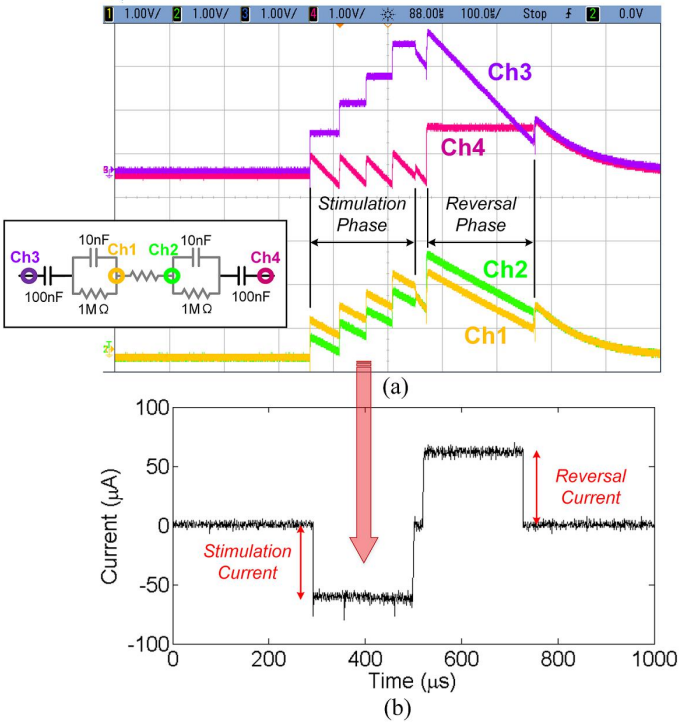


Figure 4.33: (a) The measurement of a stimulation pulse during the driving voltage adjustment. The load model is given with the measurement points highlighted. (b) The measured stimulating and reversal currents during the driving voltage adjustment.

blocking capacitor's voltages in 20 trails are overlaid in Fig. 4.34 (b), and the derived currents are plotted in Fig. 4.34 (c). The charges over-reversal in the traditional method are shown from the test in the saline solution. The measurement results verifies the theoretical analysis and the simulation.

To demonstrate that the stimulator is capable of evoking physiological activity, an *in-vivo* experiment was performed in a sedated rat. Trains of biphasic stimulus pulses (10 pulses, 5 ms interpulse interval, 0.3 ms/phase) were delivered through a pair of insulated tungsten microwires, with a 50μm diameter, implanted near the intrinsic muscles that protract the mystacial vibrissae. Whisker movements, as measured by

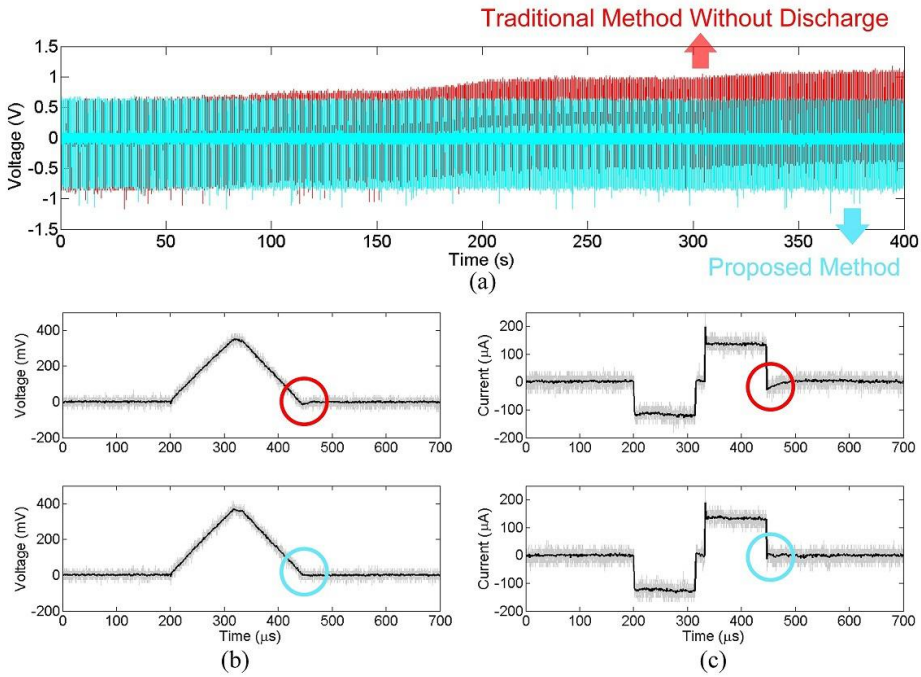


Figure 4.34: The measurement of biphasic stimulation in the saline solution using the traditional method (red) and the proposed method (blue). (a) shows a 5-min continuous stimulation without discharge. (b) and (c) show an overlay of 20 measurements of the voltage across the blocking capacitor and the derived stimulation current.

an optical micrometer, were reliably evoked as shown in Fig. 4.35 (a). The stimulator IC was programmed by a microcontroller with a wireless transceiver. Whisker displacements were a function of current intensity as shown in Fig. 4.35 (b). This experiment can be further used to implement facial reanimation for patients who suffer from facial paralysis.

The measured performance of the chip is summarized in Table 4.1.

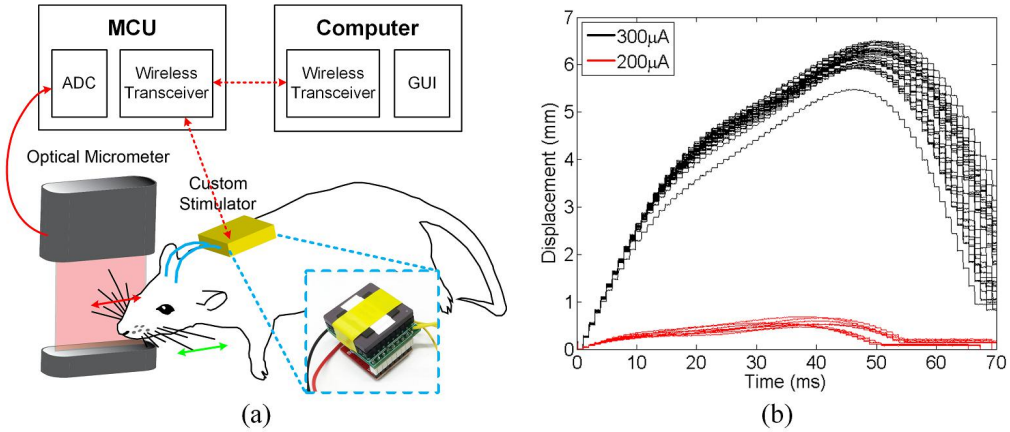


Figure 4.35: *In-vivo* experiment performed on a sedated rat. (a) Whisker movements, as measured by an optical micrometer, were reliably evoked. (b) Whisker displacements were a function of current intensity.

Table 4.1: Chip Specification Summary

Driving Site	# of sites	6
	Area per site	0.1mm^2
	Driving voltage	3.3 V
	Compliance range	97% (typical)
	Stim current	<2mA
	Stim freq.	1-500 Hz
	Charge error	<0.05% (typical)
DAC	Resolution	6-bit
	INL	0.19 LSB
	DNL	0.17 LSB
	Full scale std	<0.7%
Comparator	Resolution	40uV
	Calibration	auto-zero/4-bit DAC
	Speed	40MHz
Power	per site	136μW
	Coin battery	1.2V regulator off-chip
	Efficiency	81%

4.4.5 Conclusion and Future Work

In this section, a high efficiency, tissue-friendly net-zero charge stimulator is proposed. The net-zero charge stimulation is achieved by controlling the timing of the reversal phase based on monitoring the residual charge. Arbitrary channel configuration is achieved without a pre- or on-the-fly calibration, which enables a more dedicated stimulation position and pattern. Feedback control of the adaptive driving voltage and stimulation charge recycling are further proposed to improve stimulation efficiency. A novel current-mode DAC is implemented to suppress the process variation across the driving site array. A digital feed-forward error compensation is used to calibrate the zero-crossing detection comparator in a continuous stimulation pulse train. Both *in-vitro* and *in-vivo* experiment results are presented.

In the future, it is worth consider to design a net-zero charge neural stimulator without the blocking capacitor. The silicon area of each driving site can be further reduced. And by adding multiplexer switches, more stimulation channels can be supported with near simultaneous without increasing the number of driving sites. In addition, the power management units can be integrated on-chip for better performance and higher integration.

4.5 Conclusion

The design of electrical neural stimulator with high performance, good safety, and power efficiency has been a big challenge. A lot of research has been done in this field for several decades. With the help of advanced CMOS technology, advanced calibration techniques can be applied for achieving high current and charge accuracy.

However, all efforts will be in vain if the real physiological reactions at the electrode-electrolyte are not taken into account. Using feedback technique, a net-zero charge stimulation can be achieved without the need for a direct measurement of the underlying reactions. In addition, a high power efficiency topology can be achieved using adaptive driving. The developed chip and system have been used for neuroscience experiments and explorations.

Chapter 5

Bi-directional Neural Interface and Closed-loop Control

5.1 Introduction

A bi-directional neural interface enables simultaneous recording and stimulation with the neural system, establishing a two-way direct communication link between the brain and the external world [261]. The importance of a bi-directional closed-loop neural interface can be understood from several aspects: i) In the development of prosthetic devices, an electrical neural stimulation can provide an artificial sensory feedback to the user, allowing the user to perceive the movement and the haptic interaction with external objects [262]. This is important for the user to fully control a prosthetic; ii) In the treatment of the Parkinson's disease, the mechanism underlying the deep brain stimulation remains not clear [40]. Research shows that the application of the closed-loop stimulation has a greater effect than the conventional open-loop

stimulation paradigms [11, 41], and have the potential to be effective in other brain disorders [41]. iii) In the study of electrophysiology, the brain's response to an external stimulus is a complex combination of the activities triggered by the sensory stimulus itself and the brain's internal state, which needs more than statistical descriptions of the responses [44]. So only a closed-loop approach is effective in these studies [263].

Although the importance of the bi-directional closed-loop neural interface has been recognized [264], it has not been widely used in neuroscience research and BMI devices. The electronics design is one of the bottlenecks. There are two primary design challenges in a bi-directional neural interface system: i) The effects caused by recording and stimulating simultaneously, namely the stimulation artifacts [239, 265–267]; ii) The design of an on-chip real-time closed-loop controller [46, 80, 268–271]. Addressing both challenges is critical in a successful implementation of the bi-directional closed-loop neural interface. Thus the goal of this chapter, is to review and analyze the practical design issues related to a bi-directional neural interface and a closed-loop controller. Several novel circuit and system level designs are proposed to improve the state-of-the-art.

The chapter is organized as follows. Section 5.2 analyzes the origins of the stimulation artifacts in a bi-directional neural interface, reviews the prior works, and presents a study on the stimulation artifacts with different electrode configuration and circuitry topologies. Both *in-vitro* and *in-vivo* experimental results are presented. Section 5.3 reviews and summarizes the mechanisms of different closed-loop neural interface systems, and presents the design and testing of a commonly used PID controller for a generalized bi-directional neural interface system-on-chip (SoC).

5.2 Stimulation Artifacts in the Bi-directional Neural Interface

5.2.1 Introduction

The stimulation artifact is a known issue in simultaneous neural stimulation and recording. A long lasting stimulation artifact blanks the recording front-end, and corrupts the evoked neural response. Thus suppressing the stimulation artifact is critical in a bi-directional neural interface design. Several techniques have been proposed in the literature to attenuate or remove the stimulation artifacts, including recording front-end blanking, symmetrical electrode placement [265], temporary frequency shifting [239, 266, 267], real-time signal processing in the computer [272] or on-chip [75]. However, most proposed techniques have certain constraints, and the conclusions are not suitable for a general-purpose bi-directional neural interface design.

The goal of this study is to find the stimulation artifacts in different configurations. The combination of different stimulator and recorder configurations, namely monopolar and bipolar stimulation, single-ended and differential recording, with common and separate grounds were studied. In addition, different power supply configurations (dual-supply and single-supply), stimulator architectures (type-I and type-II) were taken into account. To the best of my knowledge, this work presents the first analysis of the stimulation artifacts considering both neural interface configuration and electronics architectures. The main sources of the stimulation artifacts were analyzed, and a custom testing board was designed to verify the analysis. Both *in-vitro*

and *in-vivo* experimental results are presented. The analysis conclusions and the design recommendations are given at the end of this section.

5.2.2 Review of Prior Work

Several studies of the stimulation artifacts in bi-directional interfaces have been reported in the literature. DeMichele *et al.* from the Sigenics Inc. and the Illinois Institute of Technology proposed a stimulus-resistant neural amplifier in 2003 [266]. The amplifier has an artifact suppression mode, which shifts the input frequency corner to 10Hz by a DC servo loop in the second stage. The work uses a low gain (x4) first-stage amplifier, which will not be easily saturated. However, most custom neural front-end designs prefer to use high gain in the first-stage amplifier to achieve a high overall noise efficiency. In these cases, the artifact suppression cannot resolve the long-lasting saturation from the first stage.

R. A. Blum and E. A. Brown *et al.* from the Georgia Institute of Technology and the University of Illinois at Urbana-Champaign proposed a stimulation artifacts model and a circuit module for the artifact removal in 2004 and 2008 [239, 267]. However, the model only considers a voltage-mode stimulation without a charge balancing design, and assumes the recording and stimulation circuits use the same electrode. A *pole-shifting* technique was used in the first stage, while the highpass pole is set to be 200Hz. A *soft-switching* technique was used to make a smooth transition between the different switch phases, which requires additional custom hardware design and optimization.

Rossi *et al.* proposed an artifact suppression device for recording the local field potential during a deep brain stimulation in 2007 [265]. The work uses separate

grounds for the recorder and the stimulator. The recording ground was placed on the scalp, and the stimulator's ground was placed on the shoulder. Besides, the stimulation electrodes were placed in the middle of the differential recording electrodes. Because the recording frequency was 2-40Hz and the stimulation frequency was 130Hz, a 10-pole lowpass filter was used to remove the stimulation artifacts. However, the recording and stimulation frequency ranges have overlaps in many other cases, and the stimulation electrode location cannot be chosen in the favor of the differential recording.

A. E. Mendrela *et al.* from the University of Michigan, Ann Arbor and the University of Minnesota, Minneapolis proposed a bi-directional neural interface circuit with an active stimulation artifact cancellation in 2016 [75]. An on-chip digital adaptive filter was used to remove the stimulus artifacts. However, the proposed design is based on the assumption that the stimulation artifact won't saturate the recording electrode or push the recording front-end out of the linear range. The digital filter itself doesn't help recover the recording amplifier from the saturation, or reduce the stimulation artifacts.

In summary, the stimulation artifacts depend on the types of the recording and stimulation electrodes, the electrical circuitry, the configuration of the ground, the characteristics of the input stage of the recording front-end, the methods and parameters used for the stimulation, and so on. Most existing investigations on the stimulation artifacts are restricted for a certain application or a certain configuration. However, it is important to understand the difference between the configurations and their effects on the stimulation artifacts, which is the goal of the following study.

5.2.3 Analysis of Stimulation Artifacts

5.2.3.1 Origins of Stimulation Artifacts

The origins of the stimulation artifacts can be quite complicated, as reviewed in the previous section. The analysis and models in this section are not intended to give an accurate electrochemical description or a precise estimation of the stimulation artifacts, but to provide insights for developing techniques for reducing or canceling the stimulation artifacts.

A. Electrode Saturation

During the stimulation phase, a portion of the charges might be stored on the double layer capacitors on the recording electrodes [20]. Ideally, after the reversal phase, a charge balance will be achieved, and all the tissue environment will return to the potential before the stimulation. However, this is not always achievable. If the recording electrode's potential is still within the input range of the amplifier after the reversal phase, the recorder will return to the normal operation reasonably fast, depending on the bandwidth of the amplifier. But if the electrode potential is pushed away from the input range, it takes a very long time for it to recover, since the input stage of the circuits usually has a very large time constant needed for high input impedance.

Due to the small amplitude of the neural signal, the neural recording amplifier is usually designed with a linear input range of tens of millivolts, and the common mode input range is usually limited to hundreds of millivolts, depending on the supply voltage and the circuit architecture. The capacitively coupled instrumentation amplifier

has extended input range in theory, however, is also limited by the ESD circuits and breakdown voltage of the input transistors. In addition, the chopping amplifiers, or the amplifiers employ a DC servo loop suffer more from the voltage headroom [92]. Even these limitations are not a problem for the neural amplifier alone, they may become a significant problem in a bi-directional neural interface.

B. Voltage Gradients

The voltage gradients can be easily understood by considering the tissue is conductive, and all electrodes inserted in the tissue environment are interconnected. Fig. 5.1 shows the circuit model proposed to analyze the stimulus artifacts in the bi-directional neural interface. A monopolar stimulation with a single-ended recording is used for illustration.

Firstly, consider the case where the recorder and stimulator share a common ground, as illustrated in Fig. 5.1 (a). The electrode impedance for the stimulator, the recorder and the common ground is Z_1, Z_2 and Z_3, respectively. The spreading resistances between the three electrodes are represented by R_{12}, R_{13} and R_{23}. Assume that the stimulator and the ground electrodes have a low impedance, while the recording electrode has a high impedance, and the instrumentation amplifier has a high input impedance, the artifacts due to the stimulus can be expressed as:

$$\begin{aligned} V_{artifact} &\approx \frac{R_3}{R_1 + R_3} V_{stim} \\ &= \frac{R_{23}}{R_{12} + R_{23}} V_{stim} \end{aligned} \tag{5.1}$$

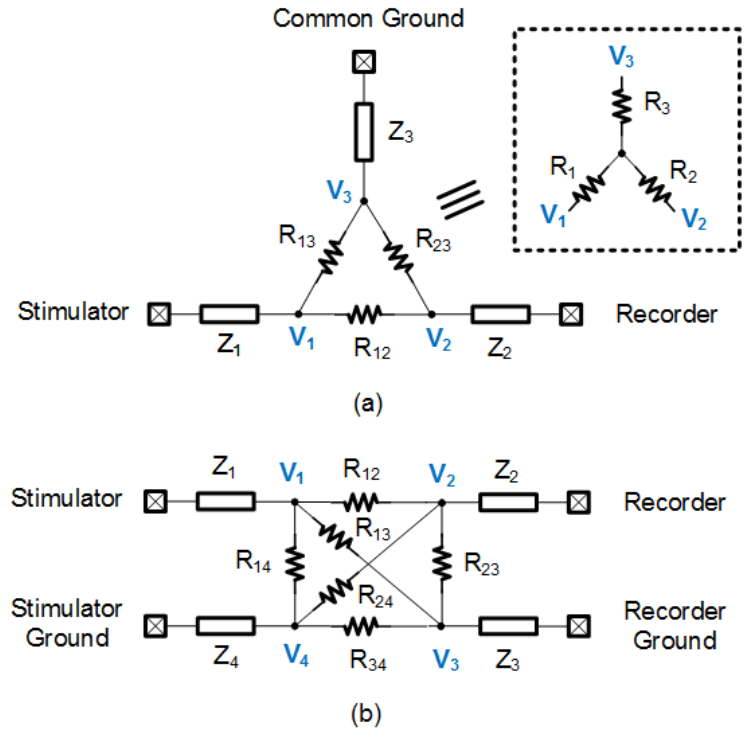


Figure 5.1: The circuit model for the bi-directional neural interface with (a) shared grounds and (b) separated grounds.

If the R_{12} and R_{23} are in a comparable magnitude, the artifact can be half of the stimulation's compliance voltage, which will easily saturate the recording electrodes or push the instrumentation amplifier out of the linear input range.

Secondly, Fig. 5.1 (b) shows the case where the recorder and stimulator have separated grounds. The artifacts due to the stimulus can be expressed as:

$$V_{artifact} \approx \frac{(R_{13}R_{24} - R_{12}R_{34})R_{23}}{R_{eq}} V_{stim} \tag{5.2}$$

where

$$R_{eq} = R_{12}R_{23}(R_{13} + R_{34}) + R_{13}R_{24}(R_{12} + R_{23}) + R_{24}R_{34}(R_{12} + R_{13}) \tag{5.3}$$

If the recording and stimulating electrodes are well separated, $R_{12} \approx R_{13}$, and $R_{23} \approx R_{24}$, then the $V_{artifact}$ is approximately zero. Even if the electrodes are not separated far away, or the recording electrodes for signal and ground have significant different impedance, the $(R_{13}R_{24} - R_{12}R_{34})R_{23}$ will still be much smaller than R_{eq}, and $V_{artifact}$ will be a very small portion of V_{stim}.

From the above analysis, the stimulation artifact due to the voltage gradient can be in the same order of the stimulating compliance voltage when the recording and stimulation share a common ground, and the artifacts can be minimized when the grounds are separated in the circuits.

C. Capacitive Coupling

The capacitive coupling between the stimulating and recording leads also contributes to the stimulation artifacts. Even though the capacitive coupling between the wires is usually less than 1pF, the simulation voltage can be six orders higher than the amplitude of the neural signal, so that the coupled signal may still be visible to the recorded signal. The coupling is worse if there is no shielding on the recording and stimulation electrodes, or the routing wires are very long.

5.2.3.2 Configuration of the Interface Circuits

The configuration of the stimulator and the recorder, and the placement of the electrodes have a high impact on the stimulation artifacts. In this section, different configurations of the neural recorder and the stimulator, and the related circuit implementations are discussed.

There are two typical configurations for a multi-channel neural signal recording front-end: the *single-ended* recording and the *differential* recording, as illustrated in Fig. 5.2 (a) and (b), respectively. In the single-ended recording, a reference electrode

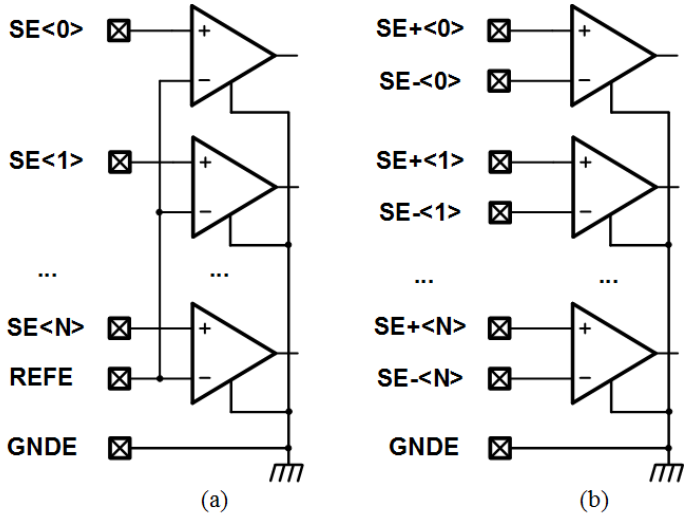


Figure 5.2: Configuration of typical multiple channels neural recording front-end. (a) single-ended recording configuration, and (b) differential recording configuration.

and a ground electrode are shared among all channels. The ground can be used as the reference in some cases. In the differential recording, two electrodes collect the signal between of them, one electrode provides the ground, and no reference electrode is required.

It should be noticed that *fully-differential* is a widely used term in the circuit design community. It refers to the circuit module which has both a differential input and a differential output. However, a differential recording doesn't require a fully differential circuit.

Similar to the recording, there are two typical configurations for the electrical neural stimulation: the *monopolar* stimulation and the *bipolar* stimulation. In a monopolar stimulation, the electrical charges are injected from one electrode for stimulation, and pulled from the same electrode to achieve the charge balancing. A low impedance counter electrode provides the return path. Usually, the ground is used as the counter electrode. In a bipolar stimulation, the electrical charges are passed between two electrodes for stimulation. A low impedance ground electrode is usually connected to the tissue and provides the electronic ground, however, it is not required since only a potential difference is needed between the bipolar electrodes to generate the current. It should be noticed that there is a difference between a *biphasic* stimulation and a bipolar stimulation. A biphasic stimulation means that a reversal phase is followed by the simulation phase. Both monopolar and bipolar stimulation can perform a biphasic stimulation.

There are two typical methods to implement the stimulator. If both a current sink and a current source are used in the working electrode to generate the stimulation and reverse phase, the stimulator circuit is referred to as a *Type-I* stimulator in this work. If only a current sink or a current source is used in the working electrode, and a voltage buffer is connected to the counter electrode, the stimulator circuit is referred to as a *Type-II* stimulator in this work. It should be noticed that:

- A Type-I stimulator can perform both a monopolar stimulation (Fig. 5.3 (a)),

and a bipolar stimulation (Fig. 5.3 (b)). Using the same supply voltage, the bipolar configuration gives twice the compliance voltage range for the stimulation.

- A Type-II stimulator usually can only be used to perform a bipolar stimulation (Fig. 5.3 (c) and (d));
- Since only a current sink or a current source is implemented in a Type-II stimulator, it saves the voltage headroom for one current source (or sink), which is an advantage over the Type-I stimulator in a low-voltage and low-power design;
- A current sink is usually easier to implement than a current source since its most circuit components are operating at a low voltage with respect to the ground. So a Type-II stimulator with a current sink (Fig. 5.3 (c)) is in fact much more popular than the design with only a current source (Fig. 5.3 (d)).

5.2.3.3 Practical Issues in the Electrical Circuits

One of the practical design challenges is the different requirements in the supply voltages for the stimulator and the recorder modules. The stimulator module usually requires high supply voltages for driving the high impedance electrodes. At the same time, the recorder and the digital modules need a low supply voltage to reduce the power consumption. Both dual-supply and single-supply are commonly used in circuit design, as shown in Fig. 5.4. A dual-supply design offers a common-mode ground, which has to be generated in a single-supply design, mostly likely by a push-pull stage. A level-shifter is required if a low voltage digital signal is to be used to control a high voltage stimulator in both configurations.

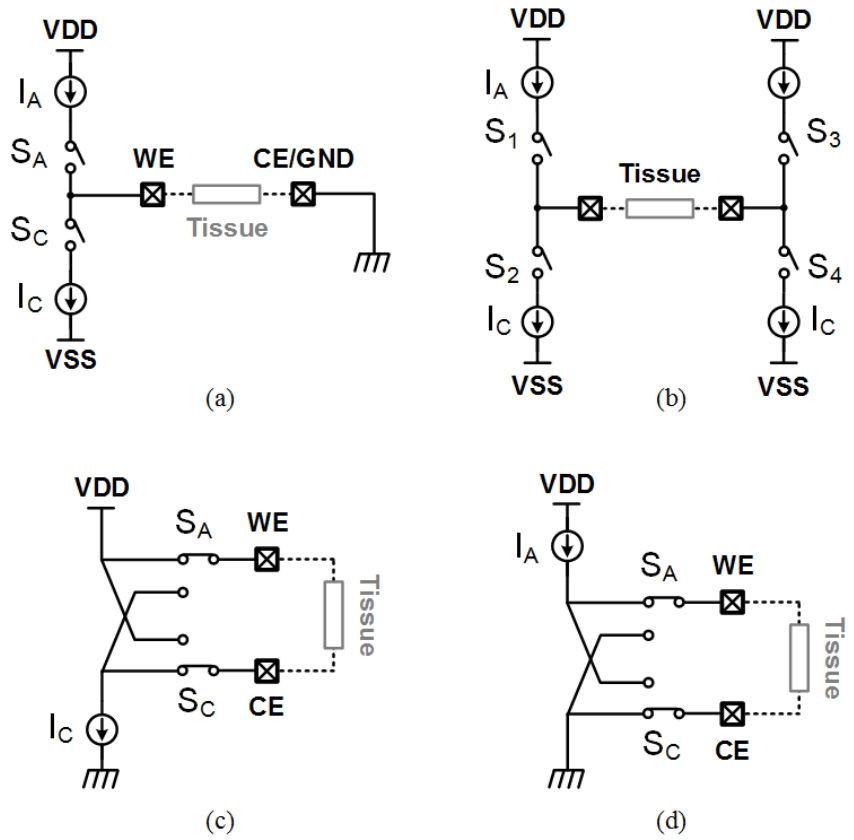


Figure 5.3: The configuration of typical neural stimulation back-ends. (a) shows a monopolar stimulation configuration, (b), (c), and (d) are bipolar configurations. (a) and (b) are using a Type-I stimulator, (c) and (d) are using a Type-II stimulator.

The two configurations are equivalent in the circuit design, however, when connecting the electronics ground to the tissue ground, there is a problem for the single-supply design. If the tissue ground is connected to the stimulator ground, the common mode signal will be much higher than the recorder's ground, and may be even higher than the recorder supply in many cases. For a DC-coupled design, or designs require a DC servo loop, what is usually not feasible. If the tissue ground is connected to the recorder ground, the tissue ground is not centered between the highest and lowest supplies of the stimulator, which will be a waste of the power and the compliance range. It should also be noticed that if the stimulator tries to discharge the local

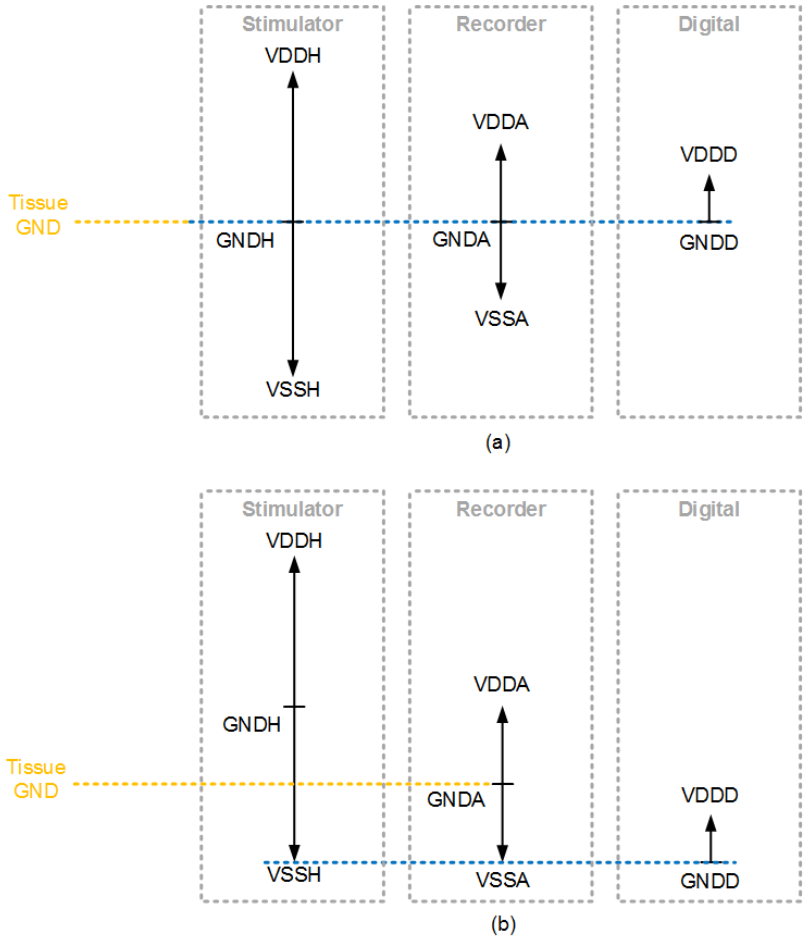


Figure 5.4: Two common ground configurations for the recorder and the stimulator: (a) dual-supply and (b) single-supply. In a dual-supply design, the grounds of all circuit modules are connected together, while in a single-supply design, the lowest supply in each module is connected together. The design is also limited by the bulk of the CMOS technology.

tissue to the stimulator ground GNDH, the potential difference between GNDH and GNDA will cause a DC current, which may damage the tissue. This must be avoided. Even though a single-supply design is simpler in certain cases, especially with bulk CMOS technology, a dual-supply design might be required for the bi-directional neural interface design.

5.2.4 Methods

A bench-test board was designed to fully study the stimulation artifacts in different configurations. The block diagram of the board is shown in Fig. 5.5, and the photograph of the assembled board is shown in Fig. 5.6.

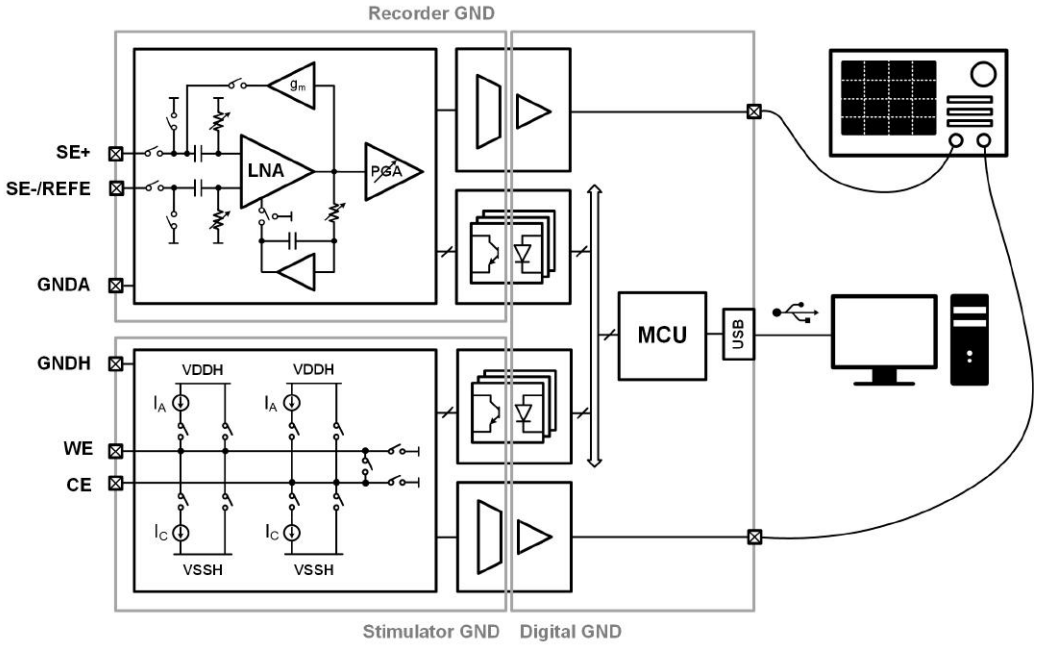


Figure 5.5: The block diagram of the bench testing board. The supplies and grounds for the recorder, the stimulator and the digital modules are intentionally separated on the board.

The board is carefully designed with ground isolation for each module. The recorder, stimulator and digital processor have individual ground and power management unit. The recorder and stimulator are powered by two 9V batteries. Positive and negative regulators are used to provide regulated supply voltages of ± 6V. The digital module is powered by a 5V USB cable, and regulated to 3.3V for the microcontroller. Optical isolators TLP292 from Toshiba [273] are used to provide a digital control signal to the recorder and simulator modules without connecting the ground.

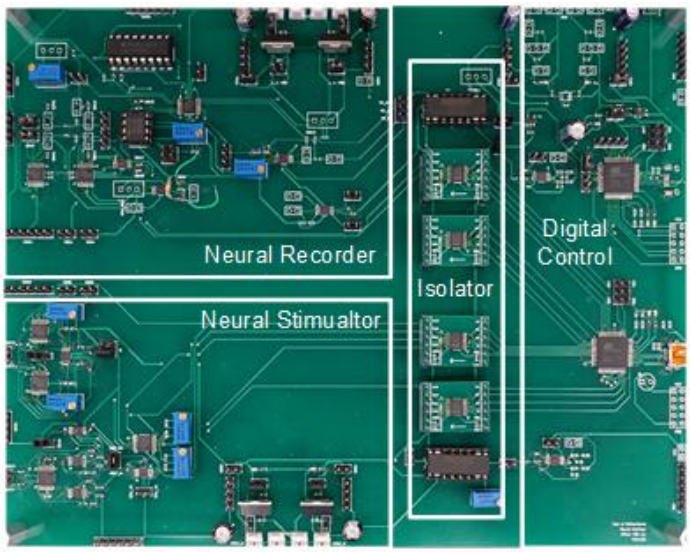


Figure 5.6: The photo of the assembled bench testing board. The dimension is 216mm × 171mm.

The isolation amplifier uses a duty cycle modulation-demodulation technique to converter the input signal to the output in separate grounds. A 50kHz two-pole lowpass filter is implemented on board to remove the ripples. The data acquisition devices and the oscilloscope share the ground with the digital module. By using the isolation amplifier, the oscilloscope can monitor the analog output of the recorder without connecting the mains ground to the animals.

The recorder has two stages. A first stage using low noise instrumentation amplifier INA111 from Burr-Brown [274]. The amplifier has a noise density of 13nV/rtHz at 100Hz. The gain is fixed at 11. The input stage is biased using large resistors to provide a very low cut-off frequency. Switches are integrated to be able to disconnect or blank the recording electrodes during the stimulation. A DC servo loop is also implemented which can move the highpass corner frequency. The second stage provides an additional gain of 50 to the recording chain. The recorder can perform either a

single-ended recording or a differential recording depending on the connection of the electrodes.

The stimulator has two output stages, each has a current source, a current sink, and two switches for shorting the electrodes to the two power lines. Additional switches are also included to short the two electrodes, or short the electrode to the stimulator ground. By controlling the timing of the switches, the stimulator can perform a monopolar or a bipolar stimulation in either Type-I or Type-II configuration. The amplitude of the current source and sink are programmable and tunable for matching. Another isolation amplifier can be connected to the electrode for monitoring the compliance voltage of the electrodes.

The digital module mainly consists of a microcontroller ATxmega128A4U from Atmel to generate the timing for the stimulator and the control signal for the recorder. The microcontroller has an integrated USB 2.0 module. A computer user interface is developed in Matlab, and commands are sent to the bench test board via USB. Besides the oscilloscope, a pre-developed data acquisition board is used to collect the data and send to the computer.

5.2.5 Experimental Results

5.2.5.1 *In-Vitro* Experiment

A series of *in-vitro* studies was conducted in 0.9g/100mil Sodium Chloride. The placement of the electrodes is illustrated in Fig. 5.7. A pair of electrodes were inserted in the saline for both the recorder and the stimulator. When performing a single-ended recording or a monopolar stimulation, only one of the electrodes was

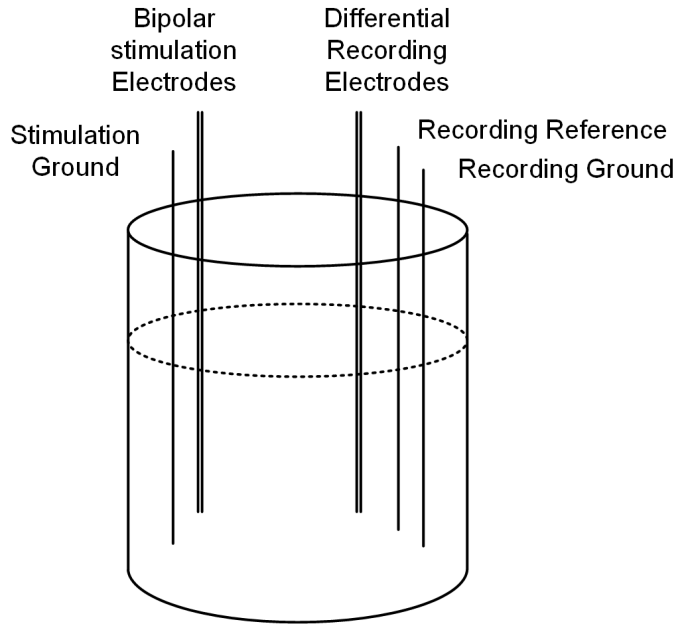


Figure 5.7: Electrode setup for the stimulation artifacts experiments.

used. A 75μm tungsten low impedance electrode was used as the reference electrode. A stripped copper wire was used as the ground.

Different configurations for recording and stimulation, different types of stimulators, with a common or separate grounds were explored in the *in-vitro* study. Table 5.1 lists the experimental setup and the corresponding results.

Table 5.1: *In-vitro* Experiments for Stimulation Artifact Study

				Recording	
				Single-ended	Differential
Stimulation	Type-I	Common GND	Monopolar	Fig. 5.8 (a)	Fig. 5.8 (b)
			Bipolar	Fig. 5.8 (c)	Fig. 5.8 (d)
		Separate GND	Monopolar	Fig. 5.9 (a)	Fig. 5.9 (b)
			Bipolar	Fig. 5.9 (c)	Fig. 5.9 (d)
	Type-II	Common GND	Bipolar (current sink)	Fig. 5.10 (a)	Fig. 5.10 (b)
			Bipolar (current source)	Fig. 5.10 (c)	Fig. 5.10 (d)
		Separate GND	Bipolar	Fig. 5.11 (a)	Fig. 5.11 (b)

Firstly, the stimulator was tested in the Type-I configuration. The stimulator was configured to do a 10-pulse biphasic stimulation. The pulse amplitude was 100μA, the pulse phase was 200μs. The pulse interval was 3ms. The discharge time was 1ms, and the discharge resistor was 1kΩ. The highpass frequency corner of the amplifier was set to be 0.159Hz by a 10nF coupling capacitor and a 100MΩ biasing resistor. The lowpass frequency was 7.2kHz, and the overall gain was 550. Fig. 5.8 shows the recorded stimulation artifacts with the recorder and the stimulator share the same ground. The figures in the top row show the results using a monopolar stimulation, and the figures in the bottom row show the results using a bipolar stimulation. The figures in the left column show the single-ended recording, and the figures in the right column show the differential recording. A monopolar stimulation with a single-ended recording gives the worst result, and a bipolar stimulation and a differential recording results in the minimum artifact.

Fig. 5.9 shows the recorded stimulation artifacts with the recorder and the stimulator having a separate ground. It should be noticed that the scale in this figure is much smaller than Fig. 5.8, one grid is only 50μV. The figures in the top row show the measurement results using a monopolar stimulation, and the figures in the bottom row show the measurement results using a bipolar stimulation. The figures in the left column show a single-ended recording, and the figures in the right column show a differential recording. Similar to the results with the common ground, a monopolar stimulation with a single-ended recording gives the worst results. However, with a separate ground, even the worst artifacts will not saturate the recording electrodes.

Secondly, the stimulator was tested in the Type-II configuration. The amplitude and timing of the stimulator pulses were the same as in the Type-I experiment. Fig. 5.10 shows the recorded stimulation artifacts with the recorder and the stimulator

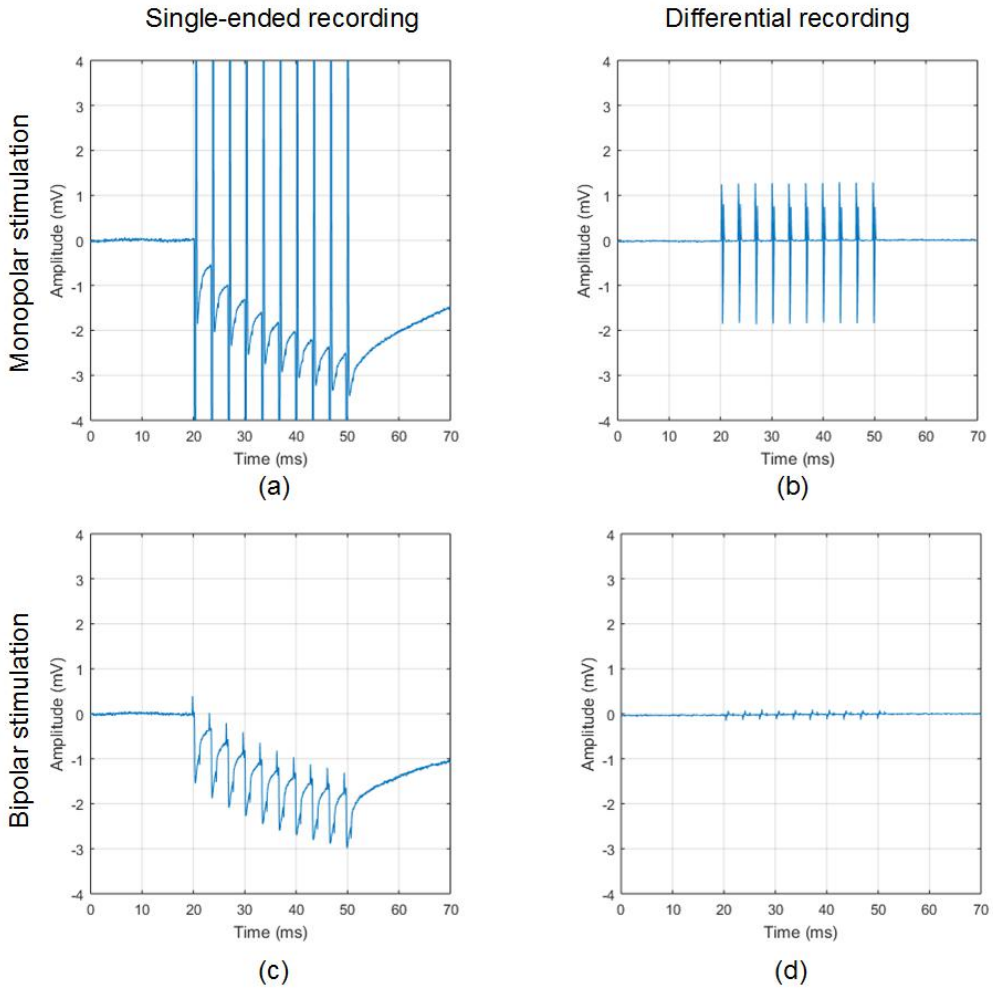


Figure 5.8: The measured stimulation artifacts of a Type-I stimulator with common ground between the recorder and the stimulator. (a) monopolar stimulation and single-ended recording, (b) monopolar stimulation and differential recording, (c) bipolar stimulation and single-ended recording, and (d) bipolar stimulation and differential recording.

share the ground. It should be noticed that the Type-II stimulator can only perform a bipolar stimulation. The figures in the top row show the measurement results using a current sink, and the figures in the bottom row show the measurement results using a current source. The figures in the left column show the single-ended recording, and

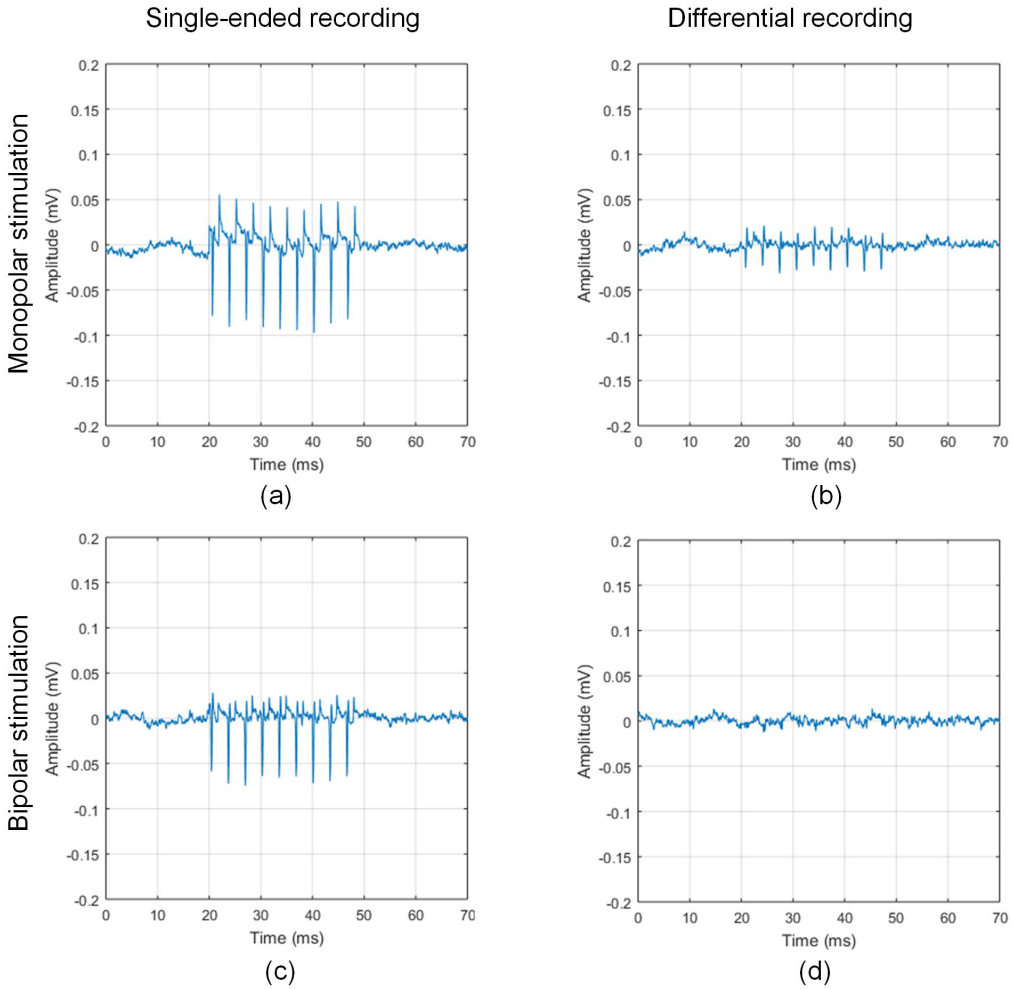


Figure 5.9: The measured stimulation artifacts of a Type-I stimulator with separate grounds between the recorder and the stimulator. (a) monopolar stimulation and single-ended recording, (b) monopolar stimulation and differential recording, (c) bipolar stimulation and single-ended recording, and (d) bipolar stimulation and differential recording.

the figures in the right column show the differential recording. In general, the Type-II stimulator using a current source gives much better results than the one using a current sink. The differential recording also gives better results in this case.

Fig. 5.11 shows the recorded stimulation artifacts with the recorder and the stimulator having a separate ground. A grid in this figure is 200μV. The current

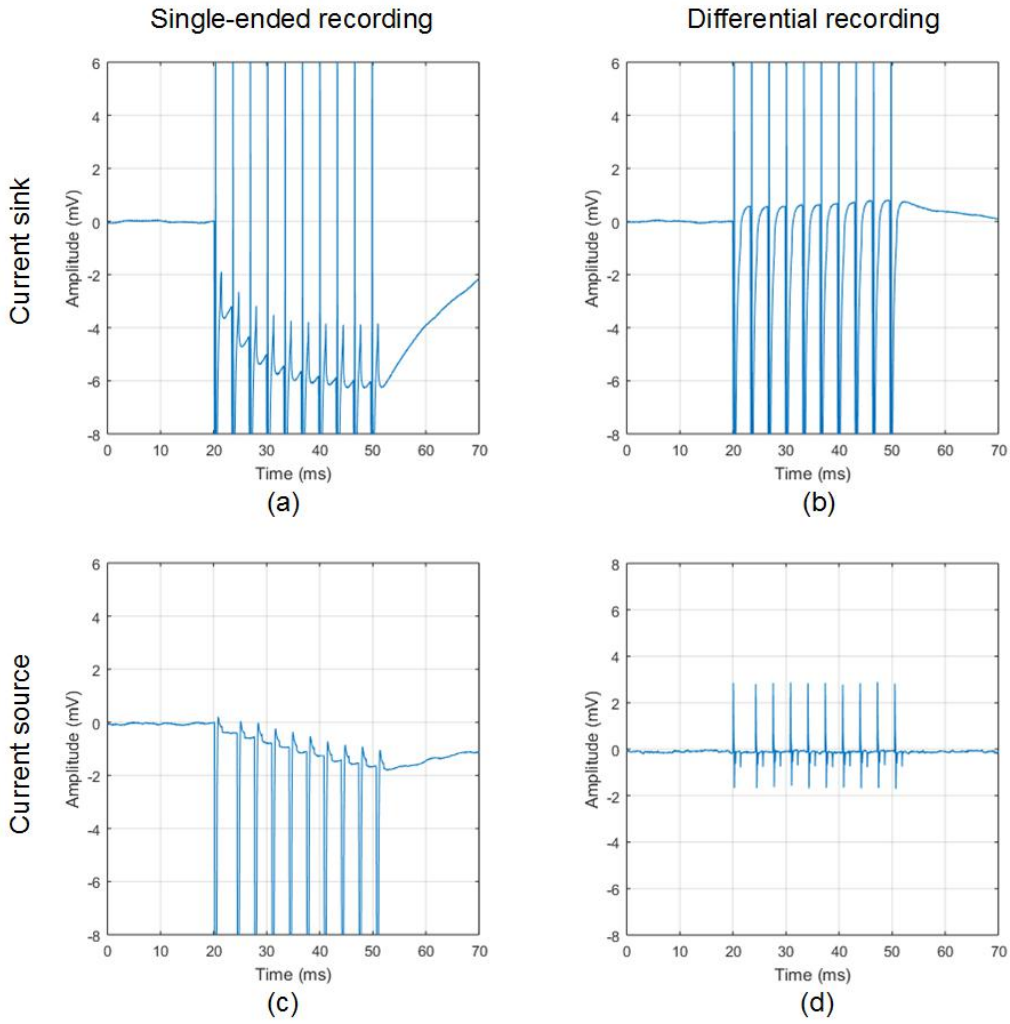


Figure 5.10: Stimulation artifacts of a Type-II stimulator with common ground between the recorder and the stimulator. The stimulator uses a current sink in (a) and (b), and a current source in (c) and (d). Recorder is configured to do single-ended recording in (a) and (c), and differential recording in (b) and (d).

source and the current sink stimulators had a similar performance when the ground of the recorder and stimulator were separate. In general, the stimulation artifacts were much smaller when the grounds were separate. It should be noticed that if the recording amplifier is not saturated, the signal processing techniques for removing the artifacts can be applied. In general, the stimulation artifacts are minimum when

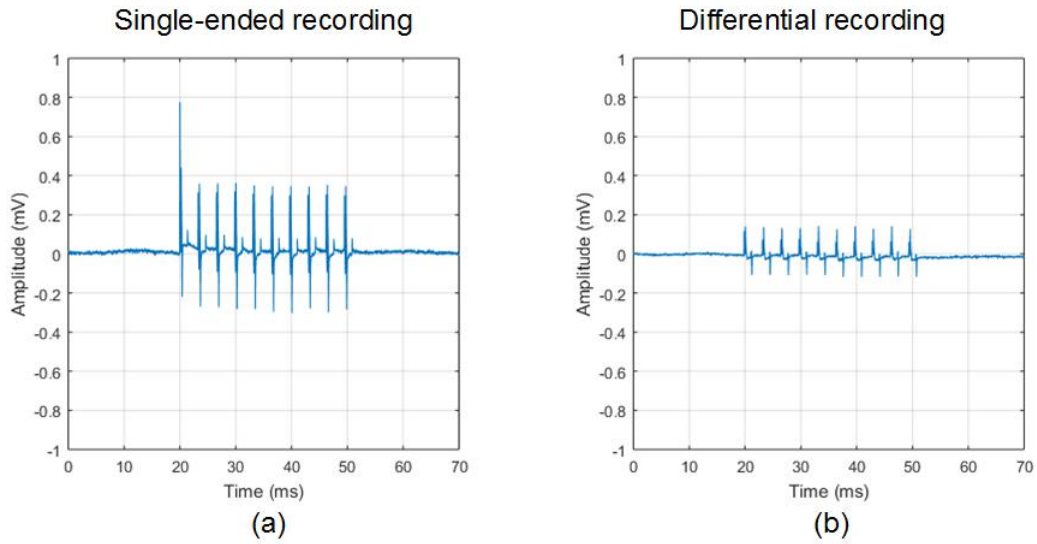


Figure 5.11: Stimulation artifacts of a Type-II stimulator with separate ground between the recorder and the stimulator. Recorder is configured to do single-ended recording in (a) and differential recording in (b).

the grounds of the recorder and the stimulator are separate. When the grounds are connected, a differential recording helps reduce the artifacts. The worst case would be the common ground with a single-ended recording.

5.2.5.2 *In-Vivo* Experiment

Two animal experiments were conducted to further verify the results from the *in-vitro* study. A female Long-Evans rat received two implants, one in the motor cortex, the other in the sensory cortex. A ground stew was connected to the skull of the rat to provide the ground for the recording. Fig. 5.12 shows the experimental setup.

A bipolar stimulation was performed in the sensory cortex while having a single-ended recording in the motor cortex. Fig. 5.13 (a) shows the recording when the grounds of the recorder and stimulator were connected together. A large stimulation

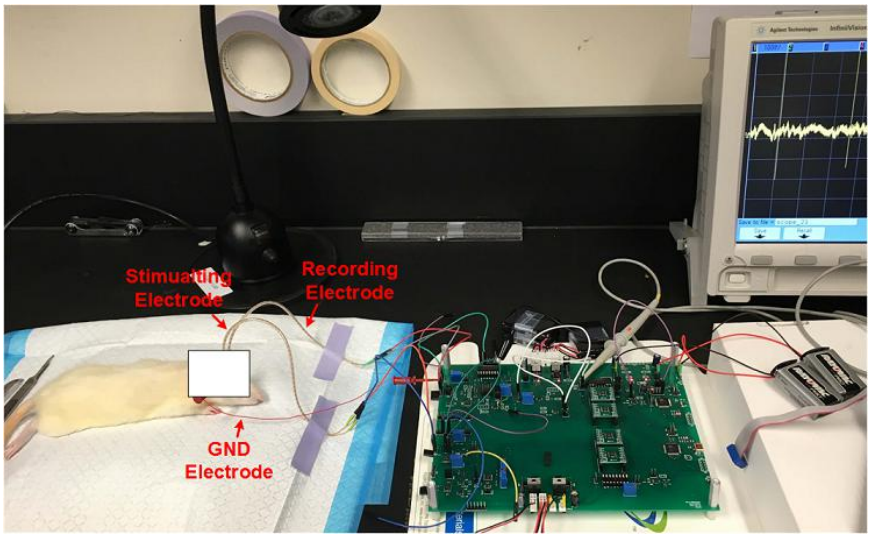


Figure 5.12: The photo of the setup of the *in-vivo* experiment. A female Long-Evans rat received two implants, one in the motor cortex, the other in the sensory cortex. The measurement was conducted using the testing board presented in section 5.2.4.

artifact appeared with a long recovery time. And Fig. 5.13 (b) shows when the grounds are separated, the stimulus artifact was minimized and an evoked potential was shown after the artifact. Notice also that the evoked potential is not visible in the top figure due to the large stimulation artifact.

Another bi-directional experiment was conducted in a macaque. The recording electrodes were chronically implanted in the left hippocampus while the stimulating electrodes in the upstream areas. The stimulation pulse train contains 10 pulses with 2mA current. Fig. 5.14 (a) shows the case when the grounds of the recorder and the stimulator were shorted together. The recording amplifier was saturated soon after the stimulation and the recovery took hundreds of milliseconds after the last pulse. Fig. 5.14 (b) shows the case when the grounds were separated, and the artifacts were minimized. Fig. 5.14 (c) shows the output of the recording with an additional low-pass filter with a frequency corner of 200Hz. The stimulus artifacts were completely

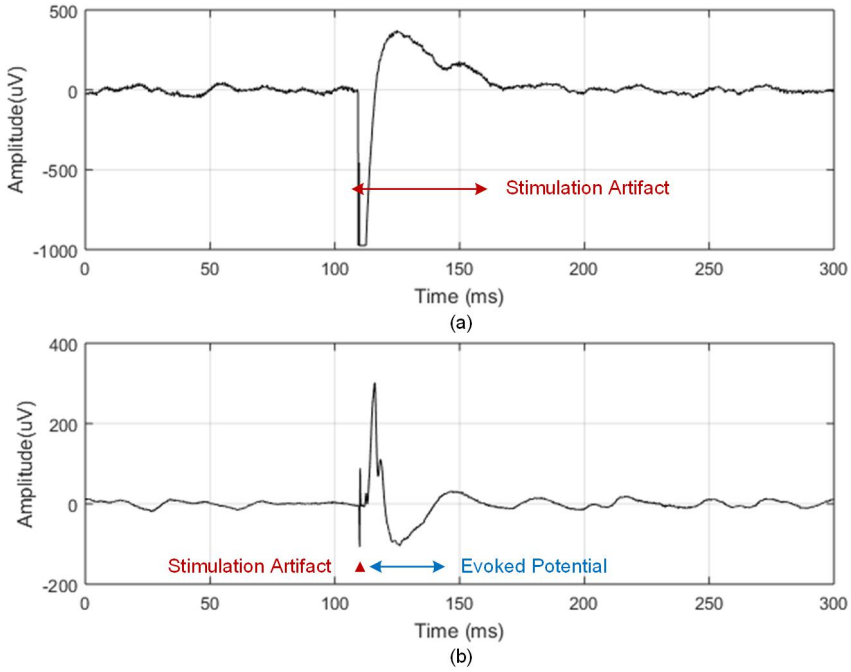


Figure 5.13: Bi-directional neural interface experiment in a Long-Evans rat. Local field potential was recorded in the motor cortex while stimulating the sensory cortex. (a) When the grounds of the recorder and stimulator were connected together, there was a large artifact and long recovery time; (b) when the grounds were separated, the artifact is minimized. An evoked potential was shown after the stimulus artifact.

removed by the filter since they were out of the signal band.

5.2.6 Conclusion

In this work, the stimulation artifact in a bi-directional neural interface has been studied. Different electrode and circuit configurations were taken into account in this study. Both *in-vitro* and *in-vivo* experiments were conducted. Several conclusions are given from the analysis of this work:

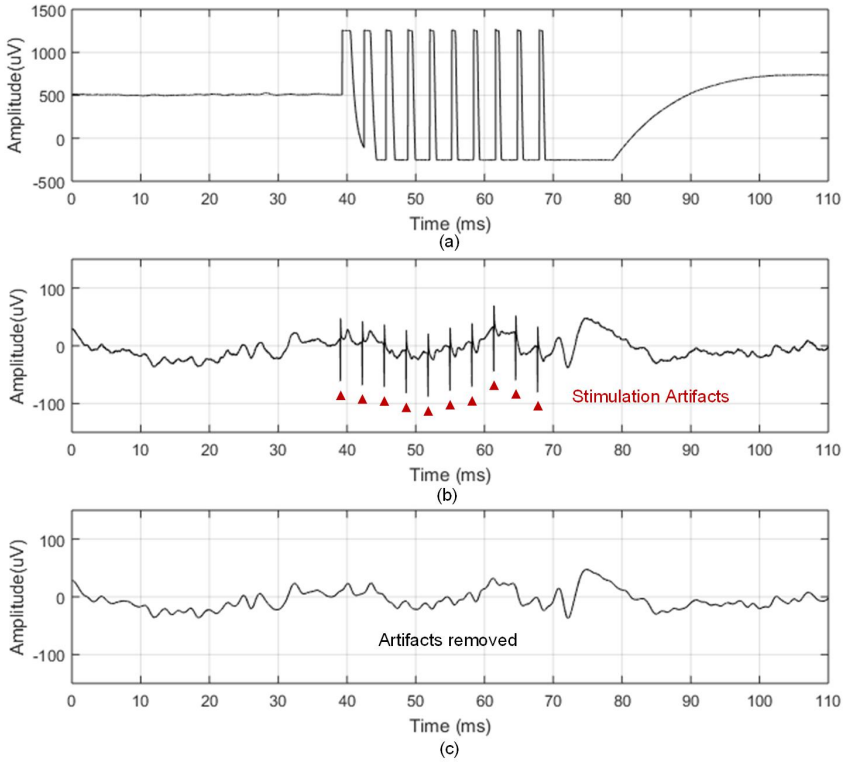


Figure 5.14: Bi-directional experiment in a Rhesus macaque. (a) The grounds of the stimulator and the recorder are shorted together, the stimulation artifact saturates the amplifier, and takes hundreds of milliseconds to return to the pre-stimulus potential. (b) The grounds of the stimulator and recorder are separated, and only minor artifact appears in the recording. (c) A 200Hz filter is applied to the recording in (b), which completely removes the artifacts.

- The stimulation artifacts can be minimized if the ground of the recorder and stimulator can be separated. Circuit techniques can be used to enable the isolation even if the bi-directional interface is implemented on a single die system-on-chip (SoC).

- A charge balanced stimulation causes much smaller artifacts in the recording amplifier. In a voltage mode stimulation without charge balancing, once the recording electrodes were saturated, they can only be recovered by the biasing circuit with a large time constant. One possible solution is to temporarily shift

the high-pass frequency corner to a higher frequency during or right after the stimulation. However, the signal will be corrupted and cannot be well recovered unless the exact timing, and frequency and phase change are known.

- A bipolar stimulation usually causes smaller artifacts in the recording than the monopolar stimulation, since the stimulation is restrained in the area between the two electrodes. However, a bipolar stimulation cannot replace a monopolar stimulation in triggering certain physiological response.

- A differential recording usually suppresses the stimulation artifacts. However, attention must be paid to make sure the differential electrodes are within the linear input range of the recording amplifier. If the signal is distorted, it might lead to wrong analysis results.

- The input range of the recording amplifier is usually limited by the supply voltages and the ESD circuits. Extending the input range of the recording amplifier is very helpful in preventing saturating the input stage. However, this is quite a challenge in a low voltage front-end design.

- When having a common ground, the Type-II stimulator with a current sink gives a large artifact, since one of the electrodes is fixed to a high voltage relative to the ground, which gives a large step input to the recording amplifier.

- Discharging the stimulation electrode might also give a step response to the recording amplifier.

In summary, separating recording and stimulation ground is highly recommended for a bi-directional neural interface design, especially if a monopolar stimulation and a single-ended recording is necessary. If a common ground has to be used, table 5.2

gives an estimation of the stimulation artifacts (both amplitude and duration). In this table, "+" means good and "-" means bad.

Table 5.2: Stimulation Artifacts with Common Ground

	Single-ended Recording	Differential Recording
Monopolar Stimulaiton	- - -	+
Bipolar Stimulation	-	++

5.3 Closed-loop Neural Interface System

5.3.1 Introduction

There are two fundamental types of control systems, the open-loop control system and the closed-loop control system [275]. In an open-loop control system, the control action signal is independent of the output of the plant under control. In a closed-loop control system, the control action signal is dependent on the output of the plant through the feedback loop. Fig. 5.15 (a) shows the block diagram of a typical closed-loop control system. The system consists of a sensor, an actuator, a closed-loop

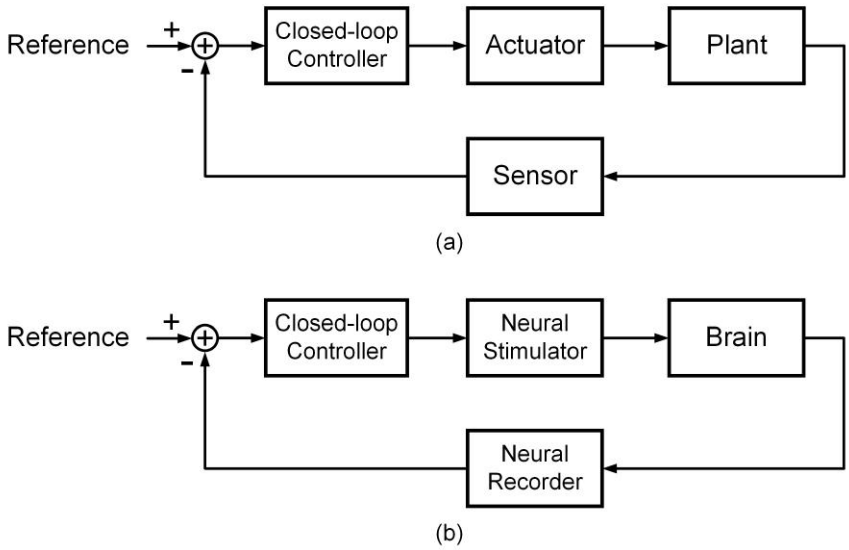


Figure 5.15: (a) The block diagram of a typical closed-loop control system. (b) The typical block diagram of a bi-directional closed-loop neural interface system. The neural recorder works as the sensor, and the neural stimulator works as the actuator.

controller and the plant under control. Ideally, the closed-loop controller generates the control action signal for the actuator to ensure the output of the plant is the same as the reference. The closed-loop control finds its applications in almost everywhere, not

only in electrical or mechanical engineering, but also in biology, climate science, social science, economics and finance, and other applications [276]. Fig. 5.15 (b) shows the block diagram of a bi-directional closed-loop neural interface system, where the neural recorder works as the sensor, and the neural stimulator works as the actuator.

There exist two types of feedback, the positive feedback and the negative feedback [276]. In a system with a positive feedback, the fed-back signal is in phase with the signal, while in a system with a negative feedback, the fed-back signal is out of phase with respect to the input signal. Both positive and negative feedback find applications in circuits and system design, but negative feedback is more applicable for its stability and accuracy of a system by correcting or reducing the unwanted changes. This is especially important for a neural interface since a positive feedback induced oscillation may cause permanent damage to the neural system.

There are many well-established control theory and stability compensation methods [121, 277]. The most commonly used closed-loop controller using a feedback mechanism is a proportional-integral-derivative (PID) controller. More than 95% of the closed-loop industrial processes use PID controllers [278]. The terms of a PID controller can be interpreted as corresponding to time: the proportional term depends on the present error, the integral term depends on the accumulation of past errors, and the derivative term is a prediction of future errors, based on the current rate of change [5].

This section presents the analysis and design of the closed-loop neural interface systems. The mechanisms of different closed-loop neural interface systems are first reviewed. Then, the design of a closed-loop neural interface with a general purpose PID controller is presented.

5.3.2 Mechanism of Closed-loop Neural Interface System

Fig. 5.16 shows the typical closed-loop brain-machine interface (BMI) systems with different control mechanisms. A prosthetic arm is used as an example of the BMI for illustration, and the PID controller represents any closed-loop controller. Fig. 5.16 (a) shows a basic bi-directional BMI. The electrical stimulator encodes the sensory information to the sensory cortex in the brain, and the brain generates the motor intent, which is decoded for the operation of the actuators. In other words, the closed-loop control policy origins from the brain. J. Liu *et al.* from the Michigan State University proposed an application aiming to improve the sensory encoding capacity of the BMI in 2011 [279]. Fig. 5.16 (b) illustrates the proposed method. The method involves an encoder mapping the sensory data acquired from the prosthetic to the desired patterns related to the somatosensory cortex activity. The errors between these desired patterns and those recorded in the somatosensory cortex are used in a PID controller to update the stimulation of the sub-cortical somatosensory areas in the thalamus or brainstem. This approach could elicit more continuous, natural sensory percepts compared to those evoked by the limited set of pre-programmed typical stimulation patterns [280].

Another closed-loop control mechanism is illustrated in Fig. 5.16 (c). In this case, the BMI system uses electrical stimulation to control a paralyzed arm rather than a prosthetic arm. The brain-controlled muscle stimulation has been shown to be a viable method of re-animating paralyzed arms in monkeys and humans [46, 268, 269]. In these studies, the muscle stimulation, and thus the arm movement trajectory, was entirely driven by motor cortex activity. However, prior work has shown that recording from pre-motor cortical areas to decode motor goals, not entire intended

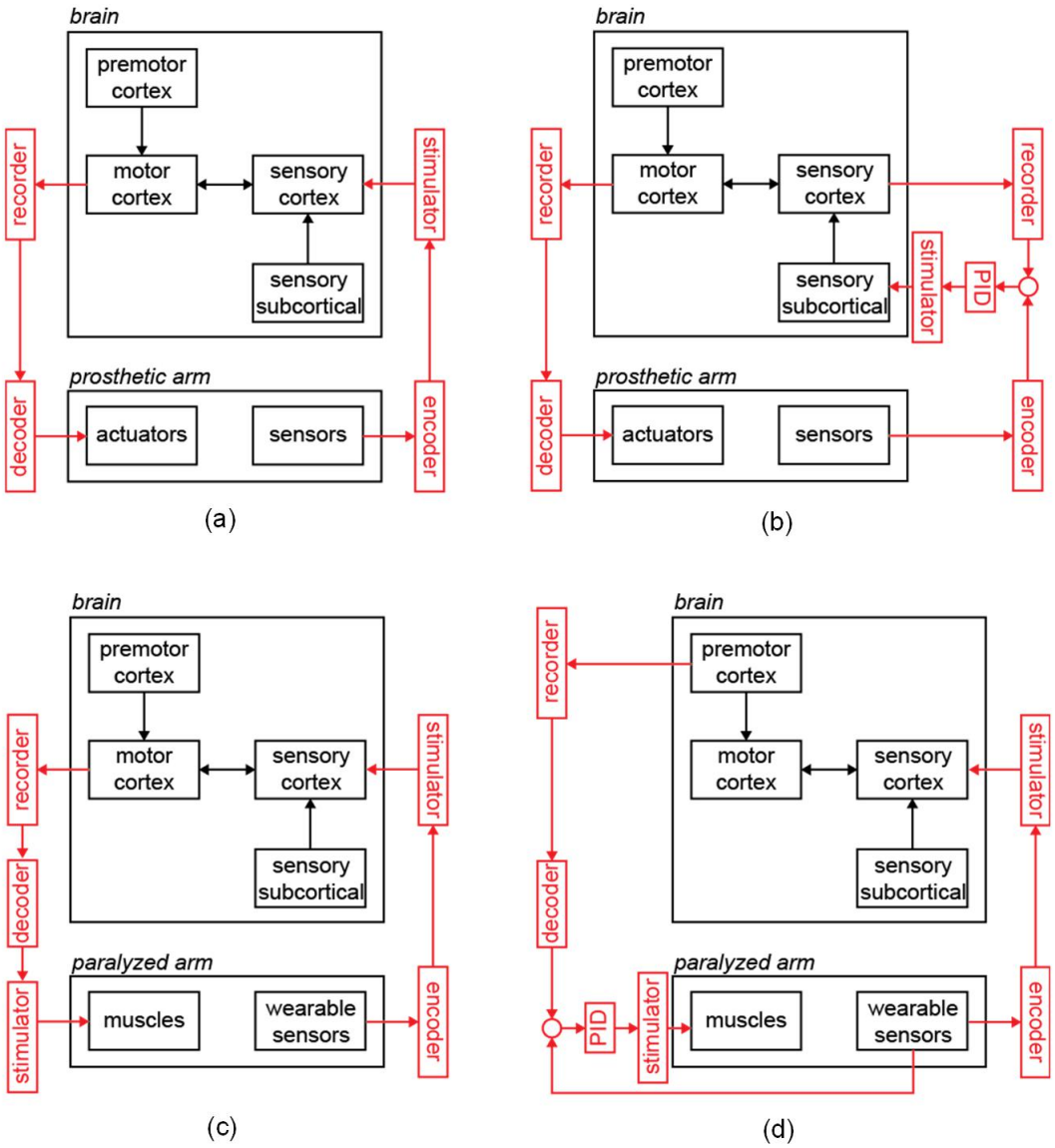


Figure 5.16: The block diagrams of different closed-loop BMI applications. The functions of the proposed neural interface system are shown in red. (a) Standard bi-directional BMI for a prosthetic arm. (b) Same as (a) but with improved sensory encoding method using a PID controller. (c) Bi-directional BMI to re-animate paralyzed arm by decoding desired arm trajectory. (d) Same as (c) but decoding motor goal and implementing arm trajectory with a PID controller.

trajectories, can improve performance and lower cognitive demand [80, 270]. Thus another potential BMI application for a closed-loop controller could be to update the

muscle stimulation based on the error between a decoded goal and the recorded state of the re-animated arm [281], as illustrated in Fig. 5.16 (d).

Besides prosthetics, other examples for the closed-loop bidirectional BMI applications include a deep brain stimulation (DBS) for Parkinson Disease and epilepsy. H. Rhew *et al.* from the University of Michigan, Ann Arbor proposed a closed-loop DBS system in 2014 [11]. The system detects the abnormal energy in the LFP and adjusts the stimulation current using a PI controller. W. Chen *et al.* from the National Chiao Tung University proposed a closed-loop neural prosthetic in 2014 [10]. The proposed design detects the seizure event and delivers a deep brain stimulation with parameters modulated from the internal brain state.

In addition, closed-loop neural interfaces are also important for electrophysiological studies [282]. A. Wallach *et al.* from the Technion proposed a neuronal response clamp in 2011 [263]. In this work, a closed-loop technique enabling control over the instantaneous response probability of a neuron was proposed using a PID controller. It has been used to characterize the input-output neuronal relationship. Sense-stimulate devices with closed-loop controllers have also been proposed for neuromodulatory applications [8]. In addition, the closed-loop stimulation of the sleep slow oscillation has been proposed to enhance memory [271].

In summary, there are many different configurations and mechanisms for using a closed-loop neural interface, and it is critical for a wide range of prosthetics and neuroscience research. Thus, the design of an energy efficient bi-directional neural interface with a closed-loop controller is highly desirable.

5.3.3 Design of a Closed-loop Neural Interface with a PID controller

5.3.3.1 Introduction

As reviewed in section 5.3.1, the PID controller is the most commonly used control loop feedback mechanism [5]. The brain is a highly non-linear, dynamic time-variant system, which can hardly be accurately modeled. A PID controller needs only the process variables and the target value, not requiring the knowledge of a system model or the underlying process. Thus the PID controller has a wide range of potential applications in the neuroprosthetic development and neuroscience research. However, it has not been reported to be integrated into the BMI hardware for on-chip closed-loop operation. In this work, a programmable PID controller in the analog domain has been designed to enable a variety of closed-loop experiments.

The basic working principle of the PID controller is briefly reviewed here. A PID controller calculates the difference between the desired reference and the measured output of the plant under test as the error value $e(t)$. The output of a PID controller $u(t)$ in the time domain is:

$$u(t) = K_p e(t) + K_i \int e(t)dt + K_d \frac{de(t)}{dt} \tag{5.4}$$

where K_p, k_i, and K_d are coefficients for the proportional, integral and derivative terms, respectively. By tuning the three parameters of the model, a PID controller can meet different process requirements.

There are four major characteristics of a closed-loop step response, including: i) the rise time, which is the time it takes for the system's output to rise to 90% of the desired level; ii) the overshoot, which is the peak level higher than the steady state, usually normalized against the steady state; iii) the settling time, which is the time it takes for the system to converge to the steady state; and iv) the steady-state (S-S) error, which is the difference between the steady-state output and the desired output. The effect of each controller parameter K_p, K_i, and K_d are summarized in Table 5.3. It should be noticed that the tuning is usually more complicated in practice.

Table 5.3: Effects of PID Parameters [5]

Parameter	Rise Time	Overshoot	Settling Time	S-S Error	Stability
K_P	Decrease	Increase	Small change	Decrease	Degrade
K_I	Decrease	Increase	Increase	Eliminate	Degrade
K_D	Minor change	Decrease	Decrease	No effect in theory	Improve if K_d is small

5.3.3.2 System and Circuit Implementation

The PID controller has been implemented in both analog [283, 284] and digital domains [285, 286]. The analog implementation has the advantages of low-power consumption and a compact layout, especially when the input and output are both analog signal. In this work, the PID controller is implemented in Gm-C blocks. The block diagram of the overall closed-loop system is shown in Fig. 5.17. The system consists of the closed-loop controller, a neural recorder, a neural feature extraction unit, a neural stimulator, and the buffers for connecting them. In the PID controller, the error signal is the difference between the extracted neural feature and a pre-set reference value. The output of the PID controller is a weighted sum of the proportional,

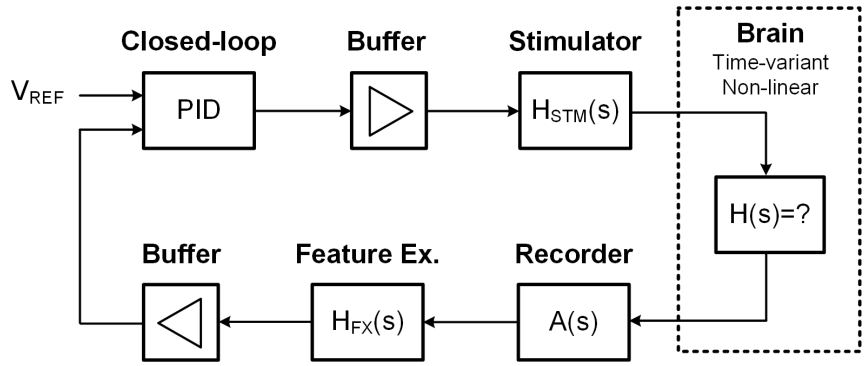


Figure 5.17: The block diagram of the overall system using the designed PID controller.

the derivative, and the integral terms of the error signal. The neural stimulator works as the actuator in the system. The output of the PID controller can be used to modulate the stimulating current amplitude, the stimulating frequency, or the stimulation pulse width. The sensor in the system is the neural recorder and the neural feature extraction unit. Any neural feature, including the spectral energy, the action potential firing rate, can be used as the input of the PID controller. In this work, the action potential fire-rate is calculated and converted to a voltage signal in the embedded MCU. However, this part can be easily integrated on-chip by using a frequency-to-voltage converter or an analog integrator.

The circuit schematic of the PID controller is shown in Fig. 5.18. The PID controller consists of 6 programmable Gm blocks and 2 capacitors. The circuit schematic of the Gm block and the biasing current generation module can be found in Fig. 3.4 and Fig. 3.6 from Chapter 3, respectively. The parameters for each of the P, I, and D components are independently programmable. The transfer function of the PID

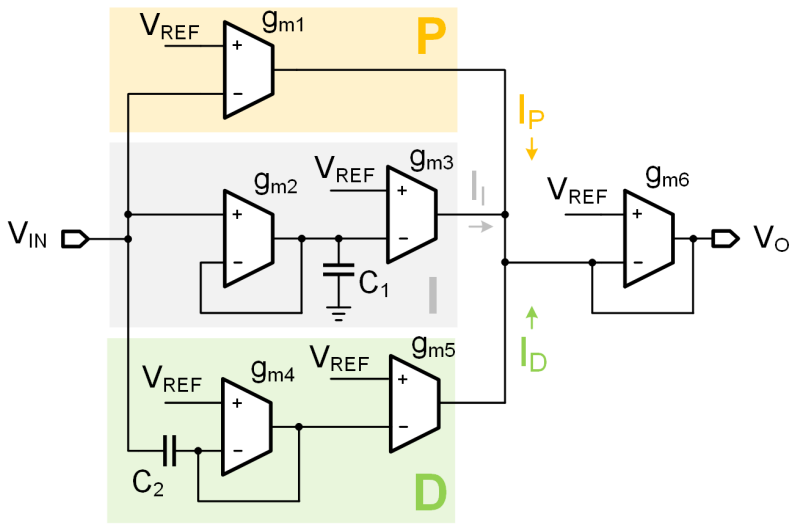


Figure 5.18: The circuit schematic of the designed PID controller.

controller is given by:

$$\frac{V_{out}(s)}{V_{err}(s)} = \frac{g_{m1}}{g_{m6}} + \frac{g_{m3}}{g_{m6}(1+\frac{sC_1}{g_{m2}})} + \frac{g_{m5}}{g_{m6}(1+\frac{g_{m4}}{sC_2})} \tag{5.5}$$

where V_{err}=$V_{ref} - V_{in}$. The gain of the P, I and D components are K_P=g_{m1}/g_{m6}, K_I=g_{m3}/g_{m6}, and K_D=g_{m5}/g_{m6}.

The basic parameter choosing and tuning of the PID controller have been reviewed in section 5.3.1. For a complex neural system where the accurate model can hardly be achieved, the plant exploration based method can be used. The initial estimation of the optimal operating points can be learned from the Zeigler-Nichols tuning method [287]. The final controller parameters can be determined by using an iterative procedure, based on the least-root-mean-square error. Considering the requirement of a BMI system, sufficient phase margin for stability must be guaranteed. Since an in-depth study of the control theory is not the focus of this work, the well

established PID tuning theory will not be discussed here. More information can be found in references [288, 289].

5.3.3.3 Experimental Results

The design has been fabricated in IBM 180nm CMOS technology. The micrograph of the fabricated chip and the layout of the PID module are shown in Fig. 5.19, with major blocks highlighted. The PID module occupies a silicon area of 100μm×75μm,

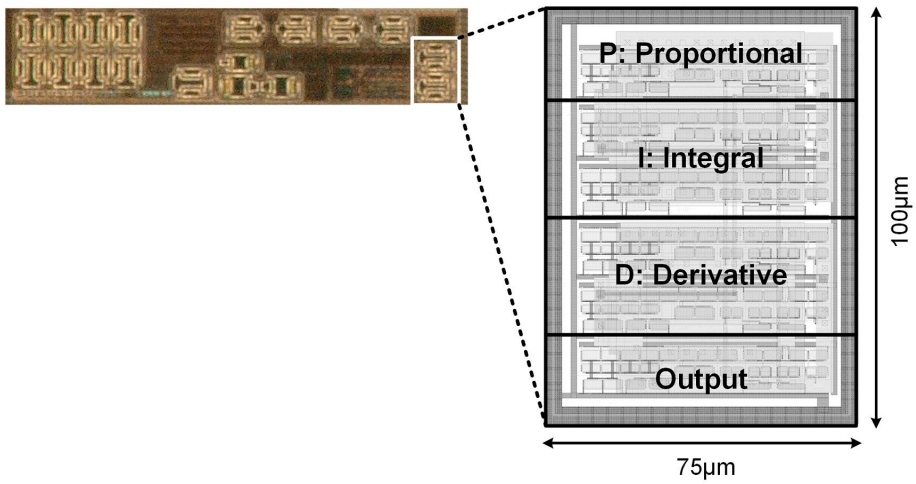


Figure 5.19: The microphotograph and the layout of the PID controller module.

including the digital registers. The bench testing was conducted to evaluate the designed PID controller. The *in-vivo* bi-directional closed-loop experiments will be discussed in the next Chapter.

The basic function and tuning of the PID controller were tested with a 2nd-order RC ladder network, as shown in Fig. 5.20. The transfer function of the RC network

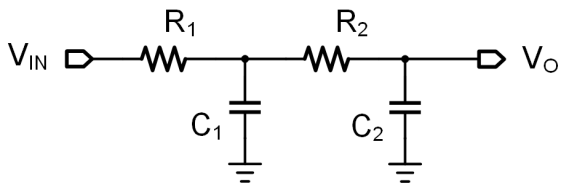


Figure 5.20: The 2nd order RC ladder network used for testing the PID controller.

is given by:

$$H(s) = \frac{\frac{1}{R_1 R_2 C_1 C_2}}{s^2 + s(\frac{1}{R_1 C_1} + \frac{1}{R_2 C_1} + \frac{1}{R_2 C_2}) + \frac{1}{R_1 R_2 C_1 C_2}} \tag{5.6}$$

The output of the RC ladder network is fed back to the PID controller and compared with a pre-set reference signal to find the error signal. A step change from 0.8V to 0.9V was given at the reference. The measured transient response of the system in different configurations are shown in Fig. 5.21. The design proves to be programmable over a large range, and useful in versatile closed-loop applications.

In addition, a closed-loop neuronal response clamp experiment [263, 282] was set up to test the proposed PID controller. The nervous system of man's and animals' response to the rapidly changing sensory information is highly variable with complex dynamics. The dynamic response is reflected from a single neuron to a neuronal network. Thus, it is important to study the behavior in a closed-loop approach in the appropriate context of a realistic input-output dependency. The voltage-, and current-clamps are well-known techniques [282] in the closed-loop electrophysiology. Recently, a dynamic neuronal response clamp technique was proposed to study the threshold dynamics of a neuron using extracellular stimulation and recording. A modified version of this technique is employed to test the proposed closed-loop system.

The diagram of the designed testing system is illustrated in Fig. 5.22. The major

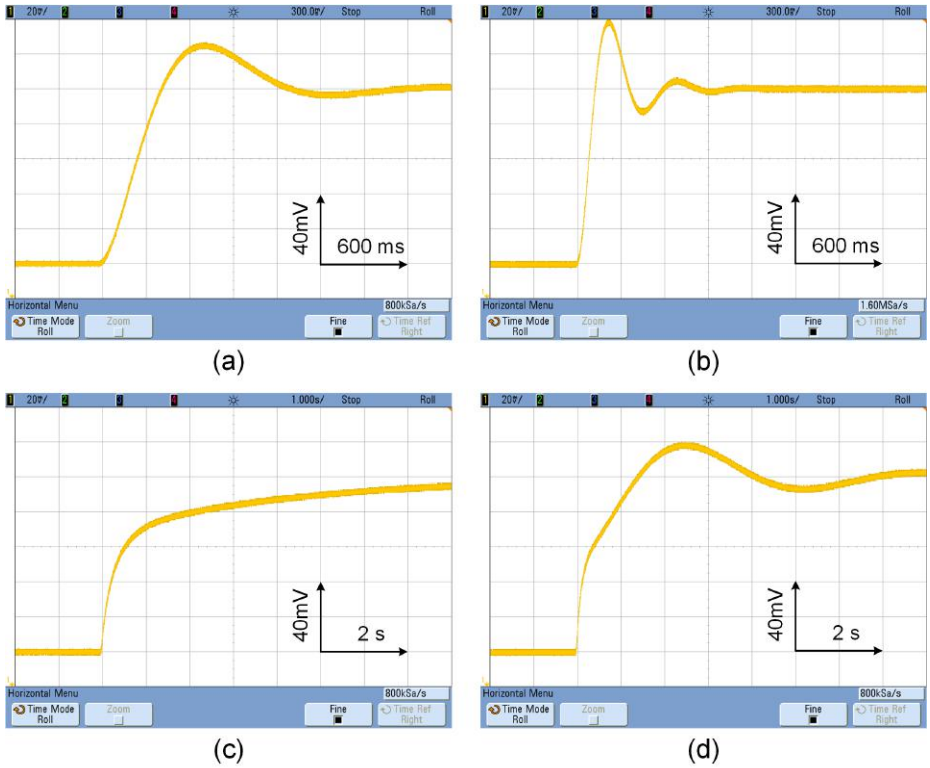


Figure 5.21: The measured transient response of the PID controller in different configurations. The ratio of the K_p, K_i, and K_d are (a) 1-2-0 (b) 2-1-0 (c) 1-0-1 and (d) 1-1-1.

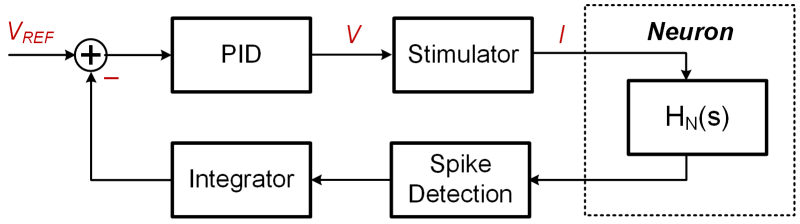


Figure 5.22: Diagram of the testing configuration of the PID controller. The major blocks used are the PID controller, a stimulator, an action potential detector, a neuron model, and a lossy integrator for finding spike rate. The on-chip stimulator is configured in the test mode to output continuous current. The lossy integrator is implemented in the MCU. The neuron is modeled by another MCU.

blocks used are the PID controller, a neural stimulator, an action potential detector, a neuron model, and a lossy integrator for finding spike-rate. The integrate-and-fire

model [206] for the single neuron employed in this experiment can be expressed as:

$$\tau_m \frac{dV}{dt} = V_m - V(t) + R_m I_s(t) \tag{5.7}$$

where $\tau_m \approx 10$ms is the membrane time constant, V_m is the resting membrane potential, $V(t)$ is the actual membrane potential as a function of time, $R_m \approx 10^7\Omega$ and $I_s(t)$ is the stimulation current. Once the membrane potential reaches a certain threshold V_{TH}, an action potential occurs and reset the potential back to its resting membrane potential. In this test, an off-chip microcontroller (Atmel XMEGA 128A4U [290]) with an integrated ADC and DAC was used to model the neuron. The MCU is running at a sampling rate of 100KHz, corresponding to a time resolution $dt = 10\mu s$. The continuous time differential equation is simplified by a discrete difference equation for the implementation in the MCU. The MCU's ADC measures the $R_m I_s[t]$, and the DAC generates $V[t]$ based on the following equations:

$$V[t] = \begin{cases} \dfrac{V_m + R_m I_s[t-1] + \tau_m V[t-1]}{1+\tau_m} & V[t-1] < V_{TH} \\ V_m & V[t-1] > V_{TH} \end{cases} \tag{5.8}$$

The stimulator was reconfigured in a testing mode to output a continuous stimulation current to meet the requirement of intracellular stimulation. The stimulation current amplitude was modulated by the output voltage of the PID controller. The neural model responded to the stimulation current, generating the membrane potential. The action potential detector module evaluated the membrane potential voltage with a pre-defined threshold voltage. The output of the detector was a PWM wave, which was sent to the integrator and converted to a voltage proportional to the spike rate. In this work, the spike rate was converted to a voltage in the embedded MCU.

The difference between the integrator's output voltage and the reference voltage was sent to the PID controller.

Fig. 5.23 shows 12 test trials with different proportional-integral-derivative parameters. The dots indicate the action potential's time stamps. The same reference

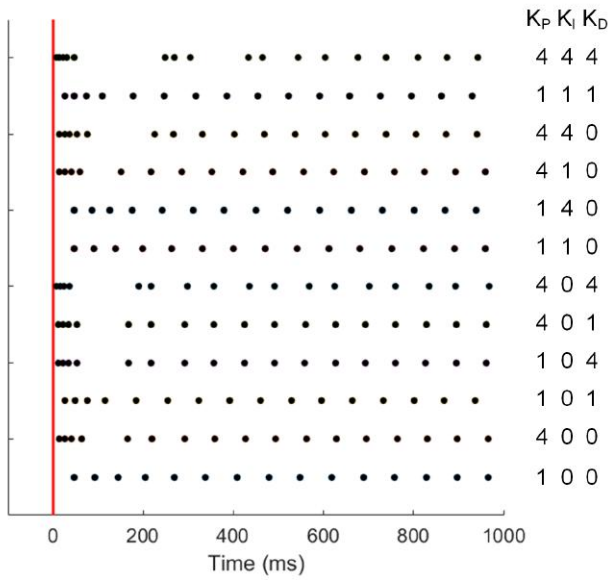


Figure 5.23: The measured transient response of the dynamic neuronal clamp with different PID parameters. Relative values of P, I and D components are shown on the right.

was set at time 0, the neuron responded to the stimulation current until it settled at a constant firing rate.

Fig. 5.24 shows 12 testing trails with different references. The neuron settled at a relative constant firing rate proportional to the reference, in a manner based on the choice of the P, I, and D terms. The test results showed that by programming the parameters, one can control the behavior of the neuron without the knowledge of the exact model [282].

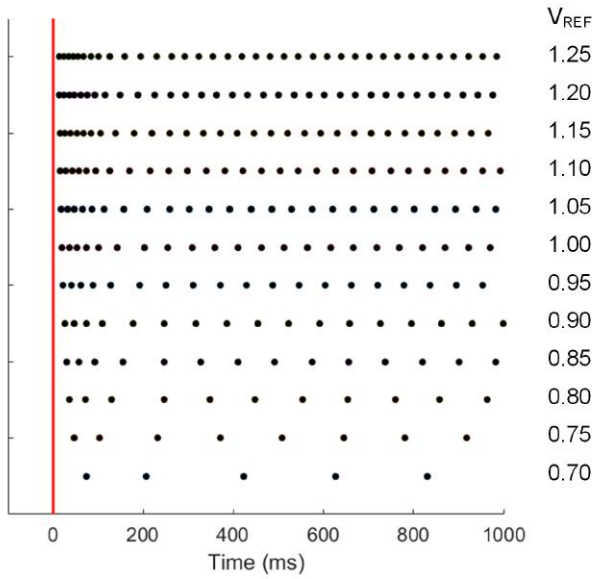


Figure 5.24: The measured transient response of the dynamic neuronal clamp with different references.

In summary, in this work, the importance of closed-loop control of the BMI device is analyzed, and the mechanisms of different closed-loop BMI devices are reviewed. As the most commonly used closed-loop controller, a general-purpose PID controller is selected to be implemented for the operation of a bi-directional closed-loop BMI device. The system and circuit implementation of the PID controller are presented, and the function of the designed chip has been evaluated in bench testing. The proposed design provides a promising solution for a wide range of neural science investigations.

Chapter 6

System Integration and Animal Experiments

6.1 Introduction

Behavioral and *in-vivo* animal experiments have been used through the history of biomedical research. The non-human animal experiments have become one of the most important methodologies in modern neuroscience research, and are highly valuable for the development of the brain-machine interface (BMI). However, most of the available medical instrumentation are designed for human medical treatments, which may not work well on animals. It is especially challenging if the designed experiment requires monitoring of the animal's brain activities or giving real-time stimulation feedback while the animal is freely behaving. In addition, the study of neural modulation and closed-loop control also requires a custom designed wearable or implantable BMI device to perform on-chip signal processing, feature extraction, classification,

machine learning, neuromodulation, and mapping. In summary, the design and integration of a wireless BMI for freely behaving animals is of great practical value and provides a very powerful tool for both neuroscience research and neuroprosthetic development.

Previous chapters have discussed the neural interface design from several perspectives. However, a complete system is more than the simple summing of individual building blocks. More importantly, many practical design issues are usually overlooked and underemphasized in the literature. A system that has been perfectly characterized in bench testing may not necessarily work well in an actual animal experiment. In this chapter, a custom system integration for animal experiments, especially during free behavior experiments is discussed. The methodologies and experimental results are presented in details.

The chapter is organized as follows. Section 6.2 presents the general purpose experimental platform, namely the PennBMBI. A custom designed command and communication protocol is presented, with a user-friendly interface. Wireless neural recording, stimulating and sensing functions have been verified in both anesthetized and awake rats. Section 6.3 presents a custom designed watermaze experiment for the study of augmenting perception through modulated electrical stimulation of somatosensory cortex. A waterproofed wireless neural stimulator and a complete animal tracking and neuromodulation experimental system are presented. Section 6.4 describes a bi-directional neural interface system for freely behaving monkeys. Long-term neural stimulation and recording during awake, sedated and sleeping monkeys are presented. A study in the hippocampal gamma-slow oscillation coupling using the developed system was also described. All procedures used in the studies presented

in this chapter were approved by the institutional animal care and use committee (IACUC) of the University of Pennsylvania.

6.2 The PennBMBI: A General Purpose Experimental Platform

6.2.1 Introduction

In this section, the design of a general purpose wireless brain-machine-brain interface (BMBI) system is presented. The system integrates four battery-powered wireless devices for the implementation of a closed-loop sensorimotor neural interface, including a neural signal analyzer, a neural stimulator, a body-area sensor node and a user-friendly graphic interface implemented on a PC. The neural signal analyzer features a four channel analog front-end with configurable passband, gain stage, digitization resolution, and sampling rate. The target frequency band is configurable from EEG to single unit activity. A noise floor of 4.69μVrms is achieved over a bandwidth from 0.05Hz to 6kHz. Digital filtering, neural feature extraction, spike detection, sensing-stimulating modulation, and compressed sensing measurement are realized in a central processing unit integrated into the analyzer. A flash memory card is also integrated into the analyzer. A 2-channel neural stimulator with a compliance voltage up to $\pm$12V is included. The stimulator is capable of delivering unipolar or bipolar, charge-balanced current pulses with programmable pulse shape, amplitude, width, pulse train frequency and latency. The system also includes a multi-functional sensor node, consisting of an accelerometer, a temperature sensor, a force sensor and a general sensor extension port. A computer interface is designed to monitor, control and configure all aforementioned devices via a wireless link, according to a custom designed communication protocol. Wireless closed-loop operation between the sensory devices, neural stimulator, and neural signal analyzer can be configured. The

proposed system was designed to link two sites in the brain, bridging the brain and external hardware, as well as creating new sensory and motor pathways for clinical practice.

6.2.2 System Overview

Fig. 6.1 illustrates the block diagram of the PennBMBI system. In general, four types of devices are required to interface with the brain, the body and the PC, including i) a neural signal analyzer, ii) a deep brain stimulator, iii) a smart sensor node, and iv) a PC interface board with the graphic user interface (GUI).

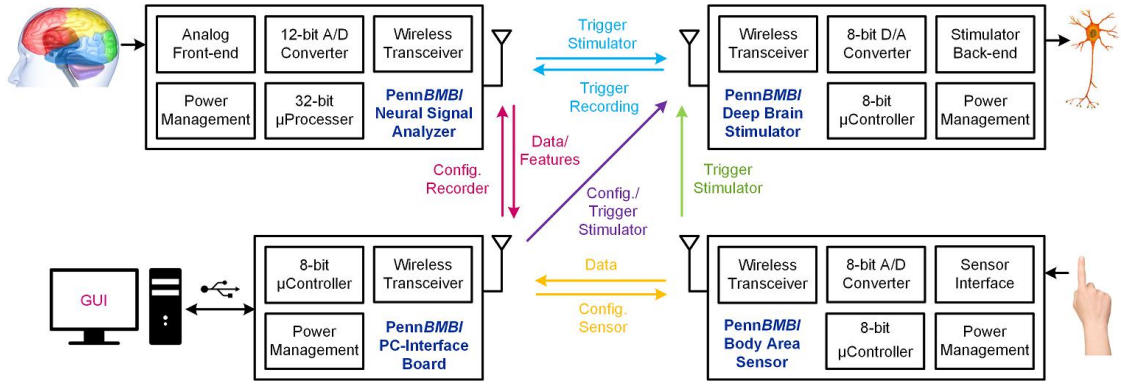


Figure 6.1: The BMBI system include four kinds of devices: neural signal analyzer, neural stimulator, body-area sensor node and computer interface. All devices can be configured wirelessly in a computer GUI. Possible closed-loop operation between devices are shown.

The first block includes a four-channel analog front-end with a high input impedance. The analog front-end records neural signals from EEG to single unit activity, with the strength of the input signal varying from less than 1μV to around 1mV, and the frequency band varying from DC to 10kHz. Configurable analog band-pass filters are used to suppress electrode offset, and to bandpass the signal in the frequency of

interest. An additional programmable gain stage and an analog to digital converter (ADC) with a programmable sampling rate and resolution are employed. Digital filtering, neural feature extraction, spike detection, sensing-stimulating modulation, and compressed sensing measurement are realized in a central processing unit integrated on board. The flash memory card is activated in the low-power operation mode, for compressed sensing recovery verification, and for data backup.

The second block is a dual-channel stimulator with a high driving capability that enables a charge-balanced current stimulation up to 400 μA with a compliance voltage of $\pm$12V for functional electrical stimulation. The device can be wirelessly controlled to deliver capacitively coupled current pulses with programmable pulse shape, width, pulse train frequency and latency.

The third block is a multi-functional body-area sensor node is included in the system, which enables communication of the sensory information to the brain with sensor-controlled wireless neural stimulation, and also enables the communication of the motor information to an effector with wireless neural recording and processing. The sensor node integrates a 3-axis accelerometer, a temperature sensor, a force sensor and a general extension port connected to an ADC, which can be used with different commercial sensors, such as pressure sensor, motion sensor, etc.

In addition, a graphic user interface has also been implemented for all device configurations, data acquisition, and simple signal analysis. A wireless link is implemented between all the devices for data transfer and on-line configuration.

6.2.3 Hardware Implementation

6.2.3.1 Neural Signal Analyzer

A Neural Signal Analyzer (NSA), with a dimension of 56mm × 36mm × 13mm, is designed to perform general neurological signal recording and analysis, as shown in Fig. 6.2. The NSA integrates a four-channel analog front-end, a central processing

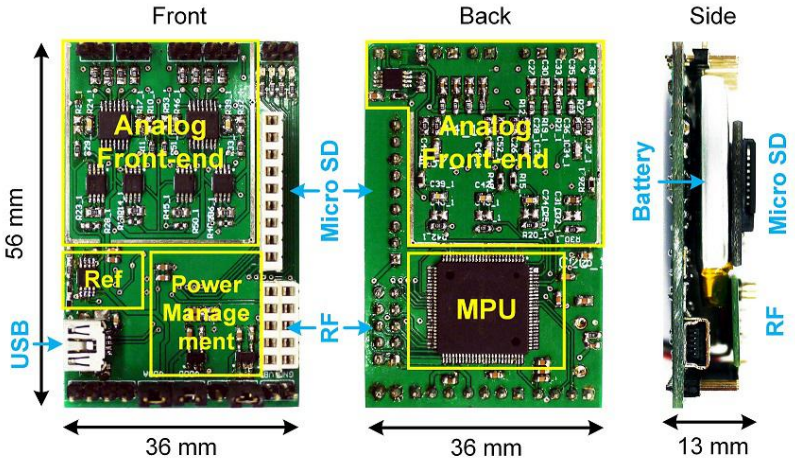


Figure 6.2: Photograph of the PennBMBI neural signal analyzer (NSA) in front, rear, and side view. The wireless module and Micro SD card are not shown in the front view.

unit (CPU), a 2.4 GHz wireless transceiver, a removable Micro SD card, a power management unit, and other peripheral circuits. An extension board holding the Micro SD socket can be plugged in through the connector as shown in the front and back view when necessary.

The analog front-end integrates four independent amplifier channels, sharing a tissue ground driving circuit. The block diagram of the front-end circuit is shown in Fig. 6.3. A supply voltage of 3.3V is used. The common mode voltage is set to be 1V. Configurable gain stages and filters are designed in order to meet the

requirement of recording neurological signals with different bandwidth and signal levels [92]. The differential input signal is AC coupled to the instrumentation amplifier with a high input impedance and a corner frequency of 0.5Hz. This is compatible with standard high impedance electrodes and removes the DC offset resulting from the electrode polarization. The gain of the instrumentation amplifier is fixed at 200, with a bandwidth of 12.5kHz [291]. Resistors with a low-temperature coefficient (TC) are used to minimize gain drift. An integrator implemented by amplifier A4 (Fig. 6.3) with a configurable capacitor is used as a high pass filter. Amplifier A5 is used to provide an additional gain stage with configurable low pass filter.

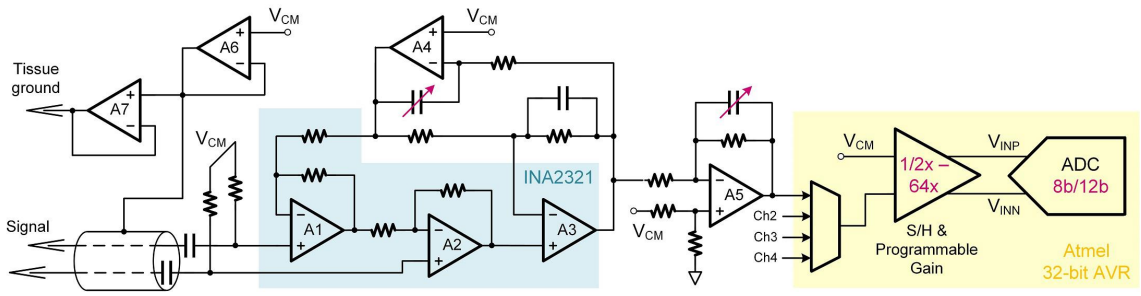


Figure 6.3: Architecture of the analog front-end. V_{CM} is the common mode voltage. A1∼A4 forms the instrumentation amplifier with high pass filter. A5 works as the second gain stage with configurable gain and low pass filter. The third gain stage has a programmable gain. The ADC digitizes the amplified neural signals at configurable sampling rates and resolution. A6 and A7 are shared by the four channels to drive the shield potential and tissue ground.

An Atmel 32-bit AVR Microcontroller AT32UC3C1512C [292] is implemented in the NSA. The MCU integrates a 12-bit pipeline ADC with a multiplexer, S/H circuit, and a programmable gain stage. A programmable gain ranging from 46dB to 102dB in total is achieved. In the recording mode, a peripheral direct memory access (DMA) controller is used for digital data acquisition, data buffering, and serial peripheral interface (SPI) accessing. Captured signals can be sent out via the wireless

module or to the Micro SD card through SPI. The DMA handshakes with peripheral interfaces directly, while the central processor core is in the sleep mode to save power.

On-line neural signal processing is performed in the 32-bit floating point DSP core in the MCU. Various function blocks are built, including,

- **Digital bandpass filter:** Type-I real Finite Impulse Response (FIR) filter is used as a bandpass filter. Six frequency bands are pre-defined, and the filter coefficients with 24 taps and 10 taps are pre-written in the flash memory for different filtering requirements.

- **Time-domain feature extractor:** Common time-domain features, such as line-length, area, energy, maximum/minimum, and zero-crossing, are extracted in real-time by a proper configuration of the sliding window length and overlay.

- **Spectral energy feature extractor:** 16/128-point FFT is used for spectral analysis.

- **Compressed sensing:** Neural signal features sparsity in certain basis/dictionaries [152], enabling a near lossless reconstruction under sub-Nyquist sampling. A signal agnostic compressed sensing measurement is implemented in the CPU. The input signal vector length N is set to be 512, and the measurement number M (M<N) can be programmable to 256, 128, 64, or 32. $y = \Phi x$ is realized as the compressed sensing measurement, where $x \in \mathbb{R}^{N\times 1}$ is the input neural signal, $y \in \mathbb{R}^{M\times 1}$ is the measurements, and $\Phi \in \mathbb{R}^{M\times N}$ is the measurement matrixes. Pseudo-random projection stored in the flash memory is used for the implementation of Φ. The reconstruction is performed on the receiver end using a convex optimization algorithm.

- **Action potential detection:** The filters are configured to first extract signals in the band of 300 to 6kHz. An amplitude threshold S_{th} is set for a rough unsupervised spike sorting for input signals in the frequency band of 300 to 6kHz. The value of S_{th} is four-times the estimation of the standard deviation of the background noise. Two time-amplitude windows are used to perform the discrimination of the action potentials after the input signal crosses the threshold with a positive derivative.

The NSA wireless transceiver can be configured to different operation modes, sending recorded raw data, neural features, spike time stamps, or compressed sensing measurements to the GUI, respectively. It also enables the sending of mapped stimuli patterns to the stimulating device, or receiving triggers for recording from other devices.

The NSA is powered by a rechargeable 3.7V lithium-ion battery (UBP002). A supply voltage of 3.3 V is used for the analog front-end, digital microcontroller and wireless transceiver. The quiescent current of the analog front-end is 380μA per channel. The CPU consumes 490μA per MHz. The 950 mAh battery supports the device for overnight continuous recording.

6.2.3.2 Neural Stimulators

A dual-channel neural stimulator with a size of 43mm $\times$ 27mm $\times$ 8mm, as illustrated in Fig. 6.4, is designed to deliver bipolar or unipolar, charge-balanced current pulses with programmable pulse shape, amplitude, width, pulse train frequency and latency. The stimulator integrates a current driving back-end, a microcontroller (MCU) with

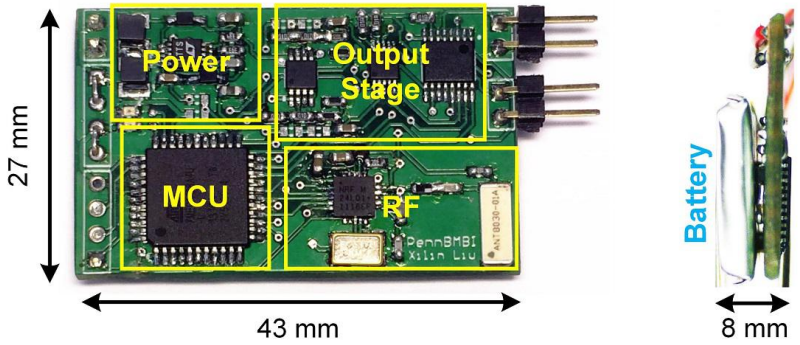


Figure 6.4: Photograph of the neural stimulator. The stimulator includes a current driving back-end, a MCU with integrated DAC and ADC, a wireless transceiver, a power management unit.

integrated DAC and ADC, a wireless transceiver, a power management unit, and other peripheral circuits.

A dual DC-DC converter is used for boosting the voltage from a 3.7V lithium-ion battery to ±12V to drive the output current stage, in order to provide a sufficient compliance voltage for stimulating through high impedance electrodes. The converter will be switched to idle mode when no stimulation is to be delivered in order to reduce power consumption. A modified Howland current source is employed as a bi-directional current driving stage, as illustrated in Fig. 6.5. Amplifiers A1 to A4 are implemented using high-voltage dual supply op-amps with JFET inputs. A resistor trimmer is used to trim the equal value resistor network to achieve good common mode rejection ratio (CMRR) and high output impedance from the Howland current source. A feedback capacitor is added for stability. Different transconductance can be selected by setting the gain resistors in order to get a large dynamic range. Amplifier A2 is a unity-gain buffer used to reduce the requirement for calibration under different gain settings.

A feedback integrator, A3, is used in an idle mode to improve the stability as well

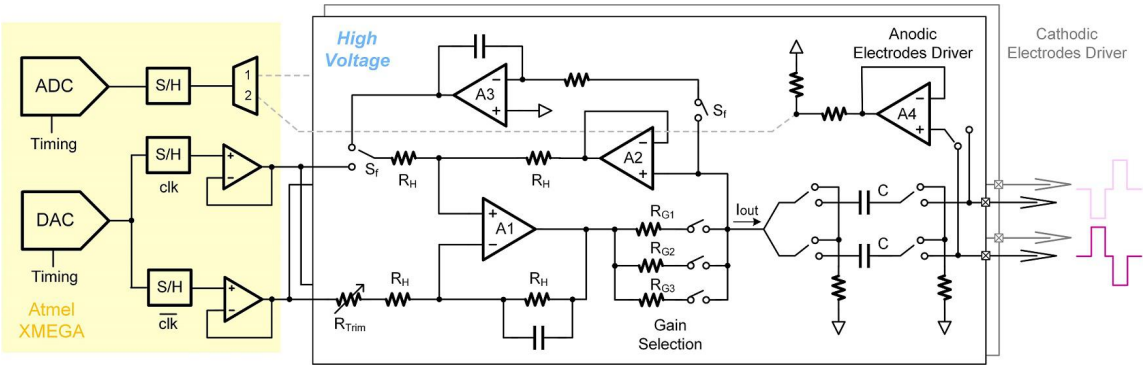


Figure 6.5: Circuit schematic of the high compliance voltage current source output stage. Arbitrary stimuli waveform is generated by a digital to analog converter. V-to-I gain is programmable to provide a high dynamic range.

as to reduce the current leakage [6]. Amplifier A4 is implemented for buffering the electrode potential, and the impedance of the electrode is calculated in the MCU. A low impedance threshold is set to stop stimulating in the case of electrode shorting. In addition, a blocking capacitor is used in each channel to prevent direct current injection and limits the maximum net charges.

The two channels are used as differential input of the Howland current source to minimize the offset. The DAC is shut down and both inputs are grounded in idle mode to reduce power consumption. The ADC is triggered twice during the stimulation phase to estimate the compliance voltage on the electrode, as well as to evaluate the impedance. If the electrode impedance is lower than a user-defined threshold, all the stimulation will be stopped and an alert will be sent to the computer.

6.2.3.3 Body Area Sensors

The multi-functional body area sensor node has a dimension of 31mm × 13mm × 8mm, as shown in Fig. 6.6. The sensor node integrates a microcontroller, a 3-axis

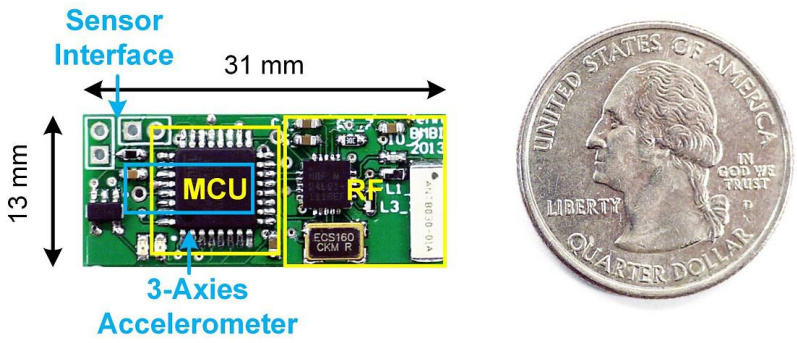


Figure 6.6: The photograph of sensor node in comparison with a Quarter.

digital accelerometer, a temperature sensor, and a flexiforce sensor.

The accelerometer interfaces with the MCU through I2C protocol. The outputs of the thermistor and the force sensor are analog signals, which are digitized using an 8-bit SAR ADC integrated in the MCU. General-purpose ports are saved for up to 12-channel potential extensions of the sensor node. The sensor node is powered by a 2.65g, 110mAh rechargeable lithium battery. The power consumption of all the modules used in the sensor node is listed in Table 6.1.

Table 6.1: Power consumption of the sensor node

Microcontroller	240μA	Accelerometer [293]	23 μA
RF Sleep [294]	0.9μA	Thermistor	$< 7\mu$A
RF Transmit	7mA	Flexiforce sensor	$< 50\mu$A
Total Working	321 μA	Total Transmit	7.3 mA

6.2.3.4 Computer Interface

A PC interface board with high-speed USB 2.0, and a Matlab-based graphic user interface (GUI), as illustrated in Fig. 6.7, have been built for wireless monitoring,

controlling and configuring all devices. The closed-loop operation can also be easily

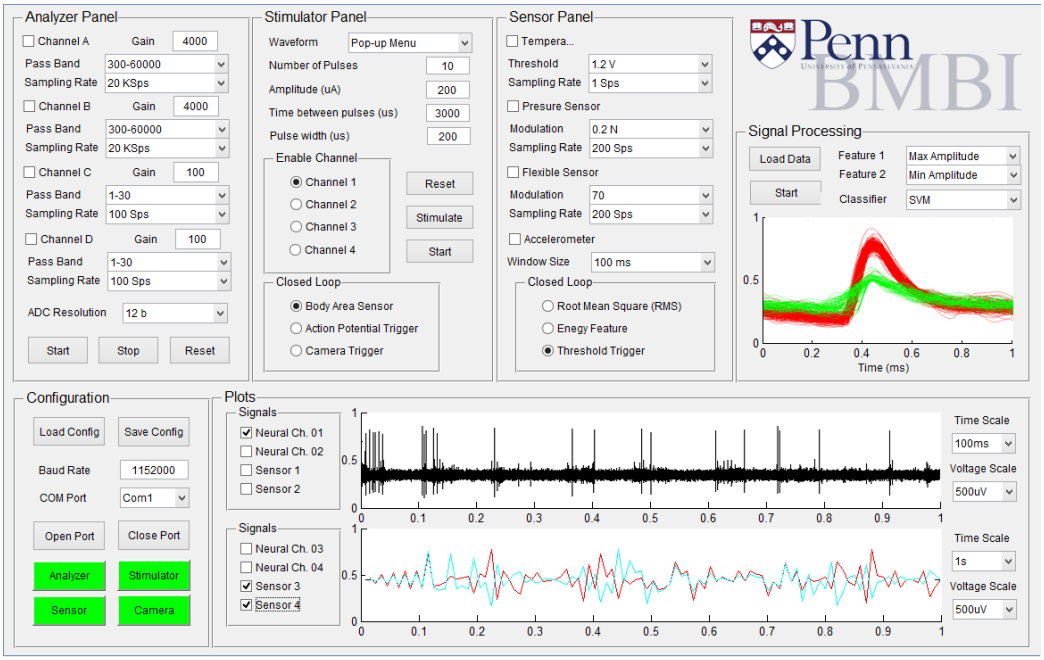


Figure 6.7: Matlab based Graphic User Interface (GUI). Six major panels are included in the GUI, which are: 1) PC configuration; 2) recording device configuration; 3) stimulator configuration; 4) body-area sensors configuration; 5) closed-loop configuration; 6) display windows.

configured in the GUI. All the devices receive interprets and commands from the PC GUI and talk to the target device through a corresponding channel/address via a wireless link. There are six major panels of the GUI, including: 1) PC configuration panel, where the communication port can be configured. All the configurations (including other panels) can be exported or loaded; 2) analyzer configuration panel, where the gain, sampling rate/resolution, filter pass-band can be configured for each individual channel. For the signal processing modes performed in hardware, the time window size and threshold for spike detection can also be configured; 3) stimulator configuration panel, where the amplitude, pulse width, pulse train number, and time interval of the stimuli can be configured; 4) body-area sensors configuration, where

parameters for sensor nodes can be configured; 5) closed-loop configuration, where closed-loop operation between different devices can be configured; 6) display windows, where the output from analyzers and sensor nodes can be displayed in real time.

6.2.4 Experimental results

6.2.4.1 Bench Testing

Fig. 6.8 shows the measured input referred noise spectrum of the analog front-end in the neural signal analyzer. The integrated noise is 4.69μVrms in a wide band.

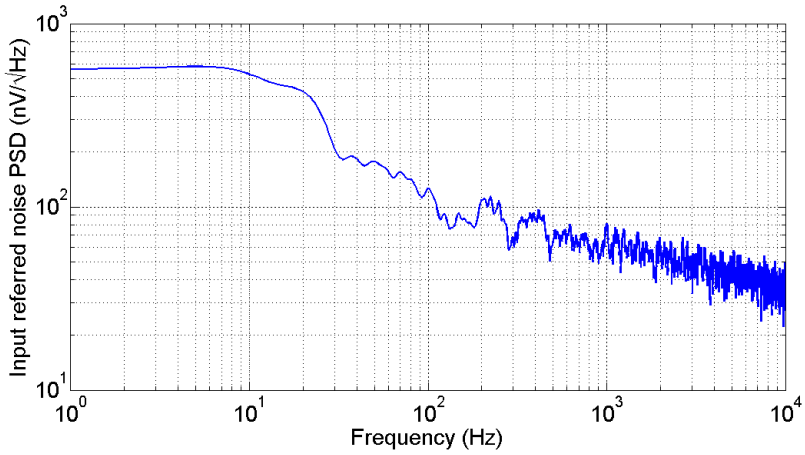


Figure 6.8: Input referred noise spectrum of the analog front-end.

The noise efficiency factor [104] is 14.6. The mid-band gain error is 0.87% and the measured CMRR at 1kHz is 67.4dB. The measured frequency responses in different configurations are shown in Fig. 6.9. Bandpass filters with different gain are integrated, e.g. 10 to 200Hz with a gain of 66dB, 300 to 6kHz with a gain of 66dB, and 10 to 200Hz with a gain of 78dB.

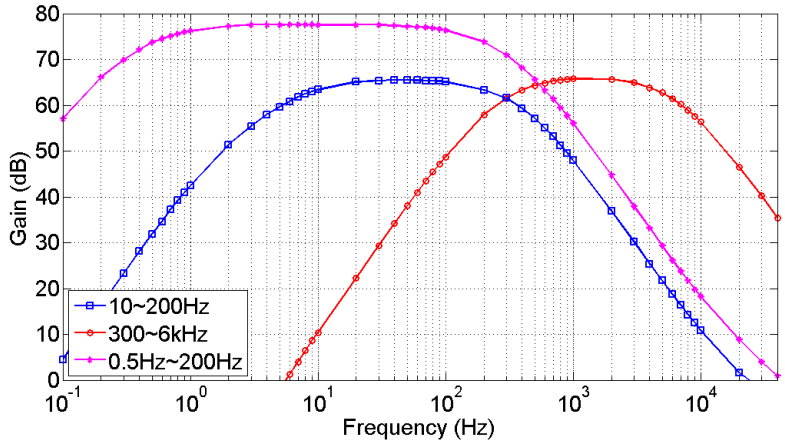
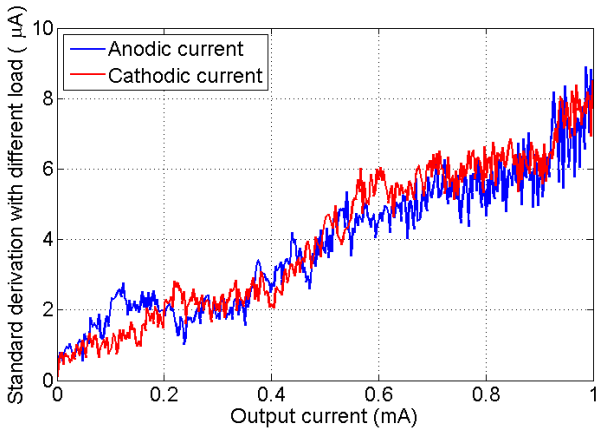


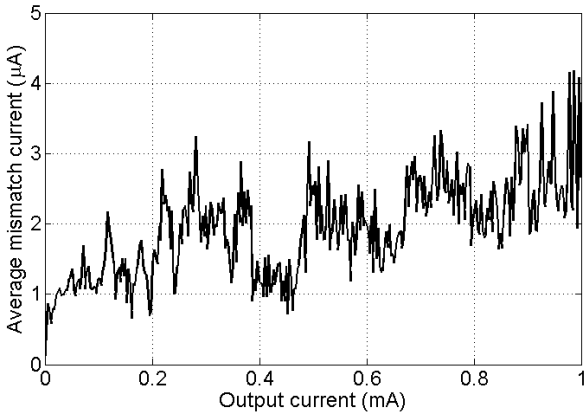
Figure 6.9: The measured frequency response of the analog front-end in different configurations. (blue) 10 to 200Hz with a gain of 66dB, (red) 300 to 6kHz with a gain of 66dB, (magenta) 10 to 200Hz with a gain of 78dB.

The output currents of the neural stimulator are measured under different loads. Fig. 6.10 (a) shows the standard deviation of the output current of anodic and cathodic drivers across the different loads. The standard deviation is calculated for each output current with all the different loads. The average of the calculated standard deviation for different currents in the output stage is 3.91μA. Fig. 6.10 (b) shows the average current mismatch between the anodic and cathodic electrodes across different loads. The average mismatch with respect to the corresponding output current is 0.75%. The stimulator is also tested in 0.9g/100mil Sodium Chloride using a 75μm tungsten electrode. Fig. 6.11 shows the measured voltage across the bipolar electrodes for different stimulation current levels. A blocking capacitor of 1μF is used. The characteristics of the neural signal analyzer and the neural stimulator are summarized in table 6.2. A lower than 10^{-3} bit error rate (BER) is measured in the wireless module for a distance of 3m in a normal animal experiment environment.

Two open-loop experiments have been performed to verify the system level operation. In the first experiment, the NSA to stimulator pathway is tested. As shown in



(a)



(b)

Figure 6.10: The measured output current for different loads (1kΩ, 2kΩ, ..., 8kΩ). (a) shows the standard deviation of the output current across different loads. (b) shows the average current mismatch between anodic and cathodic electrodes across different loads.

Fig. 6.12(a), the neural signal is first captured by the NSA. On-board AP detection is performed using a dual threshold comparison method. A pass window is generated when the input signal crosses the threshold. An AP is denoted when two pass windows are detected. Once an AP is detected, as shown in the zoomed-in view in Fig. 6.12(b), a CMD CFG is wirelessly sent from the NSA to the stimulator, triggering a group of pulse stimulation. In the second experiment, the sensor to stimulator

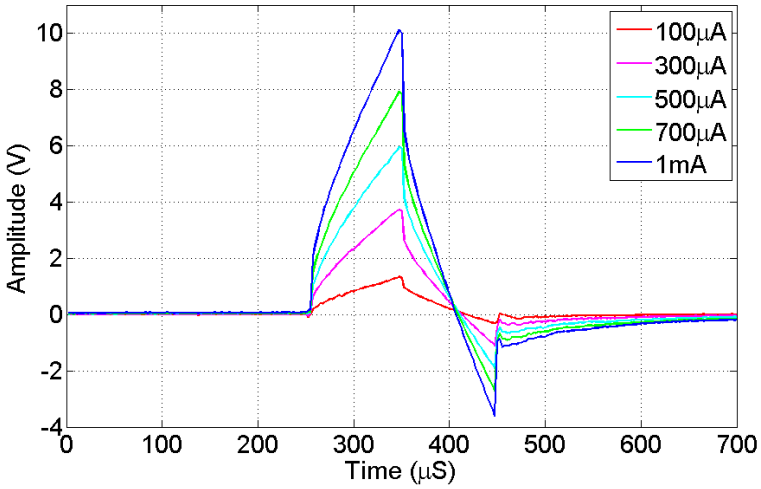
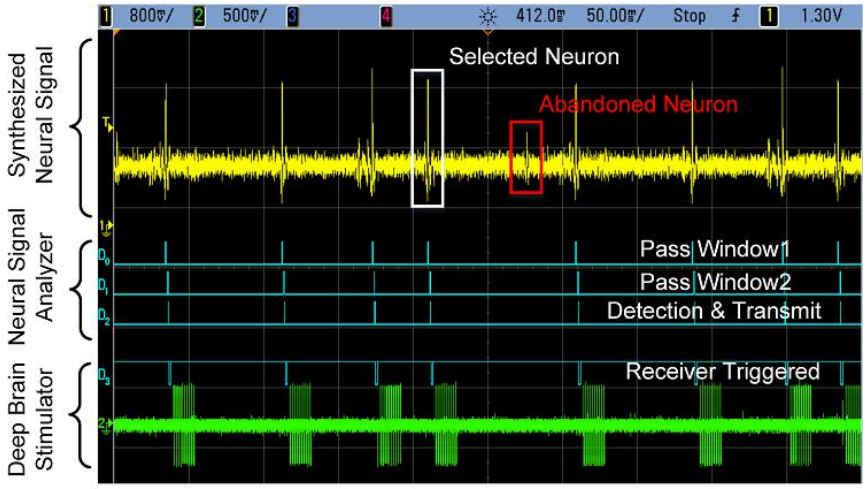


Figure 6.11: Measured stimulation pulses with different amplitudes in Sodium Chloride solution.

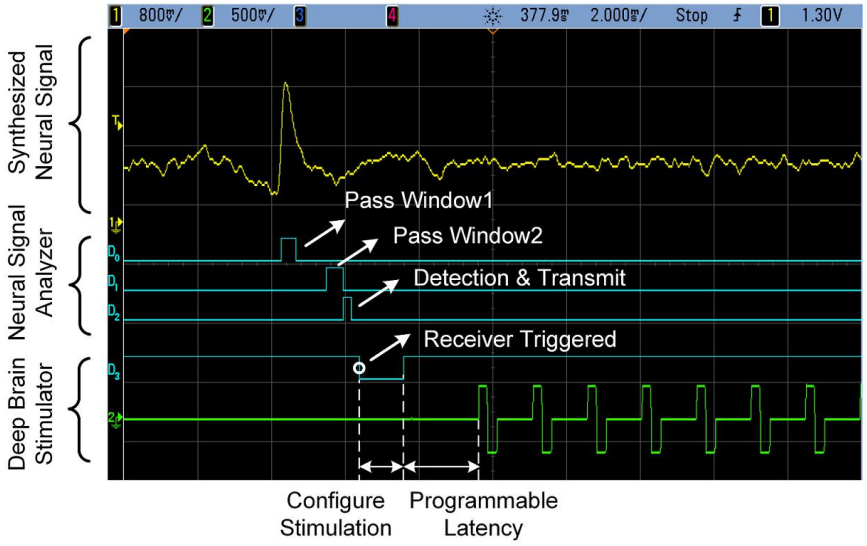
Table 6.2: Specifications of the PennBMBI system

Neural Signal Analyzer			
Supply voltage	3.3V	Supply current	380μA/ch
Input Impedance	>200MΩ	Offset tolerance	1V
I-Amp Noise floor	4.69 μVrms	I-Amp CMRR	>61dB
ADC resolution	12 bit		
Neural Stimulator			
Output current	0 $\sim$ 1mA	DAC resolution	6 bit
Compliance voltage	$\pm$12V	Output impedance	> 100MΩ
Standard deviation	1.71 μA	Driver mismatch	0.75%

pathway is tested. As shown in Fig. 6.13, the amplitude of the sensing result is encoded into the frequency of the pulses generated from the stimulator. A CMD CFG command is wireless continuously sent to the stimulator from the sensor node. The argument is encoded according to the digitized output of the sensor node.



(a)



(b)

Figure 6.12: Measured wireless closed-loop operation from the neural signal analyzer and the stimulator. Zoomed-in view of (a) is shown in (b).

6.2.4.2 *In-Vivo* Testing

To further evaluate the PennBMBI, we performed several basic tests of the wireless neural recording, stimulating and sensing functions in both anesthetized and awake rats. The neural recording was performed in an anesthetized rat with a tungsten

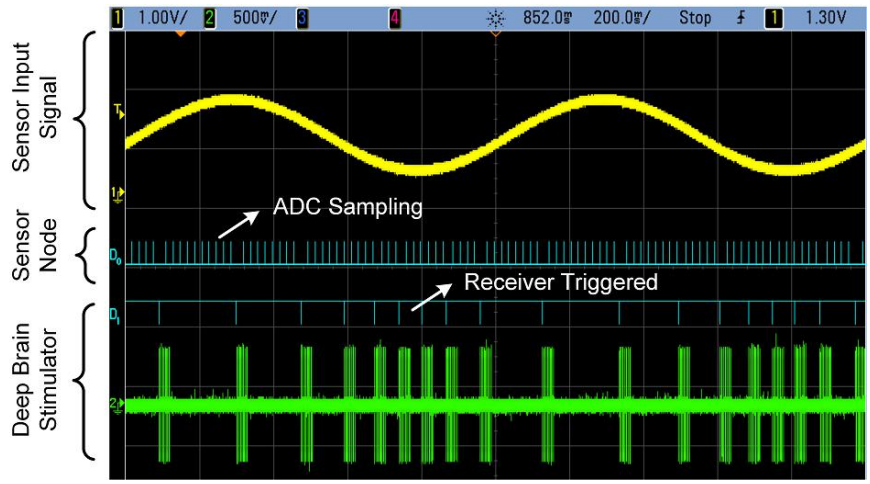


Figure 6.13: Measured wireless closed-loop operation from sensor node and deep brain stimulator.

microelectrode placed in the whisker motor cortex. The analyzer was configured to have a passband of 300∼6KHz, a sampling rate of 21 KSps, and a gain of 72dB. The recorded action potentials (APs) are shown in Fig. 6.14. The results show that the neural signal analyzer faithfully recorded the APs with a signal-to-noise ratio comparable to a commercial system.

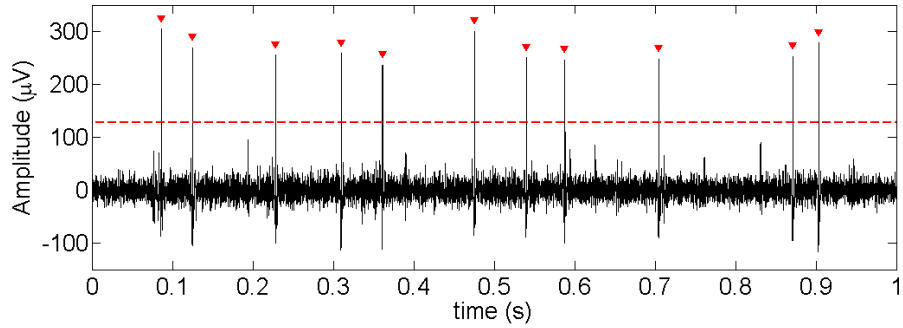


Figure 6.14: Action potentials recorded by the neural signal analyzer. Detected spikes are marked by red triangles.

In order to evaluate the quality of the captured data, the neural signal was

simultaneously recorded by a rack-mounted commercial system (RZ2 Workstation, Tucker-Davis Technologies). A comparison of the signals recorded by the two systems is shown in Fig. 6.15. The recording shows two different neurons firing APs in close

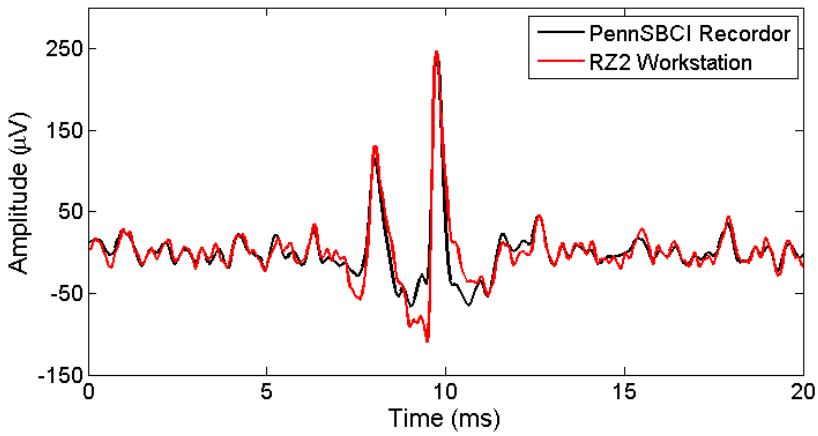


Figure 6.15: Comparison between data captured by the PennBMBI analyzer (black) and the RZ2 Neurophysiology Workstation (red).

succession.

To demonstrate the sensor and stimulator nodes, an awake rat with a chronically implanted stimulating microelectrode in the lateral hypothalamus was placed in an operant conditioning chamber with a lever press. The sensor node detected the lever press and wirelessly sent a trigger to the stimulator worn on the rat's back to deliver a stimulus train (30 of $100\mu A$, $200\mu s$ constant current pulses) to the micro-electrode. This setup allowed the rat to associate the lever press with the rewarding sensation of hypothalamic stimulation. This result provides one example of how the various nodes of the PennBMBI, in this case, the sensor and stimulator, can be flexibly combined to enable a wide range of neuroscience and neural engineering experiments in freely behaving animals.

6.3 The Watermaze

6.3.1 Introduction and Background

Sensation and perception are essential abilities for human and animals. Loss of sensation due to nerve damage prevents even basic activities of daily living. Even though many recently developed neutrally-controlled prosthetics successfully replaced motor pathways, somatosensory feedback is critical for paralyzed individuals to adequately use them, which has often been underscored. Recently, there has been an increased interest in conveying lost information through direct brain stimulation using a neuroprosthetic device [38]. These strategies rely on the brain learning to use remapped or artificial stimuli to inform actions. A common paradigm to study this learning process involves using brain stimulation to guide rats through a maze [295]. These so-called "rat-robot" studies have mapped a number of different navigation signals to brain stimulation [296]. All used land-based mazes with a discrete number of actions and goal locations. A concern with these studies is that the rats could simply memorize a few stimulus-response contingencies rather than learn a more generalized stimulus-dependent navigation strategy.

In this work, a new rat-robot paradigm is developed using a classic test of rodent navigation: the Morris water maze (MWM). In the MWM, the rat swims in a large circular tank looking for a hidden, submerged platform on which to stand [297]. In our task, the submerged platform was positioned randomly on each trial to dissociate visual cues from the platform location, and the rats navigated to the platform using only the sensation encoded from the brain stimulation. The experiment setup is illustrated in Fig. 6.16. For simplicity, the experiment system is referred to as

watermaze system. Custom hardware and software were developed to support the watermaze experiment. The findings suggest that rats can quickly interpret artificial percepts to guide behavior, which is important for sensorimotor neuroprostheses.

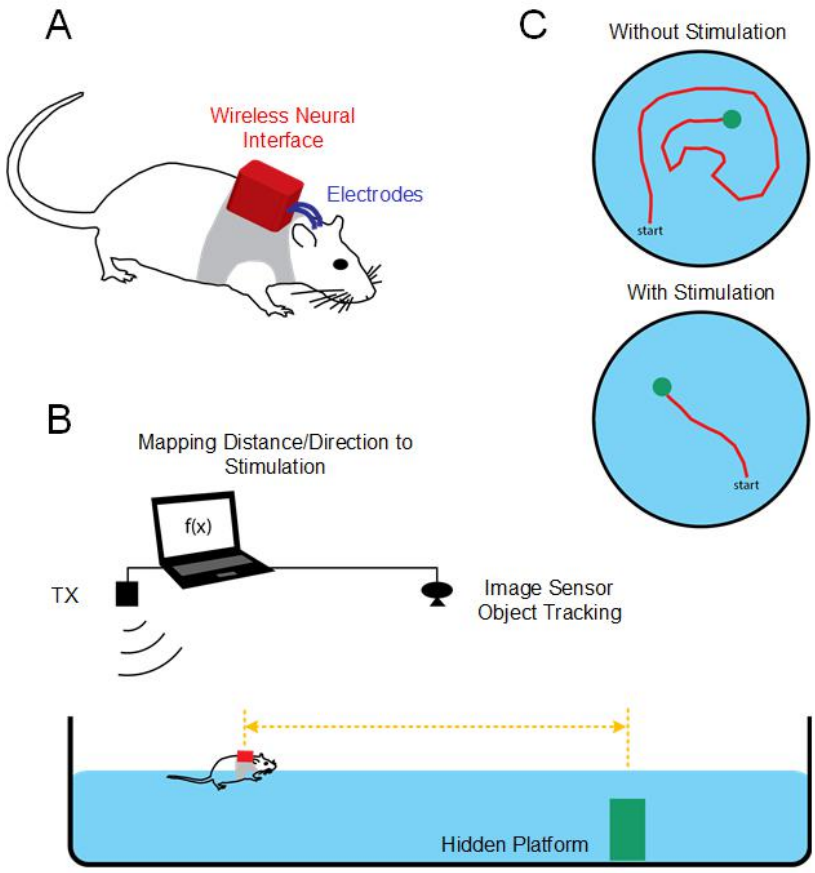


Figure 6.16: Illustration of the developed perception augmentation experiment. (a) shows a rat wearing the developed wireless waterproof neuroprosthetic. The electrodes are chronically implanted in the somatosensory cortex. (b) shows the experimental setup. A rat navigates to a hidden platform using only the perception established from the stimulation. (c) illustrates the estimated rat's swimming traces with/without the simulation guidance.

The watermaze experiment was designed by Dr. Andrew Richardson, and was conducted by Yohannes Ghenbot, Sam Deluccia, Solymar Maldonado, Gregory Boyek

and other research fellows and students in the Translational Neuromodulation Laboratory, University of Pennsylvania. The data analysis was performed by Dr. Andrew Richardson and Yohannes Ghenbot.

6.3.2 System Overview

The block diagram of the watermaze system is shown in Fig. 6.17. The watermaze

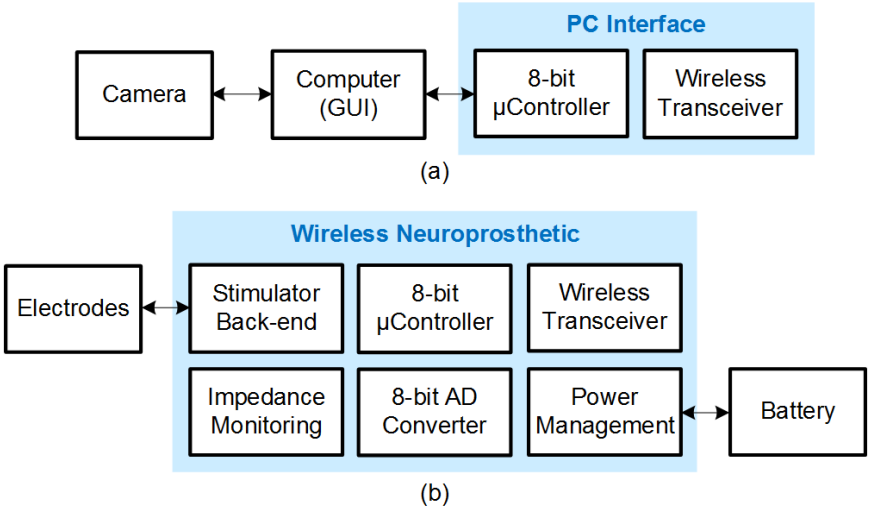


Figure 6.17: The block diagram for the watermaze system. (a) shows the computer with graphic interface, camera, and a wireless dongle to the wireless neuroprosthetic device.

system including both hardware and software. The hardware system consists of i) a wireless neuroprosthetic device, and ii) a computer with a camera and a wireless dongle. The software system consists of i) the animal tracking and modulation algorithm, and ii) the communication and stimulation protocol in the neuroprosthetic device and the computer. It should be noticed that, even though the system is optimized for this experiment, it can be generalized to perform many similar neuroscience experiments.

The developed perception augmentation experiment can also be understood as a typical closed-loop system, as illustrated in Fig. 6.18. The wireless neuroprosthetic device works as the actuator, while the tracking image sensor finds the error signal, which is the distance between the rat's location and the hidden platform, or the heading of the rat.

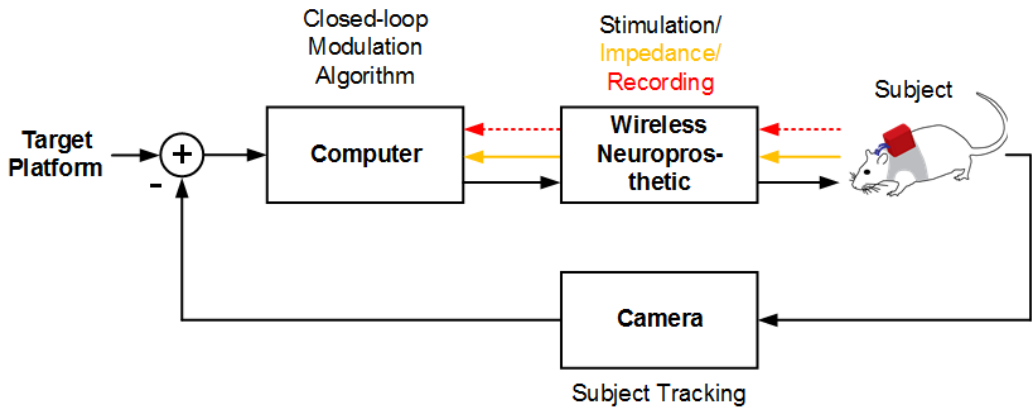


Figure 6.18: The typical closed-loop diagram for the developed perception augmentation experiment.

6.3.3 Hardware Implementation

6.3.3.1 Design of the Watermaze Stimulator

The block diagram of the watermaze stimulator is shown in Fig. 6.17. The watermaze stimulator consists of 1) a micro-controller for overall processing and control, 2) a stimulator back-end for driving stimulating electrodes, a DAC might be included, 3) a wireless transceiver for communication, 4) a power management unit for powering the whole device, and 5) an impedance monitoring module for reading back

the compliance voltage of the electrodes for monitoring the electrode impedance and guarantee the safety of the electrode during the experiment.

In this work, an 8/16 bit microcontroller ATxmega128A4U [290] from Atmel is used as the central processor. It communicates with a 2.4GHz wireless transceiver from Nordic Semiconductor nRF24L01 [294] for retrieving device configuration and stimulation commands, and sends back the compliance voltage for estimating the electrode impedance during the stimulation.

The power management unit includes a single channel LDO TPS791 [298] from Texas Instruments for powering the microcontroller and some peripheral circuits. The chip has a full-scale output current of 100mA, with a very low dropout voltage of 38mV. The RMS noise is 15μV. A dual channel DC/DC converter LT1945 [299] from Linear Technology is used to generating high voltage for powering the stimulating output stages. The chip takes input as low as 1.2V, so potentially can be powered by coin batteries. Regulated positive and negative outputs can be generated up to $\pm$ 34V, setting by feedback resistors' ratio. The converter consumes 12μA in the active mode and less than 1μA in the shutdown mode. In the first two generations of the watermaze stimulators, an inverting charge pump LTC1983 [300] from Linear Technology is used to generate negative supply. The chip gives fixed -3V with $\pm$4% accuracy with an input voltage from 2.3V to 5.5V. The full-scale output current is 100mA, with a flyback capacitor of 1μF. The battery used in the first generation is LP-402025 from Sounddon. A 150mAh Lithium Ion Polymer battery from Pkcell is used in the following generations.

The first generation watermaze stimulator has a two PCB layers structure, and a wireless transceiver with PCB antenna. The circuit schematic for the output stage

is shown in Fig. 6.19. The core circuits are two Holland current sources (A_1-A_5). The output current of the current source is the input voltage over the gain resistor, independent of the load impedance. A blocking capacitor C is used to block the DC current path to the brain. The DAC integrated in the microcontroller is used to set the input voltage. The output range of the DAC is ground to reference voltage, which is set to be VDD. A_1 works as a level shifter, shifting the output range around the ground. A trimming resistor is used to tune the output voltage. The DAC generate the stimulation voltage output waveform, while the driving sites convert the voltage waveform to the current in a programmable transimpedance by tuning the gain resistor R_G. $A_2 - A_5$ are designed using high voltage op-amps OPA2140 [301] from Texas Instruments. These op-amps operate with dual supplies up to ±18V. In this version of watermaze stimulator, the supply voltages are designed to be ±15V to provide up to $300uA$ into a load of 50kΩ. The switches S_C, S_A, and S_D are for connecting the output stages to the driving electrodes. More than one pair of bipolar electrodes can be designed by adding a multiplexer without much area and power penalty. The only drawback is the lack of ability in driving two stimulating sites simultaneously. However, a near simultaneous stimulation by switching the electrodes are more than sufficient in most cases. When no stimulus is to be delivered, the op-amps are disabled to save power, and the switches S_D short the electrodes to ground in order to prevent current leakage. Notice that since the blocking capacitors still isolate the circuits from the tissue, there is no DC current path even in this case.

The dimension of the final assembled devices is 36mm×20mm×19mm. The 3D construction of the first generation watermaze stimulator is shown in Fig. 6.20. The photo of the assembled device is shown in Fig. 6.21.

The second generation watermaze stimulator is designed on a single PCB board.

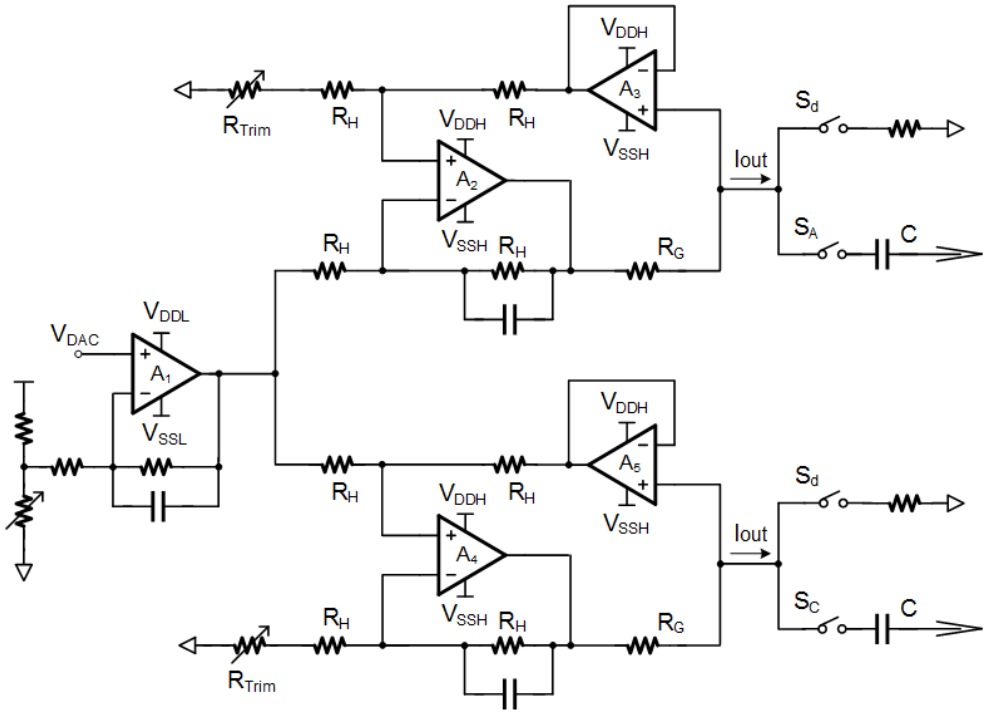


Figure 6.19: The circuit schematic of the first generation watermaze stimulator.

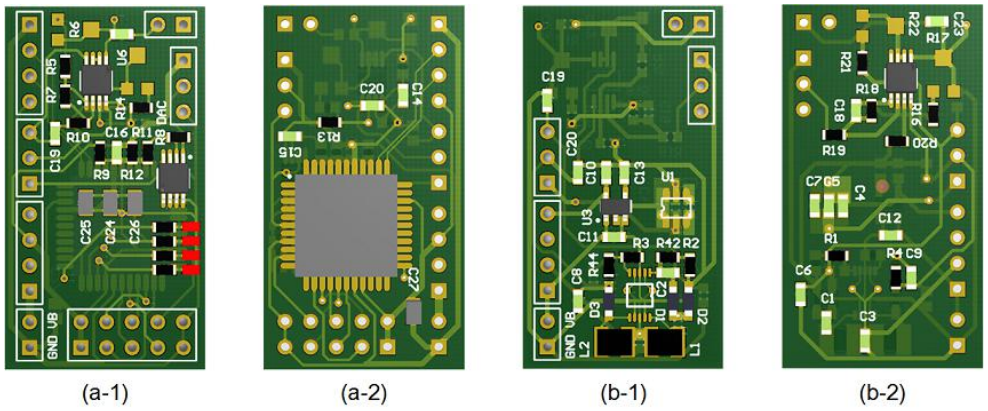


Figure 6.20: The 3D construction of the 1st generation of the watermaze stimulator board. (a-1) and (a-2) are the top boards, and (b-1) and (b-2) are the bottom board.

The circuit schematic is shown in Fig. 6.22. The second generation still uses two dual supplies (VDDH, VSSH, VDDL, VSSL), and a Holland current source. Two DAC channels from the microcontroller are used to set a differential input to the Holland

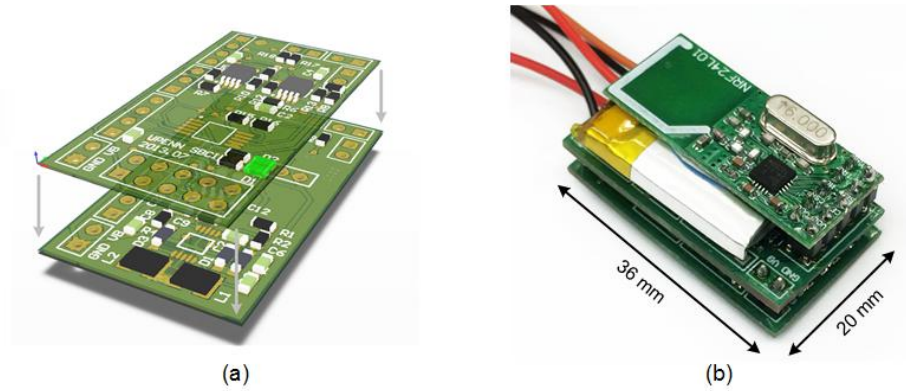


Figure 6.21: (a) The 3D construction of the 1st generation of the watermaze stimulator board. (b) the photo of the assembled stimulator board. Wires are for electrodes and battery charging.

current source, which removes one trimming resistor, and reduces the risk from a resistor drifting caused DC stimulation current. Two channels are designed in this version, using multiplexer ADG409 from Analog Devices. The wireless transceiver nRF24L01, antenna and related matching circuits are also soldered directly on the PCB.

The 3D construction of the second generation watermaze stimulator is shown in Fig. 6.23. The photo of the assembled device is shown in Fig. 6.24. The whole device is coated with PDMS for waterproofing.

The third generation watermaze stimulator is designed with a goal to simply the design and to improve the robustness. A single high supply voltage is used instead of dual supplies, and the stimulation is passed between the bipolar electrodes alternatively for generating Pseudo positive and negative compliance voltages. The circuit schematic is shown in Fig. 6.25. A_1 is a regulating op-amp which is used to produce a high output impedance. The output current is set by V_{DAC}/R_G. The V_{DAC} is programmable so the output current is also programmable. Notice that the

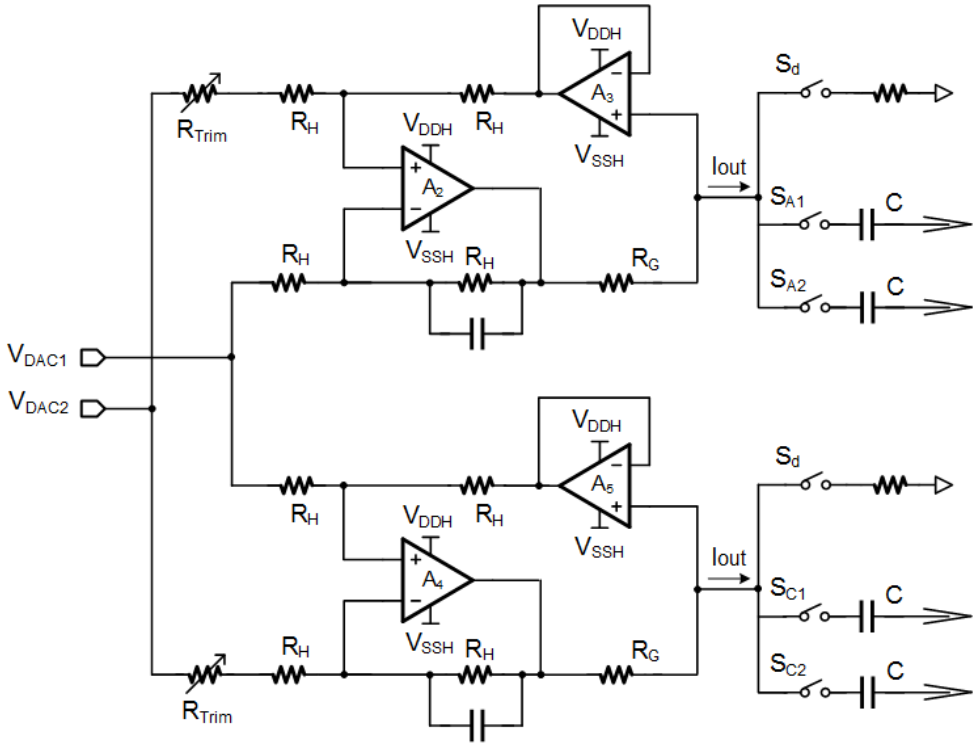


Figure 6.22: The circuit schematic of the 2nd generation watermaze stimulator.

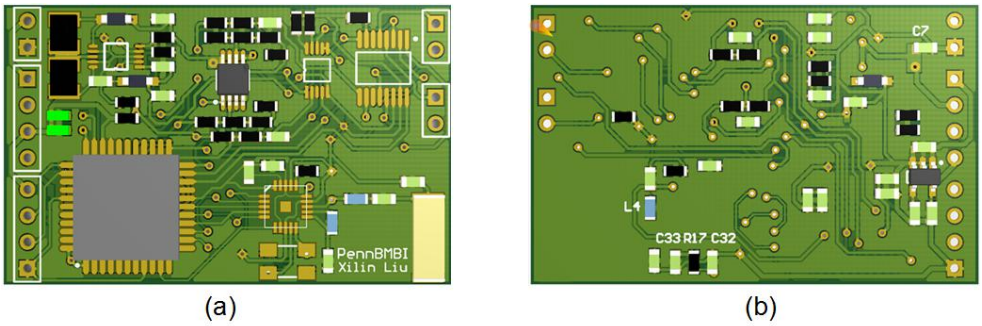


Figure 6.23: The 3D construction of the 2nd generation watermaze stimulator board. (a) is the top view and (b) is the bottom view.

op-amp is powered in the low supply voltage, and the high supply voltage is only used to drive the stimulating electrode. Switches S_3 and S_4 are used to purge the blocking capacitors and discharge the residue charges.

The 3D construction of the third generation watermaze stimulator is shown in

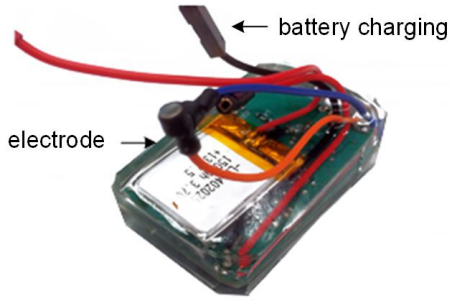


Figure 6.24: The photo of the assembled 2nd generation watermaze stimulator board. Wires are for electrodes and battery charging. The whole device is coated with PDMS for waterproofing.

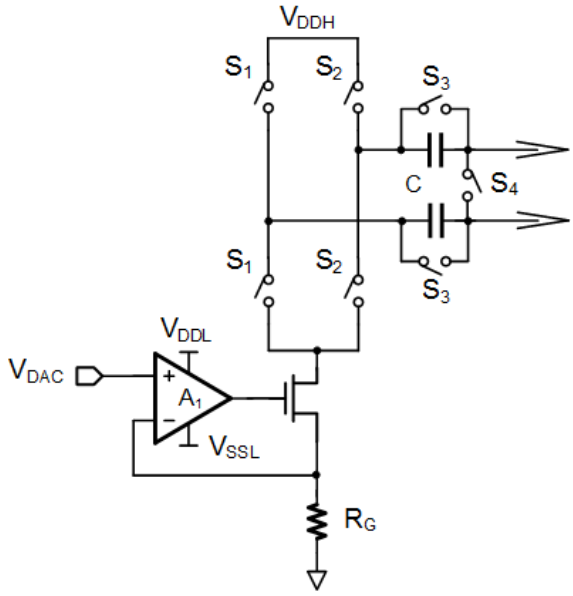


Figure 6.25: The circuit schematic of the 3rd generation watermaze stimulator. The tuning voltage V_{DAC} is generated by the microcontroller.

Fig. 6.26. The photo of the assembled device is shown in Fig. 6.27. The whole device is coated with PDMS for waterproofing.

The electrodes of the stimulator are multiplexed to the ADC. The compliance voltages of the electrodes are measured at the beginning and the end of the stimulation phase, as shown in Fig. 6.28. The spreading resistance can be estimated by

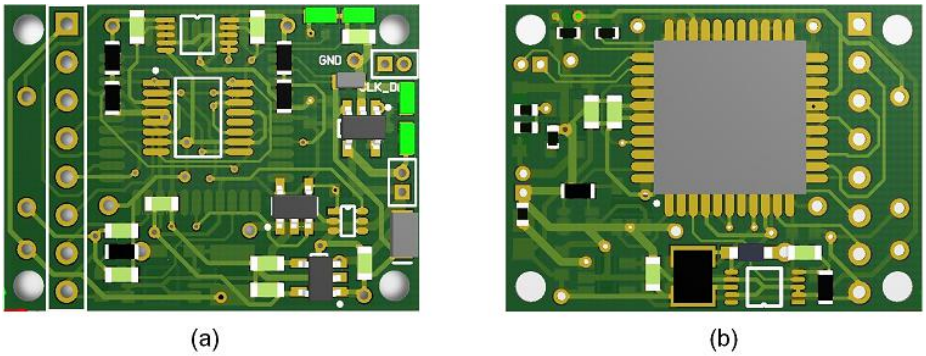


Figure 6.26: The 3D construction of the 3rd generation of the watermaze stimulator board. (a) is the top view and (b) is the bottom view.

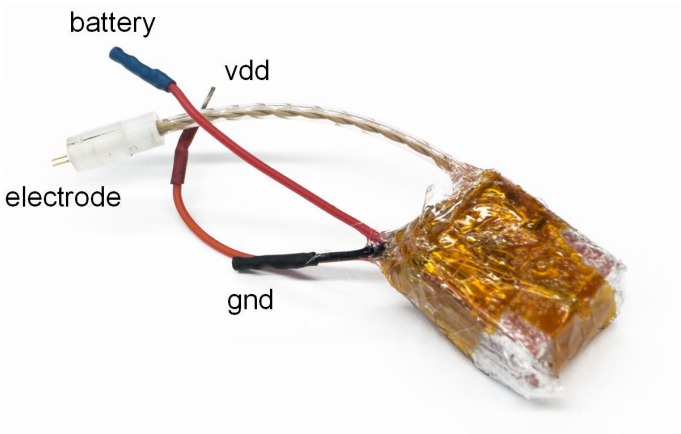


Figure 6.27: The photo of the assembled 3rd generation watermaze stimulator board. Wires are for electrodes and battery charging. The whole device is coated with silicon for waterproofing.

$R_s = V_1/I_S$, where V_1 is the voltage between the two electrodes at the beginning of the stimulating phase, and I_S is the stimulation current. An impedance baseline is measured every time before an experiment. During the experiment, if V_1 is much less than the baseline compliance voltage, it indicates that the equivalent resistance between the electrodes drops significantly, possibly because the electrodes are shorted by water. The experiment should stop since little current is actually passing between

the electrodes. On the other hand, if V_1 is much larger than the baseline, the electrodes may lose connection with the tissue, or a much larger current than the designed value is passing the tissue, possibly because of an electronic failure. The experiment must be halted in both cases to keep the animal safe from tissue damage.

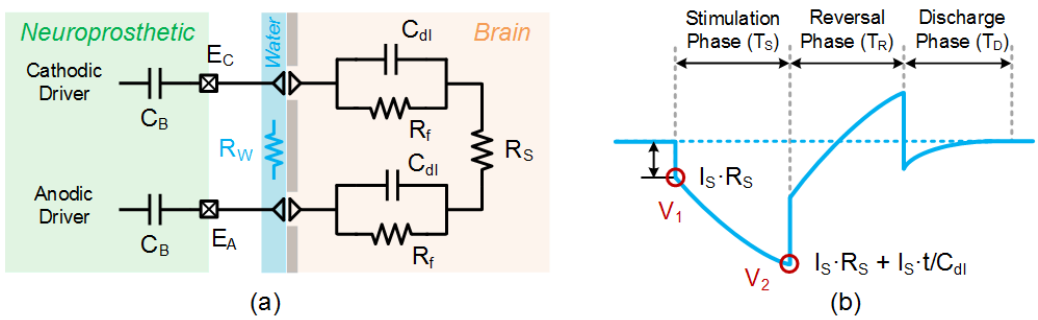


Figure 6.28: (a) The equivalent circuit model for the electrode interface. (b) A typical stimulation waveform between the electrodes E_A and E_C. Compliance voltages at the beginning and end of the stimulation phase are measured for estimating the impedance.

6.3.3.2 Electrode and electrode connector

Electrodes were chronically implanted in the sensory cortex. Different electrodes have been tried in this project, including commercial and custom-made tungsten and stainless steel electrodes. The final selected electrode is a 2-channel commercial electrode with relatively low impedance.

In practice, it takes a considerable amount of practice and time to put the jacket with the device on an awake rat every time before an experiment. The process can be somewhat easier when the rats got used to the jacket, but it can still be time consuming. So a magnetic connector was used in the early stage of this experiment, since it makes the docking process much easier. However, we later found that the magnets cannot provide a secure connection for this experiment, and the water may short the

electrodes through the magnets. The connector of the finally selected electrode has a screw thread to prevent water from shorting the electrodes. The impedance of the electrodes should still be checked before and during the experiment to make sure the animal receives the stimulation without potential danger for brain tissue damage.

6.3.3.3 Image Sensor and Computer Interface

The computer program is compatible with most USB webcams on the market. However, there are two issues worth attention: the viewing angle of the camera and the autofocus and exposure feature. Since the camera is facing the water, many webcams have trouble in focusing, and often give a wrong exposure due to the reflection of light. Different webcams have been tried, including a Microsoft LifeCam VX-5000, a Logitech HD Webcam C310, a Logitech HD Webcam C615. The camera that was decided to be used for this project is the Microsoft Q2F-00013 USB 2.0 LifeCam.

The camera has a wide angle covering the tank area, and the exposure time can be manually set in the Matlab program. The camera also can be securely mounted on the ceiling on top of the water tank. Notice that even a resolution of 1080p is supported, 640×480 is more than sufficient for the tracking purpose in this task. A higher solution will potentially cause a processing delay.

The PC interface from the PennBMBI system was also used in this project. The interface mainly consists of a 16/8-bit XMEGA microcontroller and a 2.4GHz wireless transceiver. The microcontroller has a USB 2.0 module integrated for the communication with the computer. With a full-speed USB, the communication delay can be minimized.

6.3.4 Software Implementation

6.3.4.1 Communication Protocol

The software is implemented using Matlab on a computer, and in C language on the microcontroller. The flowchart of the computer program is shown in Fig. 6.29. The program mainly has three operation modes, i) the testing mode ii) the animal training mode, and iii) the experiment mode.

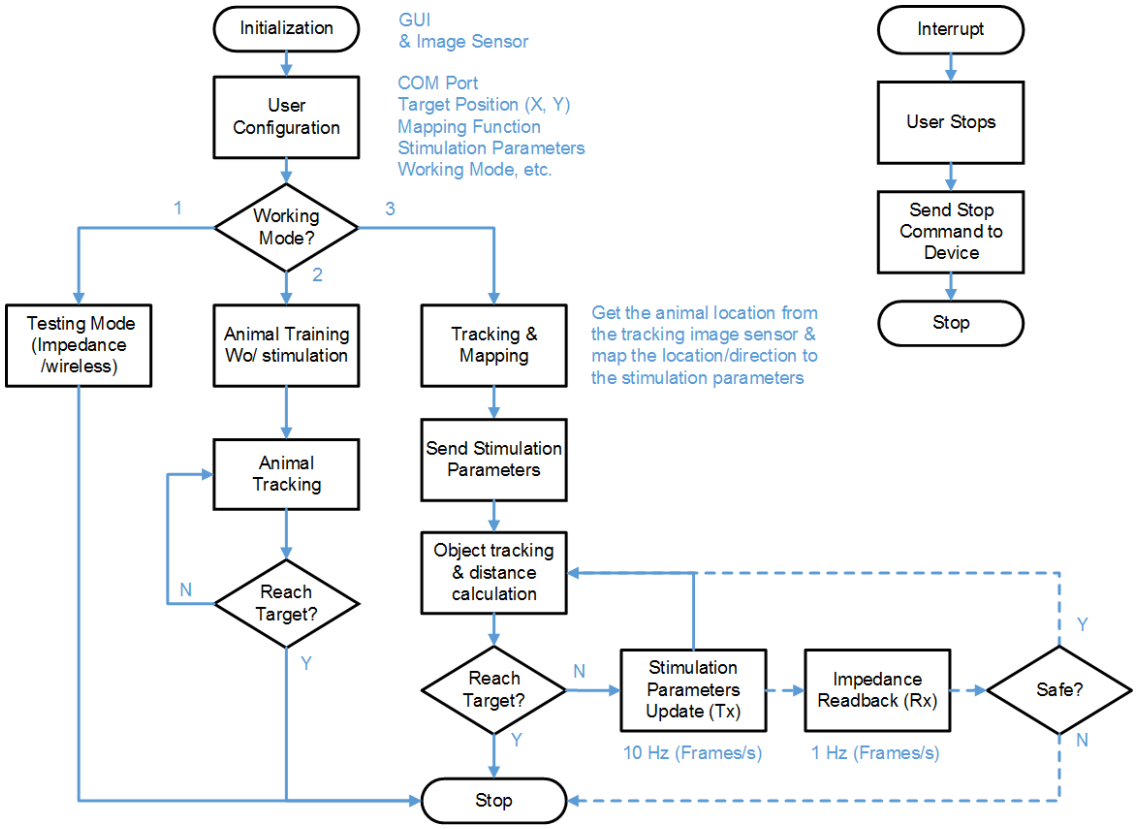


Figure 6.29: The flowchart of the computer program.

The testing mode includes both wireless communication and electrode impedance test. Both tests need to be run every time before the animal is set into the water

for the experiment. In the animal training mode, the image sensor tracks the rat swimming but no stimulation is delivered. The training mode helps the animal get used to swimming in the water tank, and learn the existence of a hidden platform in the tank. After the rat reaches the hidden platform, it will be rewarded to be motivated. Also, the control data is collected in this mode for analysis and comparison purposes. In the experiment mode, the image sensor tracks the rat swimming, and the computer maps the location and/or direction of the rat to a stimulation sequence. The established mapping algorithms include i) binary mapping, ii) linear mapping and iii) Gaussian mapping. In the binary mapping, the rat receives a simulation train only if it's heading towards the hidden platform. In the linear mapping, the simulation frequency is modulated by the distance between the rat's location and the platform in a linear relation, as

$$f_{stim} = \alpha \cdot \sqrt{(x - x_0)^2 + (y - y_0)^2} + \beta \tag{6.1}$$

where x, y are the location of the rat, x_0, y_0 are the location of the platform. α is the gain factor, and β is the offset parameter. Notice that if α is positive, the rat receives a higher frequency stimulation when it swims away from the target. If α is negative, it receives higher stimulation frequency when it swims towards the platform. The offset β should be set so that the stimulation frequency is a positive parameter in the range of 0.5Hz to 300Hz. In the Gaussian mapping, the animal's distance to the platform maps to the stimulation frequency according to the Gaussian distribution, as

$$f_{stim} = f_{max} \cdot e^{-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2}} \tag{6.2}$$

where f_{max} is the designed maximum frequency, and σ is the standard deviation. Versatile mapping functions can be easily implemented in this program.

The computer program updates 10 frames per second, and sends the updated stimulation parameters to the wireless neuroprosthetic. If the animal reaches the hidden platform, the program stops. The user can set the radius of the target. The program checks the load impedance every second to guarantee the safety of the animal. The experiment stops immediately if the measured impedance or compliance voltage is out of the safe range. In addition, an interrupt service allows the user to halt the experiment at any time. The program sends the stop command to stop the stimulation before the computer program ends.

The flowchart of the neuroprosthetic device is shown in Fig. 6.30. The program has a main routine and an interrupt service routine. After powering on, the device performs the initializations. The wireless module will be configured in the receiving mode and then the CPU will be put in the sleep mode to lower the system's power consumption. Once a RF package is received, the device first checks if this is a stop command. Once the stop command is received, the device disconnects the output driver from the electrodes to prevent any potential damage to the animal. Next, the device sends a signal back to the computer indicating the stop command has been executed. The stop command is also used for testing the wireless communication. A wireless communication is established if the computer can successfully read back the response from the device. The computer program retries to establish the wireless handshake ten times before timeout.

If the received package is not a stop command, the device checks the working mode, and performs accordingly. In the impedance testing mode, the device delivers

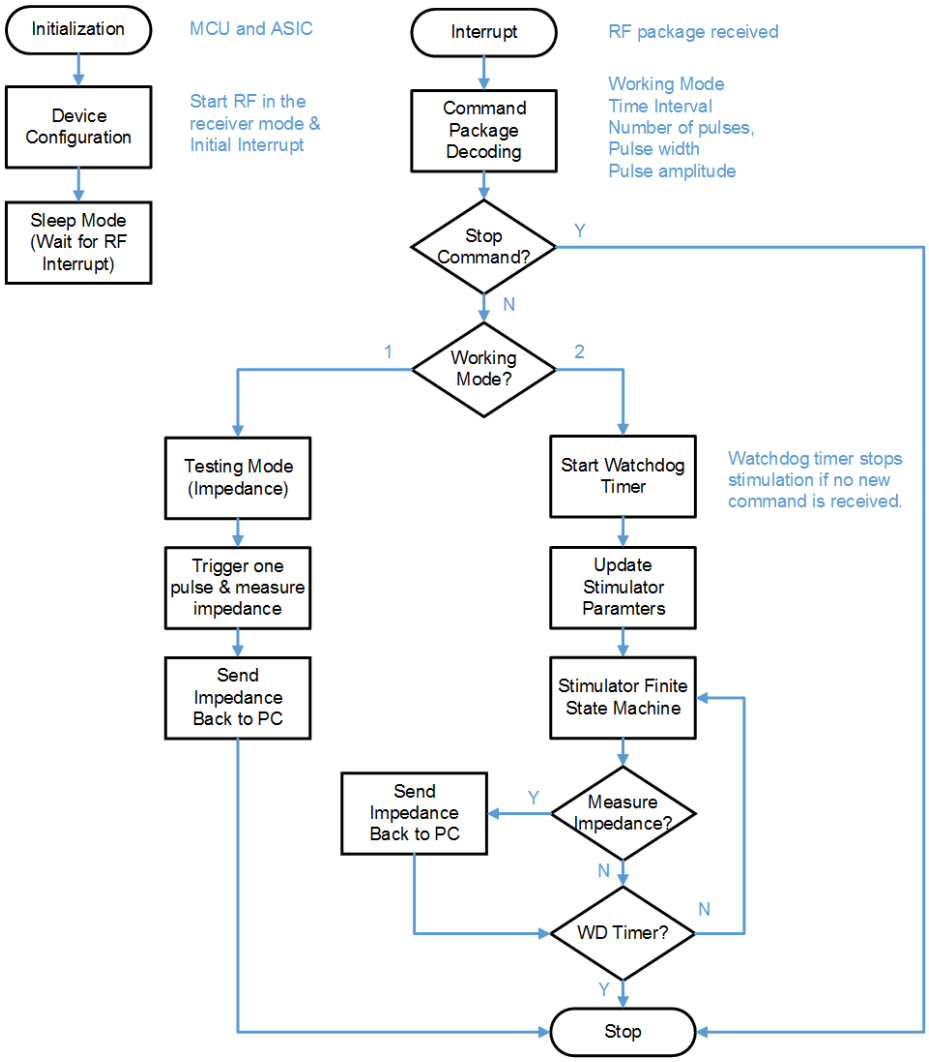


Figure 6.30: The flowchart of the program implemented in the neuroprosthetic device.

one pulse train according to the received parameters. Since the impedance testing mode is a manually triggered simulation mode, it can also be used for studying the animal's reaction to the stimulation out of the water. In the regular experiment mode, a watchdog timer is first started. The timer counts 6 seconds, and if no new RF package is received, the stimulation stops. This is to prevent the failure of wireless communication during an experiment.

The DAC is set according to the stimulation amplitude, and the local finite state machine is set according to the timing parameters. The stimulator runs according to the current time interval until the next command is received. Fig. 6.31 illustrates the timing for the stimulator in two scenarios, when the stimulation time interval is shorter, or longer than the time per frame. The finite state machine will update the real stimulation frequency according to the most recent wirelessly received command.

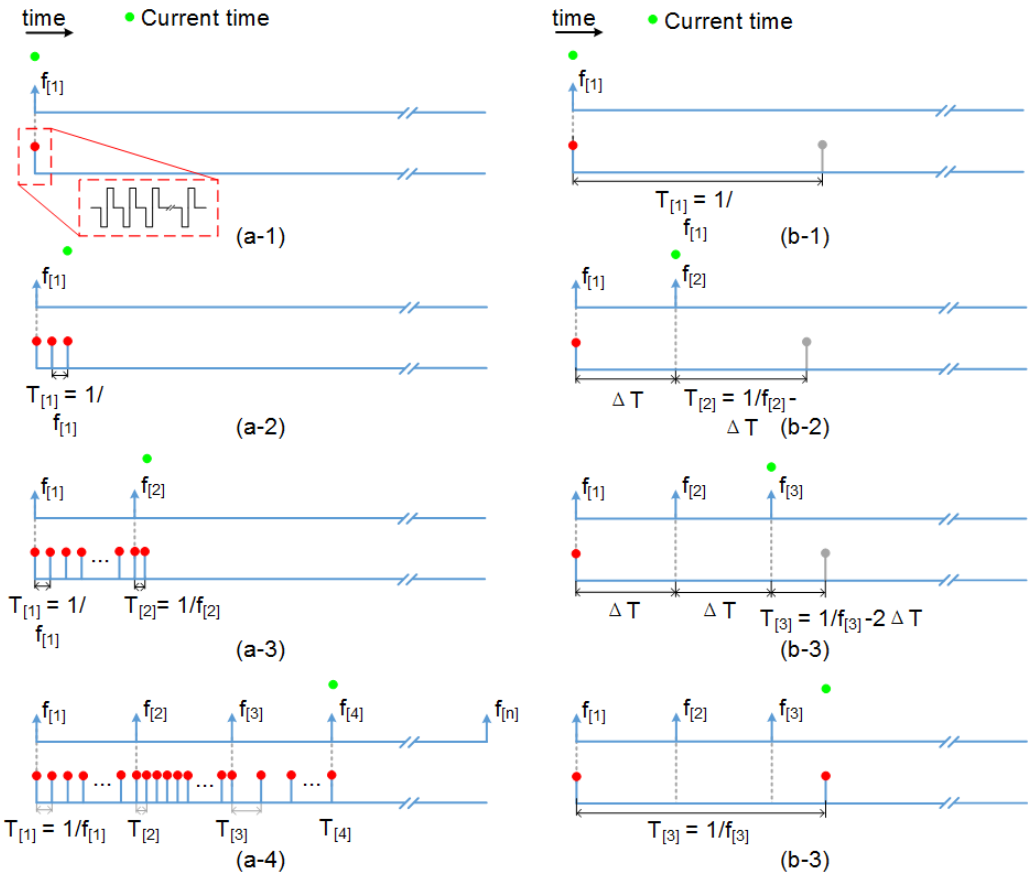


Figure 6.31: Illustration of the timing of the stimulator in the neuroprosthetic. Each red dot is a stimulation pulse train, and the green dot indicates the current time in each subplot. (a) shows the delivered pulses when the stimulation interval is less the time per frame, and (b) shows the delivered pulses when the stimulation interval is larger than the time per frame. The finite state machine will correct the stimulation time interval according to the latest command.

6.3.4.2 Animal Tracking

The animal tracking algorithm was implemented in the Matlab program. After acquiring the image frame from the camera, the program first extracts the red components of the image. Notice that since each pixel consists of R (red), G (green), and B (blue) components, the extracted image has the same dimension as the original image. Then the image is filtered by a median filter to suppress the noise. The image is then converted to binary using a predefined threshold. The threshold can be used to tune how sensitive the algorithm is and should be adjusted according to the environmental light condition. The center of the detected area is then used as the location of the object. The algorithm only considers the object with smaller coordinations when multiple areas are detected. So a drawback of this algorithm is that it won't tell if multiple red objects existed in the scene. This error is avoided by not placing other red objects in the scene. Since most part of the scene is the water tank, the error is not hard to be avoided.

6.3.4.3 User Interface

A friendly user interface has been designed for using the program, as shown in Fig. 6.32. There are mainly 6 panels in the GUI. The communication panel sets the port used to communicate with the PC interface. The used port can be found in the Device Manager in Windows. A Target Position panel where the location and radius can be set. Initially, these positions will display NA. In the target setting mode, a set of random locations and start angles will be generated. But the user can always manually set these parameters.

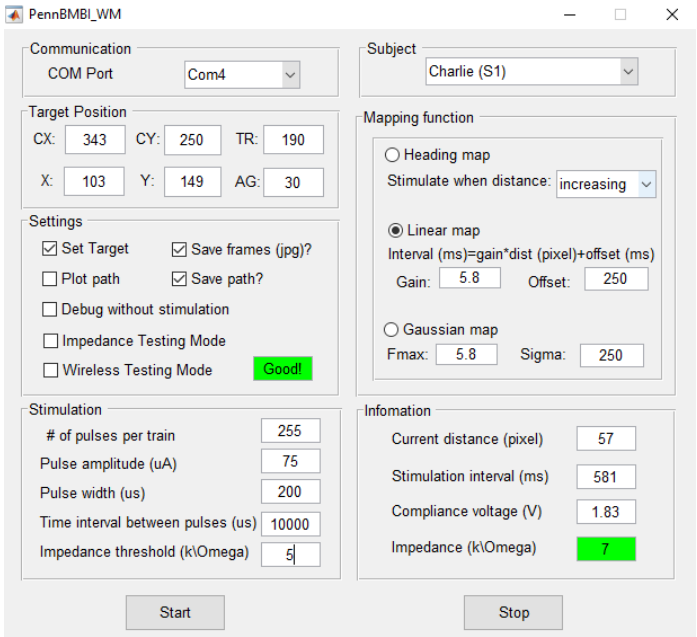


Figure 6.32: Matlab based Graphic User Interface (GUI).

In the settings panel, there are several options for running and control the program. Firstly, there are options for saving the frames, saving the path, and plotting the path. Fig. 6.33 shows a frame of the captured video during the experiment. The large green circle shows the submerged platform, the yellow dot shows the start location. The red curve shows the swimming trace. The left top corner shows the current distance to the platform and the total distance the rat has been traveled in this trial.

There are several working modes one needs to choose before running the experiment. Check the wireless testing mode box, then press the "Start" button, the program will try to communicate with the device. If it successfully reads back from the device, the edit window turns green and shows "Good!". If the communication cannot be established after ten tries, the window turns red and displays "Bad!". Check the impedance testing mode box, then press the "Start" button. The program

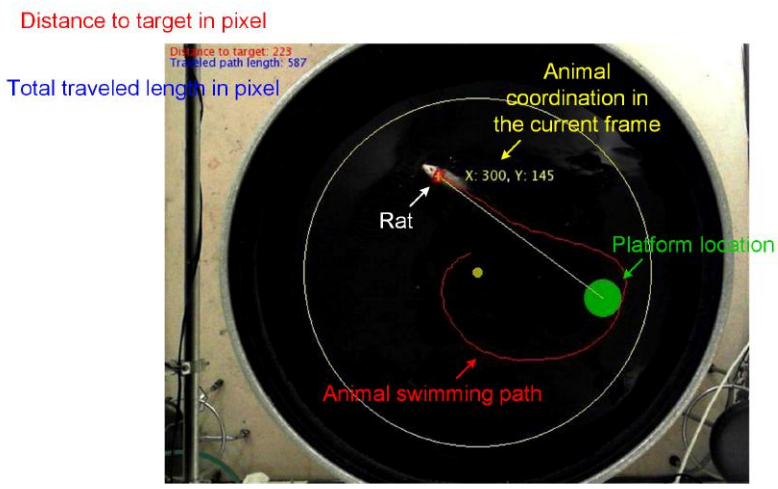


Figure 6.33: One frame of the captured video during the experiment with displayed information highlighted. The large green circle shows the submerged platform, the yellow dot shows the start location. The red curve shows the swimming trace. The left top corner shows the current distance to the platform and the total distance the rat has been traveled in this trial.

will send stimulation commands and read back the compliance voltage for estimating the electrode and load impedance. Check the debug without stimulation mode box, then press the "Start" button, the program will load the camera and track the rat swimming, but no stimulation will be delivered to the rat. This mode can be used for training the animal and getting the control data.

In the simulation panel, several parameters for the stimulation can be set. The parameters include i) number of pulse per train (8 bit), ii) pulse amplitude in μA (8 bit), iii) pulse width in μs (8 bit), iv) time interval between pulses in μs (16 bit), and a threshold for the electrode impedance in kΩ. If the detected impedance is lower than the threshold, the impedance window in the information panel will turn red for a warning. The input number will be truncated to the maximum number of bits allowable for the designed registers in the microcontroller.

In the mapping function panel, the user can select which function to use for mapping the detected rat's location to the stimulation frequency. The panel also includes windows for entering parameters for the mapping functions. The established mapping functions include a binary mapping, a linear mapping, and a Gaussian mapping. More mapping functions can be added in the Matlab program.

The information panel is for displaying the tracking results and compliance voltage readouts in real-time. If the read back impedance is lower than the user-defined threshold, the window turns red. If the compliance voltage cannot be read back, the window will turn yellow and display "999", which is the error code. Otherwise, the window is green in normal operation.

6.3.5 Experimental Results

The experimental results using the developed watermaze device will be presented in this section. A pair of electrodes was implanted in the somatosensory cortex of a Long-Evans rat. Fig. 6.34 shows a rat wearing the wireless waterproof neuroprosthetic device. The jackets used for housing the device are dyed red for the color-based object tracking.

A couple paradigms, modulation algorithms and parameter combinations are studied in this work. Initially, the rats swam in random directions until the platform was found or the trial timed out (60 s). A typical example of the swimming trace before the simulation is shown in Fig. 6.35 (a). The performance significantly improved over the course of about 50 trials, as shown in Fig. 6.35 (b). The poor performance on the catch trials, in which no stimulation was delivered, confirmed that the learned behaviors were guided by the stimulation.

Figure 6.34: A rat wearing the wireless waterproof neuroprosthetic device.

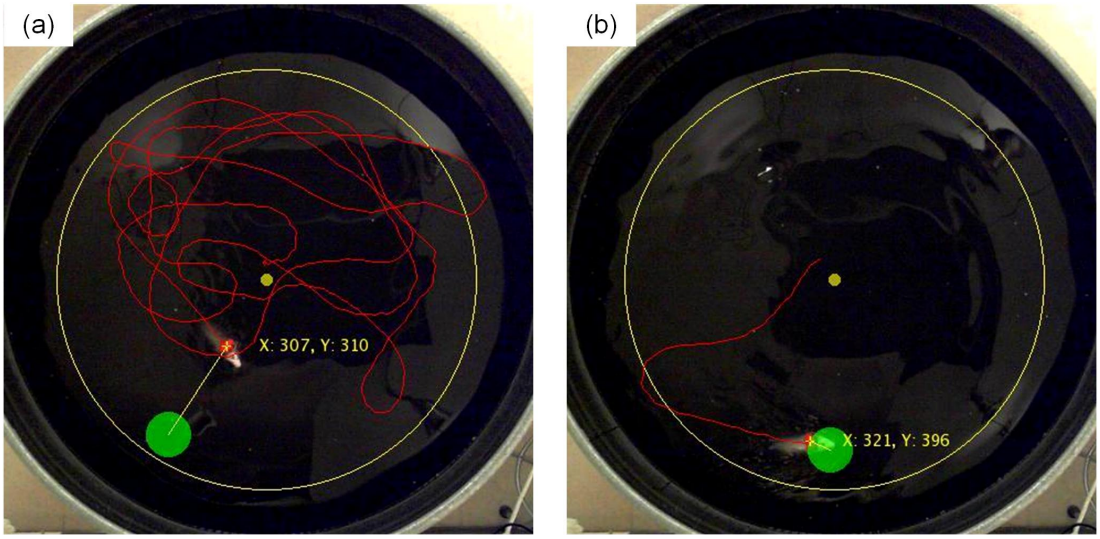


Figure 6.35: The *in-vivo* experimental results. (a) and (b) are webcam captured frames during the experiments. The small yellow and large green circles indicate the start and platform locations, respectively. These were superimposed on the video frame and not visible to the rat. (a) shows the swimming trace of the rat without the simulation, and (b) shows the swimming trace with the simulation guidance.

Fig. 6.36 shows the stimulation pulses versus time, together with the rat's distance to the target platform. Each red vertical line indicates a stimulus pulse. During

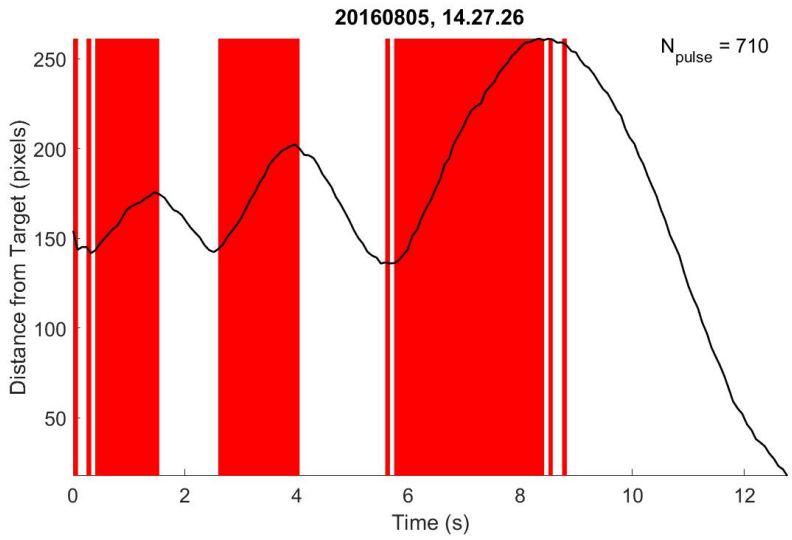


Figure 6.36: A typical trial with stimulation. The animal receives a stimulation when it swims away from the platform. The stimulation pulses are marked by red vertical lines in this figure. It clearly shows the animal turned the direction when it received the stimulation.

this experiment, the rat receives a stimulation when it swims away from the platform. There are in total of 710 stimulus pulses delivered in this 12s trial. The result clearly shows that the rat learned to turn around when it received the stimulation.

In the following analysis, the platform's locations are restricted to one in each quadrant for better quantization. The four locations are separated by 90 degrees with equal distance to the center of the tank, as illustrated in Fig. 6.37. In each trial, the platform is randomly placed in one of the four locations. The rat is initially set free at the center of the water tank, and it had no visual clue of the location of the platform. A total of 139 trials are conducted in this setup, including 124 trials with stimulation and 15 catch trials without stimulation. Naturally, the chance for the rat to visit each of the four locations should be equal, since the platform is randomly placed. Fig. 6.38 compares the percentage of the trials when the rat finds the correct location on the 1st visit to one of the four locations with and without the stimulation.

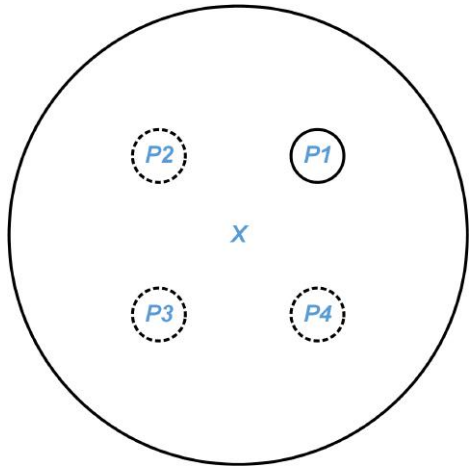


Figure 6.37: Illustration of the four possible platform locations $P1 - 4$. In this experiment setup, the platform was randomly placed in one of them. The rat was initially set free at the center of the water tank in each trial.

Without the stimulation, the percentage for the correct visit is around 20%, while with the stimulation, the percentage is 65%. This result clearly indicates that the rat has learned to use the stimulation.

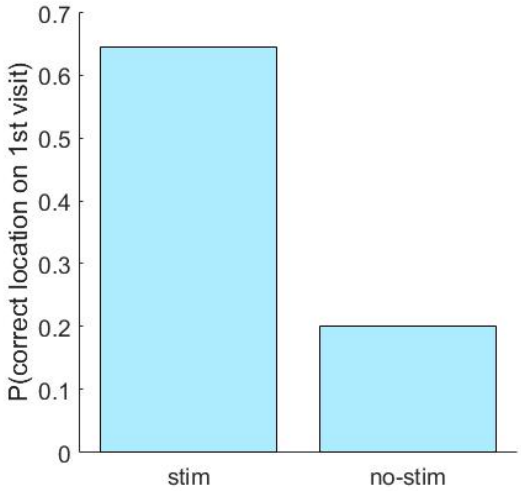


Figure 6.38: In a total of 139 trials consisting of 124 stimulation trials and 15 catch trials, the percentage of trials in which the rat reaches the correct platform in its first visit is 65% with stimulation, and 20% without stimulation.

Fig. 6.39 compares the percentages of trials when the rat's first visited location is the platform's location in the previous trial with and without the stimulation. In

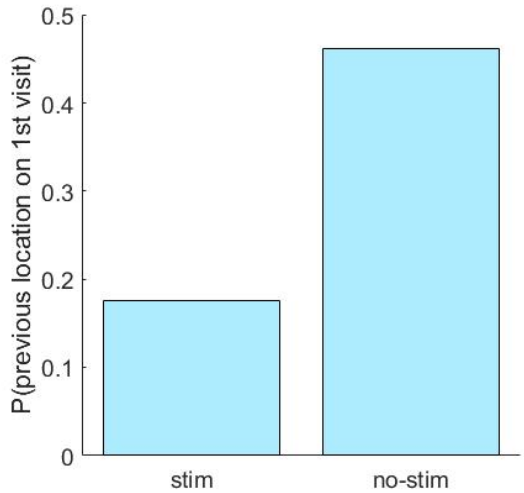


Figure 6.39: The chance for the rat's first visit is the platform location in the previous trial is 17.5% with stimulation, and 46.2% without stimulation.

17.5% of the trials with stimulation, the rat's first visited location is the platform's location in the previous trial, while in 46.2% of the trials without stimulation, the rat's first visited location is the platform's location in the previous trial. The result indicates that the rat relies mainly on its memory to find the location of the platform if no stimulation is presented.

Finally, Fig. 6.40 compares the number of locations the rat visited until it found the actual platform. The result is again the average of the 139 total trials conducted in this setup, including 124 trials with stimulation and 15 catch trials without stimulation. The average times of trials with stimulation are 1.8 times, while trials without stimulation are 3.7 times. The error bars show the standard deviation of the data. The result indicates that for the trials with the stimulation, the rat finds the platform much faster than those trials without the stimulation.

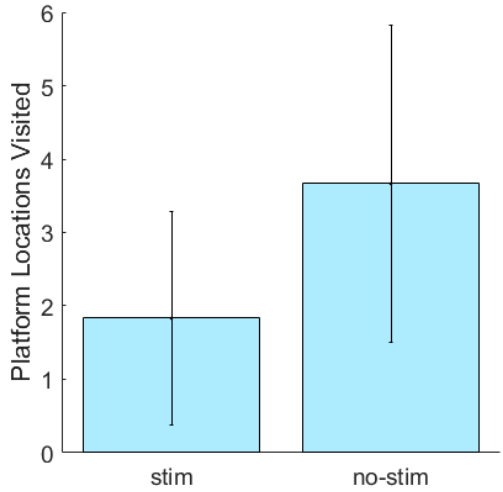


Figure 6.40: The average times of platform visits are 1.8 times with simulation, and 3.7 times without stimulation guidance. The error bars show the standard deviation in the data.

In summary, this section has presented a custom designed wireless BMI platform consisting of a wireless waterproof neuroprosthetic, an animal tracking system and a user interface. The design features a failure prevention mechanism for animal safety. A custom software framework has also been developed to support the experiments. The experiment is the first wireless sensory encoding experiment conducted in a freely swimming animal. The experimental results indicate that animals can quickly interpret artificial percepts to guide behavior. The result is important for the development of sensorimotor neuroprosthetics. More importantly, with the fully programmable wireless interface to the neuroprosthetic, the developed system can be used as a general purpose platform for investigating different sensory encoding experiments in freely behaving animals.

6.4 Bidirectional Neural Interface for Freely Behaving Macaque

6.4.1 Introduction and Background

Sensations and actions are inextricably linked. Behavioral goals are achieved by sampling the environment with the available sensory modalities and modifying actions accordingly. Somatosensory feedback is especially important to the dexterous hand movement control. Recent developments in hand prosthetics with motor pathway replacement alone have shown not to be adequate enough for use of a paralyzed hand [37]. Artificial sensation restoration is needed for this technology to meet the performance required for clinical adoption. The sensation may be restored with direct electrical microstimulation of the brain [38]. Fig. 6.41 illustrates the envisioned bidirectional clinical hand neuroprosthesis with motor function restored through brain-controlled stimulation of hand muscles, and somatosensation restored through sensor controlled electrical stimulation of the brainstem. The cuneate nucleus (CN) in the dorsal brainstem carries fine touch and proprioceptive information from the upper body, and is a suitable sensory encoding site. Besides, its compact representations may be reliably activated artificially. Recently, the Translational Neuromodulation Laboratory (TNL) at the University of Pennsylvania demonstrated the first successful chronic interface to the CN of macaques [302], which allows us to investigate the sensation encoding with CN microstimulation in the monkeys.

In this section, the design and integration of a bidirectional BMI system are presented. The system is custom designed for the operation in freely behaving monkeys.

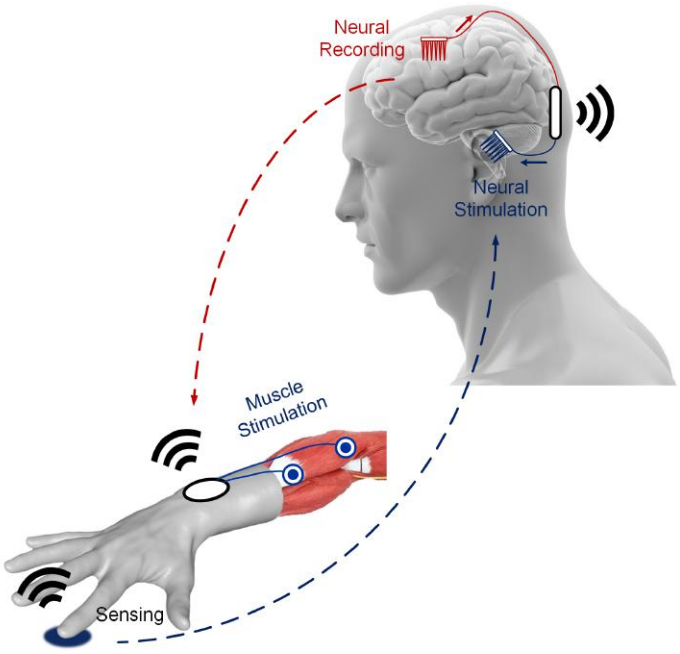


Figure 6.41: Envisioned bidirectional clinical hand neuroprosthesis. Motor function is restored through brain-controlled electrical stimulation of hand muscles, and somatosensation is restored through sensor controlled electrical stimulation of the brain. The motor pathway replacement has been extensively studied, and this work is focused on the sensation restoration.

It has been used to explore the aforementioned sensory mapping, hippocampal oscillation during sleep and awake, and related research. The detailed system architecture, integration method and the animal experimental results are presented.

6.4.2 System Integration

Fig. 6.42 shows an illustration of the custom designed BMI device for freely behaving monkey experiments. The detailed system configuration varies for different experiments. The overall BMI system includes multiple custom ICs and discrete electronic components. The electrodes are chronically implanted, with connectors cemented on the skull to mate with the BMI device. The custom IC performs the noise sensitive

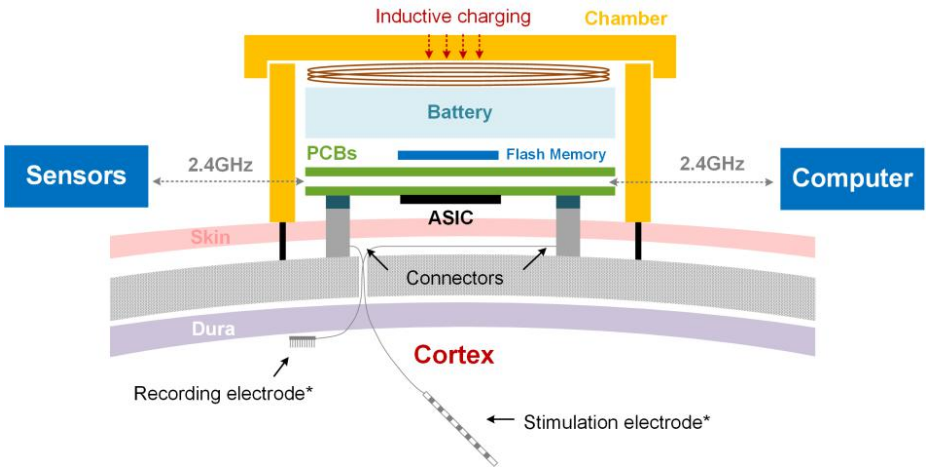


Figure 6.42: Illustration of the BMI device housed in a chamber (not to scale). The electrodes will be chronically implanted with universal nano-connecters through the skin, secured by dental cement. (*electrode type varies with different applications).

neural signal recording, energy-efficient neural feature extraction, and high safety electrical stimulation.

A general-purpose low-power microcontroller (MCU) is integrated into the system for the configuration and control of the ICs, handling of the data packets, and performing certain closed-loop algorithms. Several MCUs have been used in this design, with different performance, hardware interface and features. Table 6.3 compares three of the most commonly used MCUs in this work. The 32-bit Tiva TM4C123GH6PM [303] from Texas Instruments is the mainly used for the monkey project. Compared with the 32-bit AT32UC3C0512C from Atmel, which is previously used to upgrade the PennBMBI system, the AT32UC3C0512C is more power efficient and features more open source libraries. However, the AT32UC3C0512C and several DSP from Analog devices and TI have more powerful signal processing ability, which will be important for heavy duty on-chip neural feature extraction. The ATxmega128A4U [290] from Atmel has been introduced in the previous sections for the PennBMBI system

Table 6.3: Comparison of MCUs used in this work

Features	TM4C123GH6PM	ATxmega128A4U	nRF51822
CPU	32-bit	8/16-bit	32-bit
Voltage	3.3V	1.6-3.6V	1.8-3.6V
Clock	80MHz	32MHz	16MHz
Flash	256kB	128kB	128kB
RAM	32kB	8kB	16kB
EEPROM	2kB	2kB	NA
UART	8x	5x	1x
SPI	4x	2x 12-bit	2x
USB	USB 2.0	USB 2.0	NA
ADC	2x 12-bit	12x 12-bit	8x 10-bit
DAC	NA	2x 12-bit	NA
Wireless	NA	NA	2.4GHz
Package	LQFP 64pin	TQFP 44pin	QFN 48pin

and the watermaze project. Compared with the AT32UC3C0512C, ATxmega128A4U features smaller package, lower power consumption, and lower cost. It also includes a 12-bit digital to analog converter which is useful to generate configuration voltage reference. At the same time, both AT32UC3C0512C and ATxmega128A4U don't have wireless module integrated, which means an additional wireless module has to be integrated on board. So a wireless MCU, nRF51822 is used in a couple of applications with space restraint. The wireless protocol in the nRF51822 is air compatible with the nRF24L01+, which is used in the PennBMBI system and the watermaze project. And the wireless module also features Bluetooth 4.0 protocol for communication with workstations and mobile devices. The nRF51822 has a 32-bit ARM Cortex M0 core, with rich hardware interfaces including UART, SPI, and ADC. However, the processing ability is still much weaker than the AT32UC3C0512C and ATxmega128A4U, so limited on-chip signal processing can be performed, especially when streaming data near the full wireless data rate.

A Micro-SD card module is integrated in the system for wireless recording where there is no need for real-time display or processing. An FAT32 file system is implemented on the Micro-SD card. An open source Generic FAT file system module FatFs [304] is modified for this work. The FatFs is a generic FAT/exFAT file system written in ANSI C and independent of the platform. Additional IO interface layers are written so that both the AT32UC3C0512C and ATxmega128A4U can use the file system. Notice that wiring into the SD card doesn't require a file system. However, writing data without a file system will result in a limited address range of sectors. According to the SD protocol, the maximum data sector reached without a filesystem is 2GB. Moreover, no file system means no direct access from a computer file system. The reading and writing of the SD card would need a custom hardware or software. In this work, the FAT32 file is implemented for easy access, better organization, and future extension. A configuration file can be easily edited and saved on the card. The parameters for the recorder and stimulator can be set in the configuration file, including sampling frequency, buffer size, file length, file name, stimulation pulse width, pulse interval, the number of pulses per group, and the number of groups with different pulse configurations.

The BMI device is powered by 3.7V Lithium batteries. A reliable and high capacity battery is a key component in a wearable device. A lot of different batteries are available on the market. The Polymer Lithium batteries from Adafruit are used in this work. The batteries are lightweight and offer the highest energy density among Lithium batteries on the market. Coin batteries and non-rechargeable batteries have also been used in this project for several extremely small sensor nodes. The power management module on chip includes i) battery protection, charging and management circuit; ii) power management modules including switching converter and LDOs. The

AT32UC3C0512C and ATxmega128A4U are powered by 3.3V, and the nRF51822 is powered by 1.8V. So nRF51822 has the power advantage over the other two. The most power hungry components being the wireless modules (36mW during transmitting) or Micro-SD card (40mW during writing).

Fig. 6.43 (a) shows a photograph of monkey D with one of the custom designed chamber. The chamber has a diameter of 30mm and a height of 40mm. The cap of

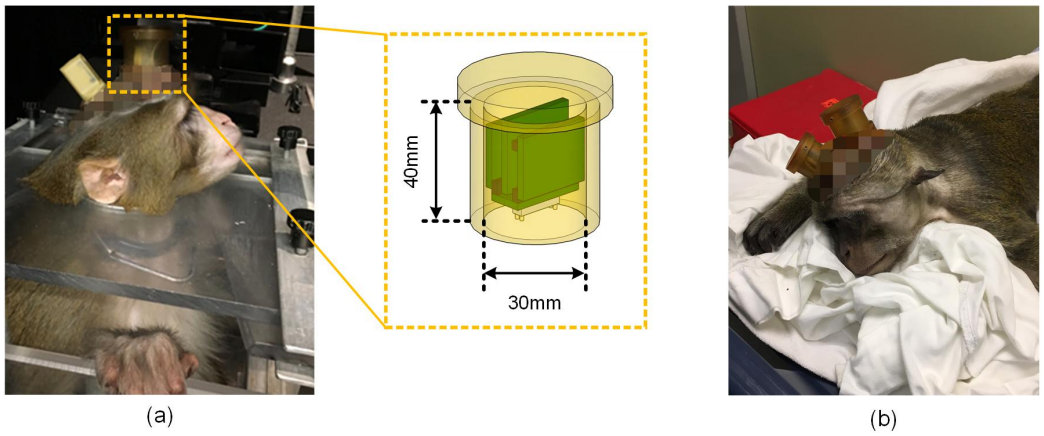


Figure 6.43: The photograph of (a) monkey D with one of the custom designed chamber and (b) monkey M with two of the custom designed chambers. Each chamber has a diameter of 30mm and a height of 40mm.

the chamber is secured by screws. There are three six channel electrodes implanted with connectors in the chamber. Fig. 6.43 (b) shows a photograph of the monkey M with two of the custom designed chambers. There are four 32-channel electrode arrays implanted, with two connectors in each chamber.

A couple of devices have been developed for several different experiments. Fig. 6.44 shows the photographs of several devices. Fig. 6.44 (a) shows a bi-directional BMI with one recording channel and one stimulation channel. The device has an on-board Micro-SD card for data storage. The recording and stimulation modules

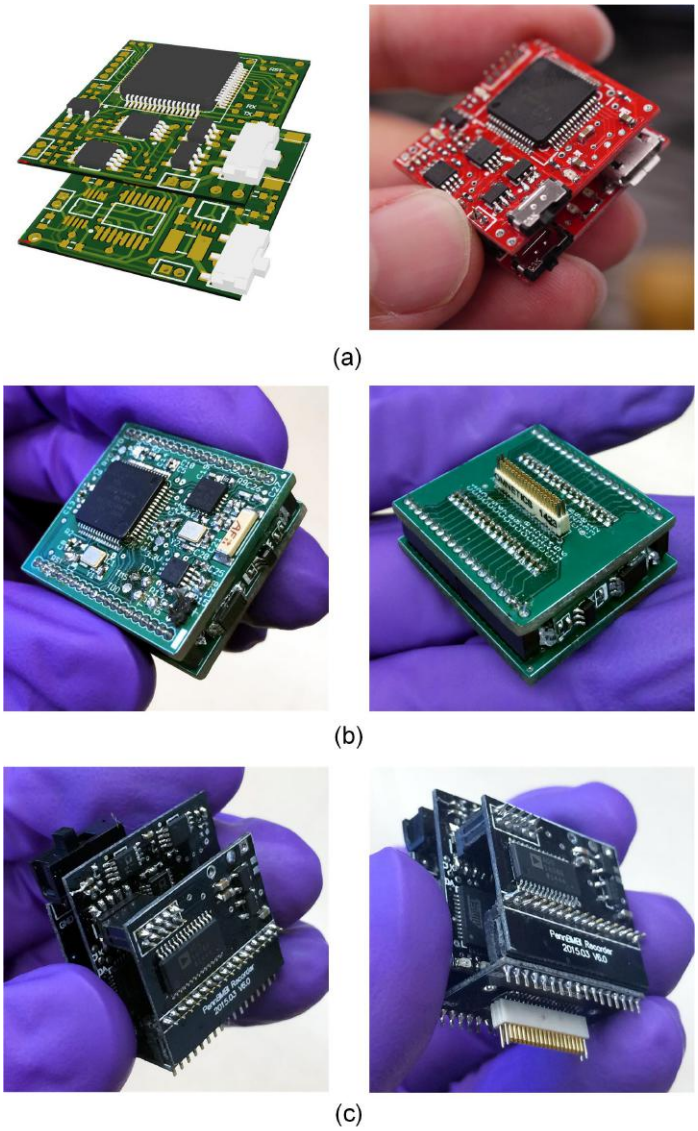


Figure 6.44: Photographs of several assembled devices for the chamber. (a) A bi-directional BMI device with micro-SD card, (b) a 32-channel wireless neural recorder, and (c) a 16-channel wireless bi-directional BMI.

have separated grounds for stimulation artifacts suppression. Fig. 6.44 (b) shows a 32-channel wireless neural recording device. The device features a continuous 12-hour recording with real-time data streaming. The MCU integrated is AT32UC3C0512C. Fig. 6.44 (c) shows a 16-channel wireless bi-directional BMI device. The MCU

integrated is nRF51822.

Custom designed ASIC have been used in these BMI devices. Fig. 6.45 shows the block diagram of the basic version of the bi-directional custom IC. The main building

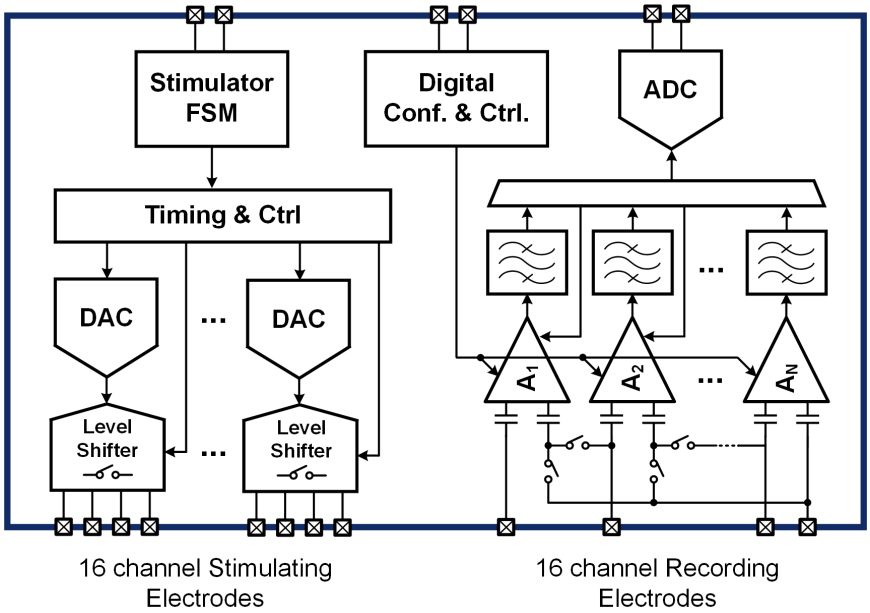


Figure 6.45: The block diagram of the basic version of the bi-directional BMI ASIC.

blocks include i) analog front-end, ii) stimulator back-end, iii) data converters, and iv) peripheral modules.

Fig. 6.46 shows the block diagram of the proposed bi-directional neural interface with energy-efficient neural feature extraction and on-chip PID closed-loop controller.

The SoC mainly consists of 1) 16-channel neural front-end with neural feature extraction units and closed-loop controller, 2) 16-channel programmable neural stimulators, 3) data converters, 4) power management, analog references, and peripheral circuits. The detailed circuit implementation has been presented Chapter 2 to 5. The

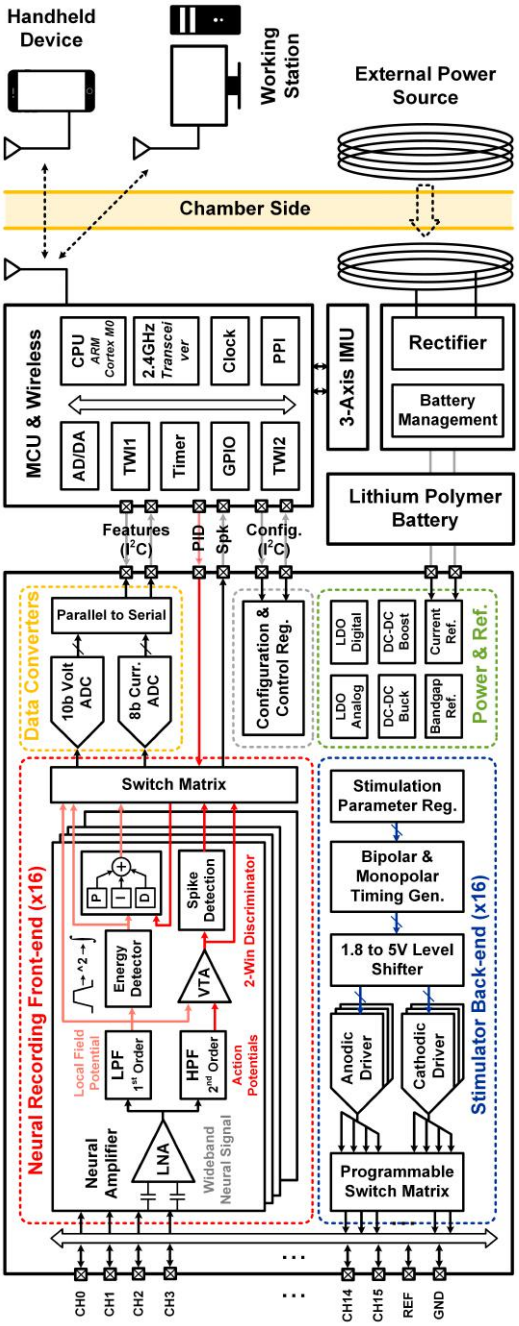


Figure 6.46: Architecture of the bi-directional, closed-loop brain-machine interface system. The system includes a custom system-on-chip (SoC) and supporting electronics.

configuration of the SoC is stored in the flash memory of the MCU, and it can be programmed wirelessly by the Bluetooth link. Once the device is powered up, the MCU first reads the default configuration in the flash memory, and then configures the chip accordingly. The interface between the MCU and the chip is shown in Fig. 6.47. The configuration and data readout are through a simplified two-wire interface (TWI) module. The TWI module supports standard I^2C protocol [305] which is compatible with most general purpose MCU. The MCU works as the master and the chips work as slaves. The MCU first sends the address and the chip with the same address response. Only two pads are used to set up the address, thus, the current implementation can support up to 4 chips (64 channels in total). This can be easily expanded in the future to support 127 chips (full 8-b address). The $START$, $STOP$ and $ANSWER$ commands are also shown in Fig. 6.47.

T3168 and XKT510 are used as the wireless power transmitter and receiver ICs. The wireless charging uses a switching frequency of 125kHz. MC73831 is used for battery management. A 3-axis accelerometer ADXL345 has also been integrated into the system, with 3-wire SPI interface to the MCU.

6.4.3 Experimental Results

The proposed system has been fabricated in standard printed circuit board (PCB) and CMOS technology. The PCB features FR-4 material, 2 or 4 layers, 0.8mm thickness, minimum trace and spacing 0.15mm, minimum hole diameter 0.2mm, minimum via diameter 0.15mm. The PCB surface uses hot air solder leveling (HASL) lead-free finishing.

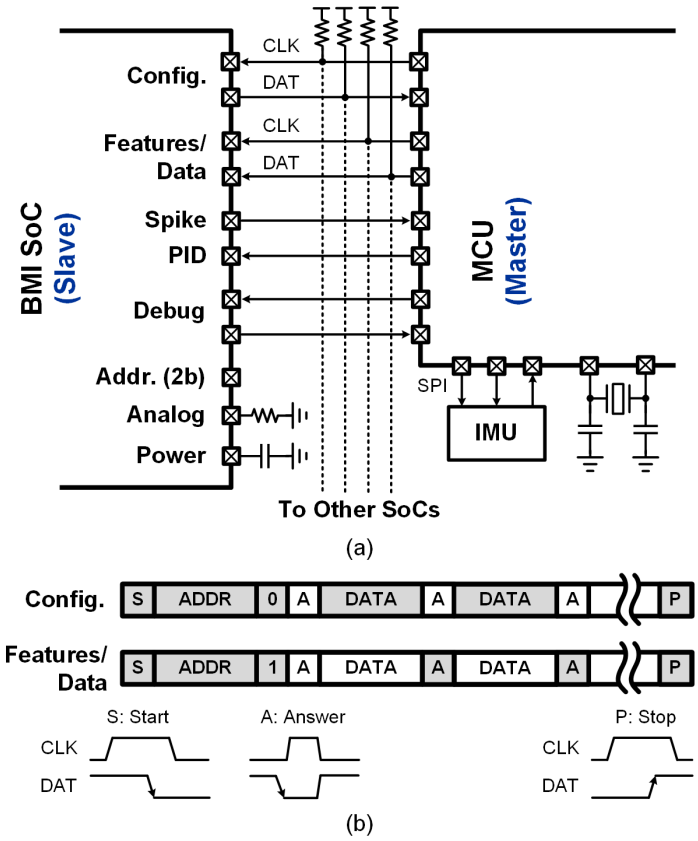
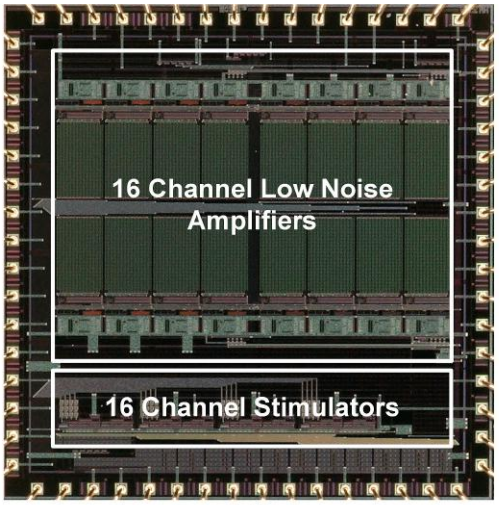
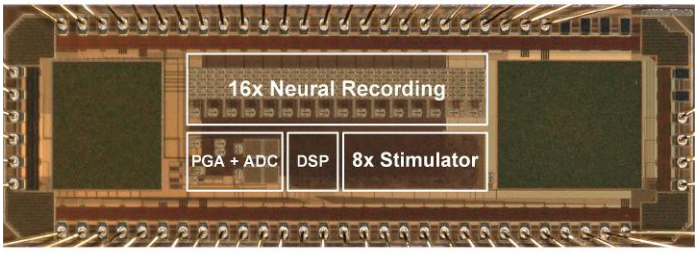


Figure 6.47: (a) Communication interface between the chip and general purpose MCU (not all pads are shown), (b) communication data format. The MCU (master) writes the gray sectors.

The basic version of the neural interface SoC have been fabricated in both On-Semi 0.5μm CMOS technology and IBM 180nm CMOS technology. Fig. 6.48 (a) shows the die photo of the implementation in 0.5μm CMOS. The major building blocks are low noise amplifiers and neural stimulator back-ends. The chip occupies a silicon area of 3mm$\times$3mm including the IO pads. The supply voltage is 3.3 to 5V. Fig. 6.48 (b) shows the die photo of the implementation in IBM 180nm CMOS technology. The chip occupies a silicon area of 4.5mm$\times$1.5mm including the IO pads. The major building blocks are i) analog front-end, ii) stimulator back-end, iii) data converters, and iv) peripheral modules.



(a)



(b)

Figure 6.48: A micrograph of the basic version of the bi-directional neural interface SoCs in (a) On-Semi 0.5μm CMOS technology, and (b) IBM 180nm CMOS technology. Major building blocks are highlighted.

Fig. 6.49 shows the die photos of the proposed bi-directional neural interface with energy-efficient neural feature extraction and PID closed-loop controller. Fig. 6.49 (a) is the first version including 12 channels with debugging and testing structures, and Fig. 6.49 (b) shows the second version including 16 channels. Both chips occupy a silicon area of 4.5mm$\times$1.5mm, including the IO pads. The major building blocks are highlighted in the figure, including: 1) neural front-end with neural feature extraction units and closed-loop controller, 2) programmable neural stimulators, 3) data converters, 4) power management, analog references, and peripheral circuits. Table 6.4 summaries the key specifications of the bi-directional neural interface SoC.

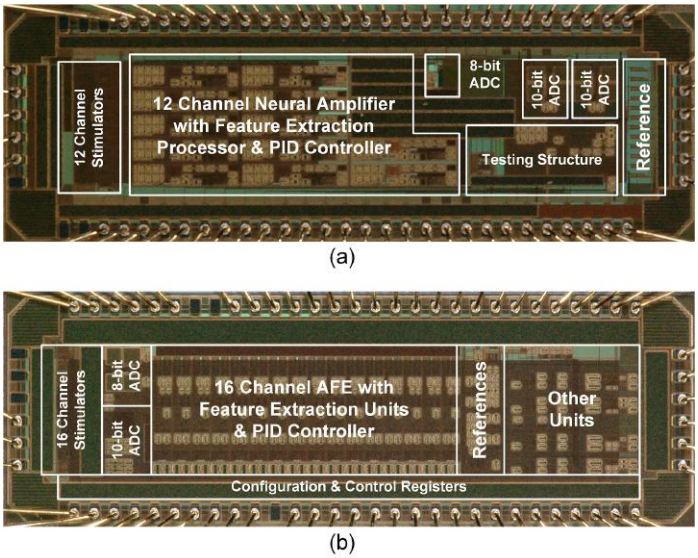


Figure 6.49: The micrographs of the bi-directional neural interface with the proposed energy-efficient neural feature extraction and PID closed-loop controller. (a) shows the first version with 12 channels with debugging and testing structures. (b) shows the second version with 16 channels. Major building blocks are highlighted.

The developed BMI devices have been used in a few experiments with freely behaving monkeys. The experimental results presented in the following analysis were conducted in the monkey (Macaca mulatta) O, D, and F (8-12 kg), with a focus on a study of the hippocampal (HIPP) gamma-slow oscillation coupling in macaques during sedation and sleep. The slow oscillation (SO) of non-rapid eye movement sleep plays a critical role in the consolidation of newly-formed memories [306]. There is substantial behavioral evidence linking the amount of SO activity after learning to the strength of both procedural and declarative memories [307, 308]. These effects of SO on HIPP activity have been studied in rodents and cats, but have not been documented in primates. In this work, we recorded the HIPP field potentials during sedation and during natural sleep. In addition, electrical stimulation was delivered to HIPP afferents in the parahippocampal gyrus (PHG) during sedation and awake to study the effects of the sleep like SO on excitability.

Table 6.4: Key Specifications of the Bi-directional Neural Interface SoC

Analog Front-end	LNA Gain	40dB
	LNA Bandwidth	0.3Hz - 7kHz
	LNA Integral Noise	4.57uV
	LNA Power	9uW
	LNA NEF/PEF	4.77/41.1
	THD (10mVpp Input)	-61dB
	CMRR/PSRR	81dB/71dB
	PGA + Filters Power/ch	8uW
Energy Extraction	Center frequency	1Hz - 200Hz
	Tuning steps	64 natural log
	Quality Factor	1 - 8
	Window length	10 - 500ms
	Ex + PID Power/ch	7uW
Spike Discriminator	Algorithm	Window discrimination
	Amplitude Thresholds	6-bit
	Latency	10us
	Avg. Power/ch	4uW
ADC (Volt Mode /Curr Mode)	Sampling Rate	1MSps/250KSps
	ENOB	9.1/7.9
	FoM(fJ/step)	34.2/10.7
	Power (at 200KSps)	7uW/0.5uW
Stimulator	Stim. Current	4mA/200uA
	Amplitude Res.	6-bit
	Pulse width	1us - 255us
Total Power	Chip Power/ch	56uW/ch
	MCU + Wireless (avg)	8mW

In preparation of the monkey for this experiment, a post was first attached to the skull with screws and acrylic, and magnetic resonance (MR) images of the brain were acquired with fiducial markers. A sterile surgery was performed to implant the electrode arrays using an MR-guided neuronavigation system. Fig. 6.50 illustrates the implanted electrode array. The platinum electrode sites are shown as red boxes and the electrode trajectories are shown as black outlined rectangles. The dimensions of the electrode sites are indicated at the bottom. Typical MRI and CT image

with the visible hippocampal array is shown in Fig. 6.50 (b). Colored regions indicate different neuroanatomical areas. Since the goal of this project is to identify

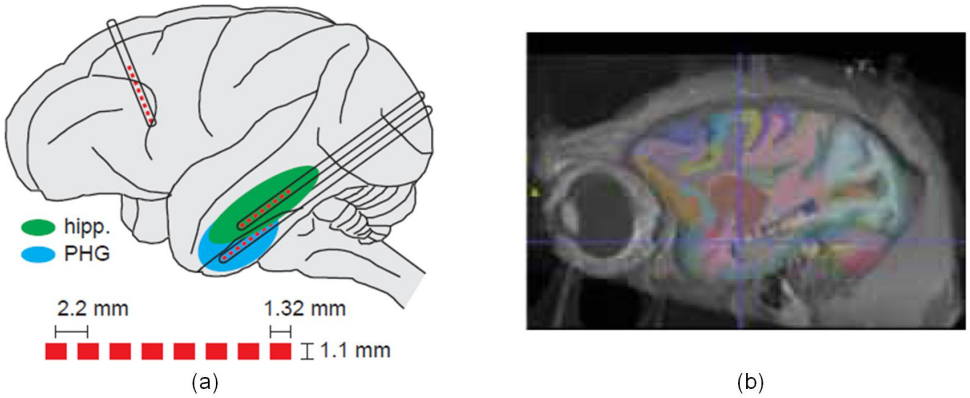


Figure 6.50: (a) Illustration of the implanted depth electrode arrays. (b) MRI and CT image with the visible hippocampal array.

potential closed-loop stimulation paradigms for modulating memory for human patients, clinical microelectrodes were used in this study, rather than the conventional microelectrodes. Single neurons cannot be collected from these electrodes, but the field potentials were sufficient to document the regionally-specific SO and effective PHG-HIPP connectivity for the objective of this study.

Fig. 6.51 shows the power spectrums recorded from the three electrode arrays in the hippocampus, entorhinal cortex and medial septum, respectively. The signal was recorded using the developed BMI device on a Micro-SD card. The device was placed in the chamber during the sedation, and retrieved the next time the monkey was brought to the lab. Different brain states, from sedation, recovery, awake and sleep can be told from the spectrums. Both time and frequency domain features and chewing artifacts verify the reliability of the recording.

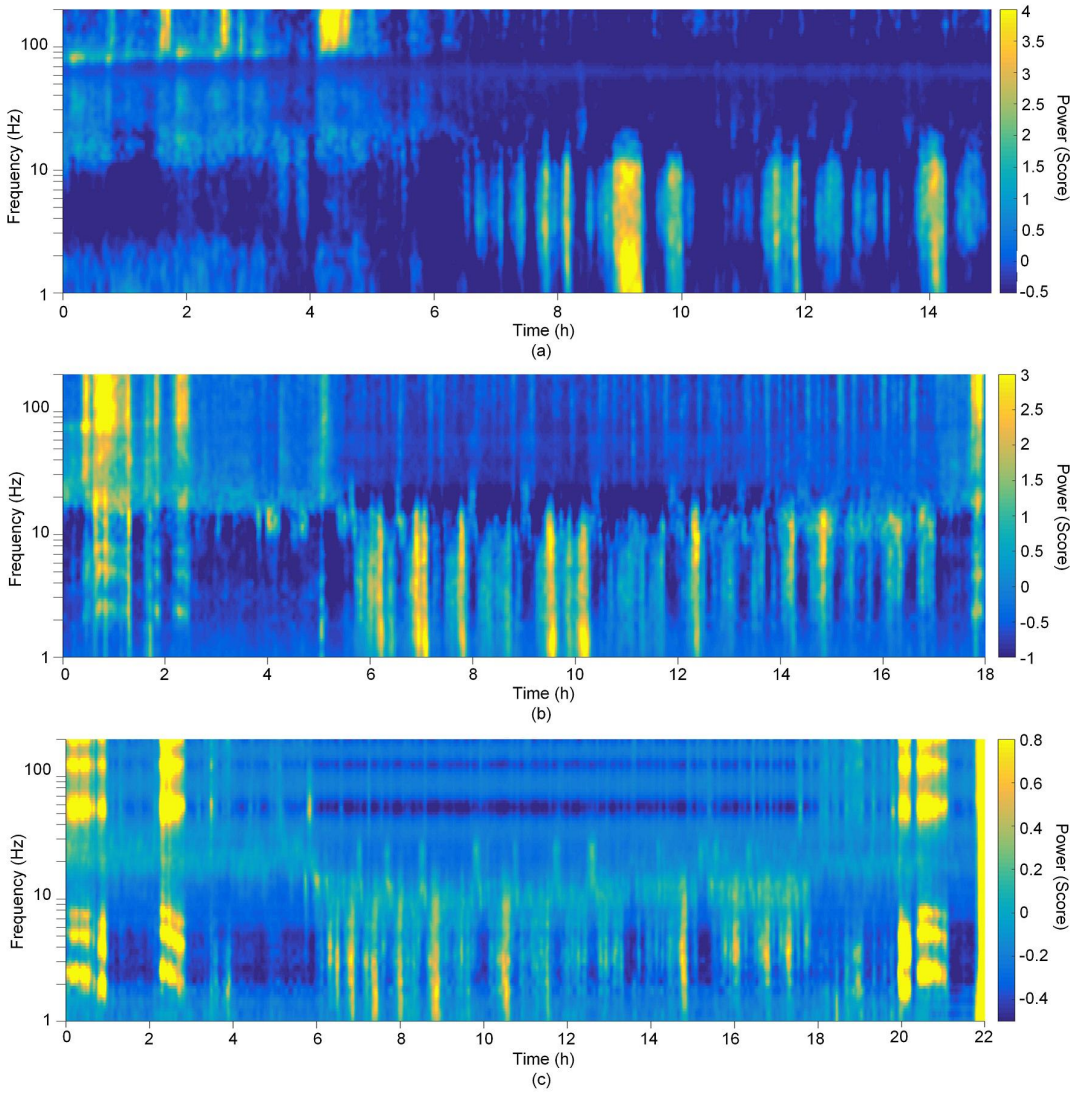


Figure 6.51: The power spectrums of the long-term recordings of the monkey D from the three electrode arrays: (a) hippocampus, (b) entorhinal cortex, and (c) medial septum.

Fig. 6.52 shows the recorded spectrum of the recovery process from anesthesia in monkey D and monkey F, respectively. Both recordings show a period of an increase of the high-frequency oscillation after the sedation, which is the effect of the ketamine-dexmedetomidine.

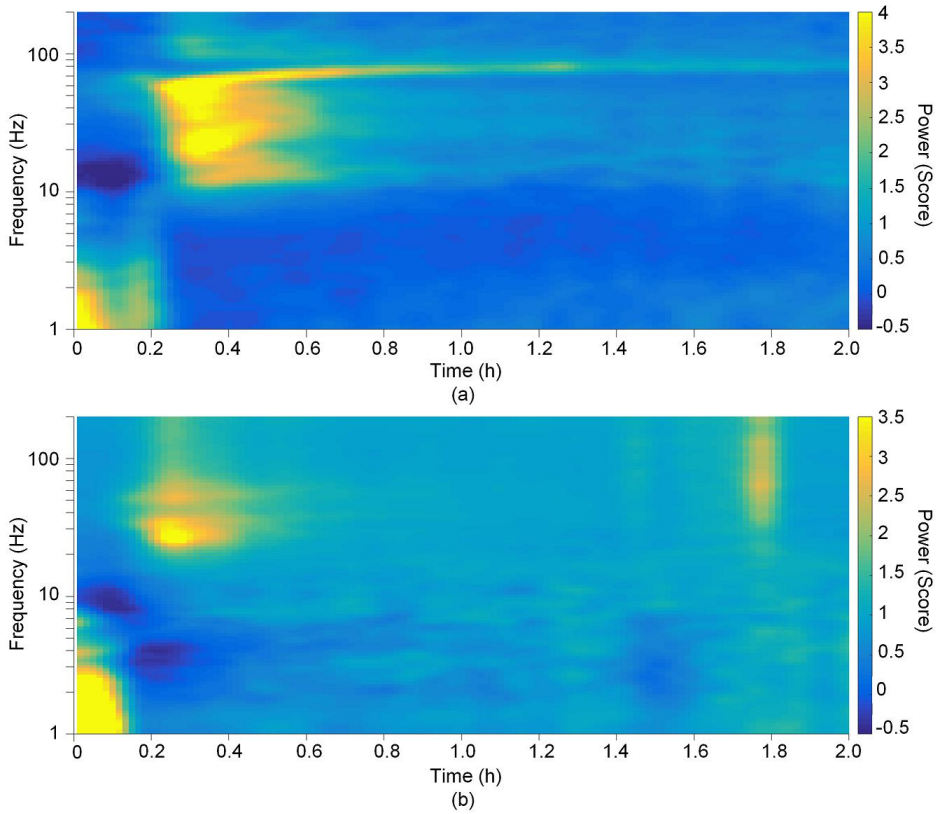


Figure 6.52: The recorded spectrum of the recovery process from anesthesia in (a) monkey D and (b) monkey F.

Separate sessions were conducted with a focus on quantifying neural connectivity. In order to study during stimulation reversal and awake states, the developed bi-directional BMI device was configured to deliver a single bipolar charge-balanced pulse with an amplitude of 2mA in every 30 seconds to the entorhinal cortex. The same device recorded the evoked response in the hippocampus. In this way, the PHG-HIPP connectivity in consistent states defined by oscillatory activity was studied. Fig. 6.53 shows a stacked plot of 278 responses, aligned by the on-set the of the stimulation. The stimulation artifacts are marked by the red arrow. Thanks to the fast artifact recovery design as described in Chapter 5, the evoked potentials can be clearly seen

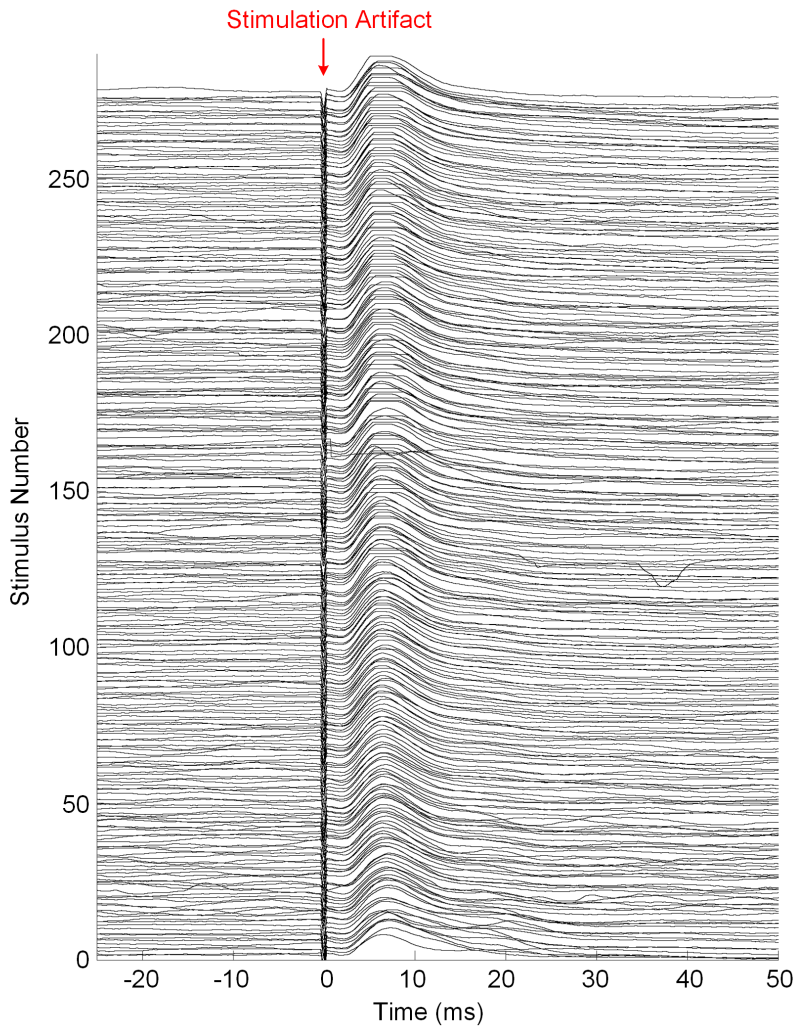


Figure 6.53: The stacked plot of in total of 278 stimulation triggered evoked potentials recorded using the developed BMI device.

from the recording without any signal corruption. The experiment was repeated on Monkey D approximately two months. These results demonstrate the reproducibility and stability of the effects. Fig. 6.54 shows the average waveform of the responses.

In the next experiment, a programmable pulse train was delivered to the medial septum. The stimulation frequency was switched between 20 to 80Hz during one session. Fig. 6.55 compares the evoked potentials from three states: sedation, recovery

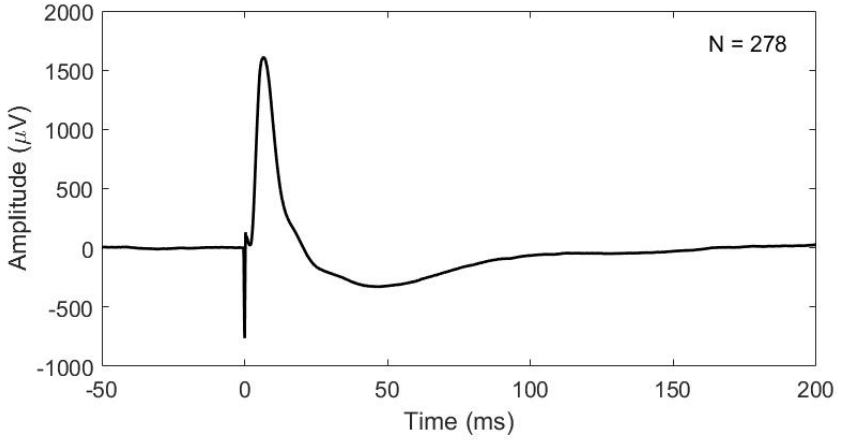


Figure 6.54: The average response waveform of the stimulation triggered evoked potentials in the hippocampus.

and awake. The plots are the average of over a 3-hour recording with stimulation

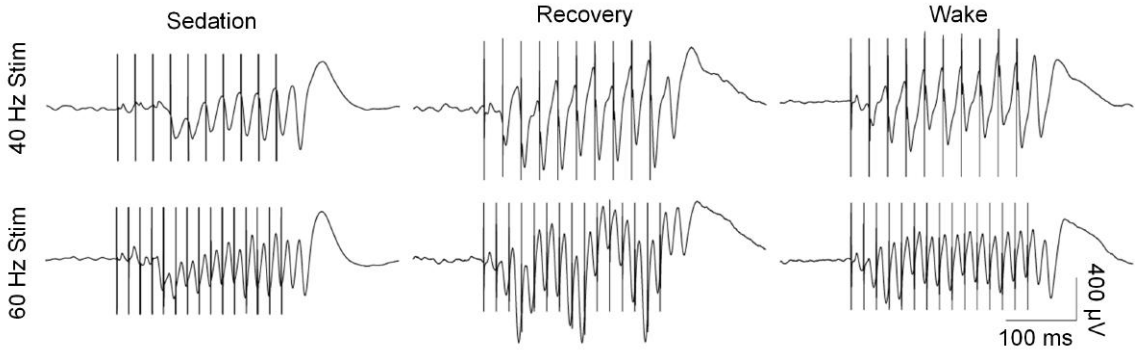


Figure 6.55: Recording of the stimulus-evoked potentials. Stimulus trains of 40 Hz and 60 Hz in different brain states are shown. The plots show an triggered average over 3 hours recording in total, with stimulation every 30 seconds.

delivered every 30 seconds. In all three states, the stimulation pulse train evoked an oscillation in the hippocampus at 40 and 60 Hz. The oscillation continues for at least one cycle after the stimulation pulse train. Due to the high stimulation current (2mA) and the high compliance voltage (12V), the battery typically last 3-4 hours for the stimulation-recording sessions.

Fig. 6.56 illustrates the Gamma-dependence of hippocampal EPs in sedated versus awake animals. Time course of EP peaks and pre-stimulus gamma power in

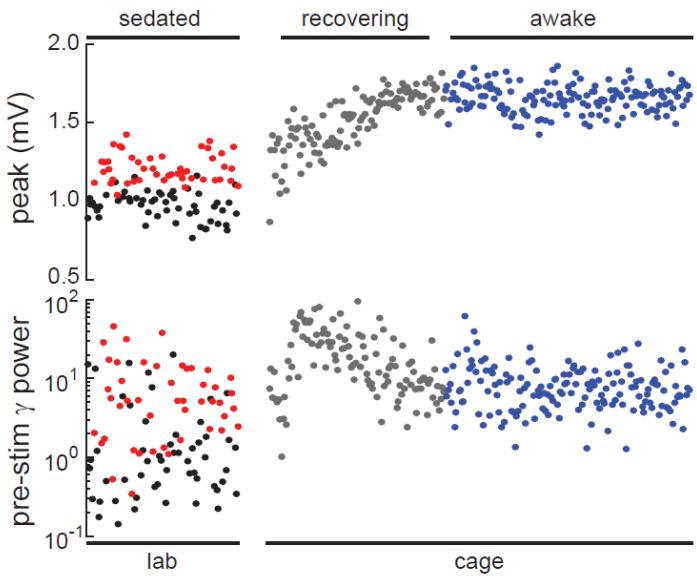


Figure 6.56: Time course of EP peaks and pre-stimulus gamma power in across three behavioral states: sedated (red, black circles), recovery (gray circles), and awake (blue circles). Gamma (30-50 Hz) power was calculated in a 300-ms window preceding each stimulus.

monkey D across three behavioral states: sedated (red, black circles), recovery (gray circles), and awake (blue circles). Gamma (30-50 Hz) power was calculated in a 300-ms window preceding each stimulus. The horizontal axis is scaled to stimulus number, not absolute time. PHG stimuli during sedation in the lab were delivered every 5s by a commercial stimulator. PHG stimuli during recovery and awake periods in the home cage were delivered every 30s by the custom designed BMI stimulator. Stimulus amplitude was 0.5mA in both cases. The reversal agent, atipamezole, was given between the sedated and recovering states. In both monkeys the recovery periods show a transient increase in HIPP gamma-band activity and steady increase in the peak amplitude of the EP.

Thus the gamma-dependence of the evoked HIPP responses were specific to the sedated state and the responses overall were weaker than during the awake state. The results were replicated in Monkey F using instrumentation recording. Fig. 6.57

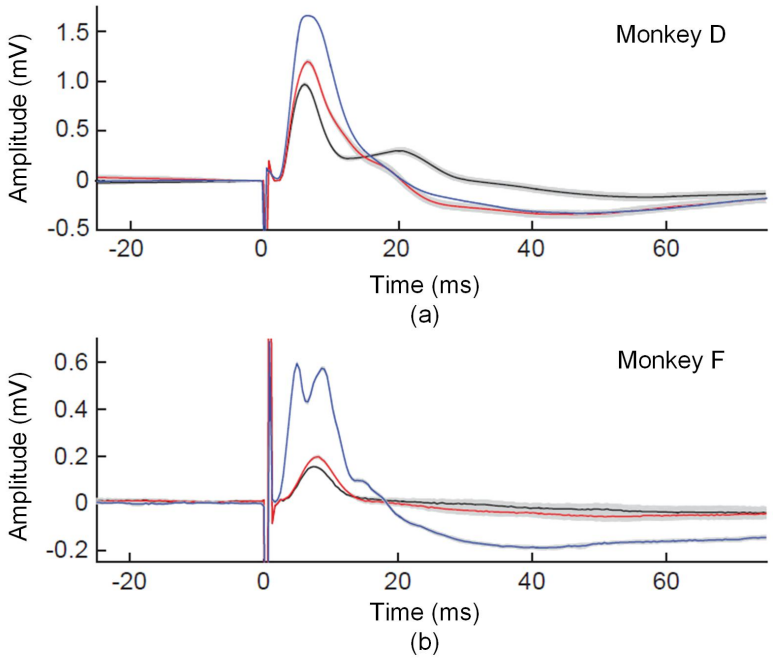


Figure 6.57: Mean EPs during sedation (red, black) and awake (blue) for (a) monkey D and (b) monkey F. 95% confidence intervals on the mean are shown in gray.

compares the EPs during sedation and awake in for Monkey D and Monkey F. 95% confidence intervals on the mean are shown in gray.

Fig. 6.58 (a) shows the power spectral density of the hippocampal recordings during recovery (gray) and awake (blue) states in monkey F. Monkey F exhibited similar HIPP oscillatory activity during recovery from sedation as Monkey D, with an increase in gamma power and a decrease in low-frequency power relative to the awake states. Fig. 6.58 (b) shows the distribution of gamma amplitude across sedated (red, black) and awake (blue) recording sessions in monkey F. The sedated distribution

was colored black and red to highlight its bimodal nature and correspondence to the bimodal EP response amplitudes as in Fig. 6.56 and Fig. 6.57. The higher mode of

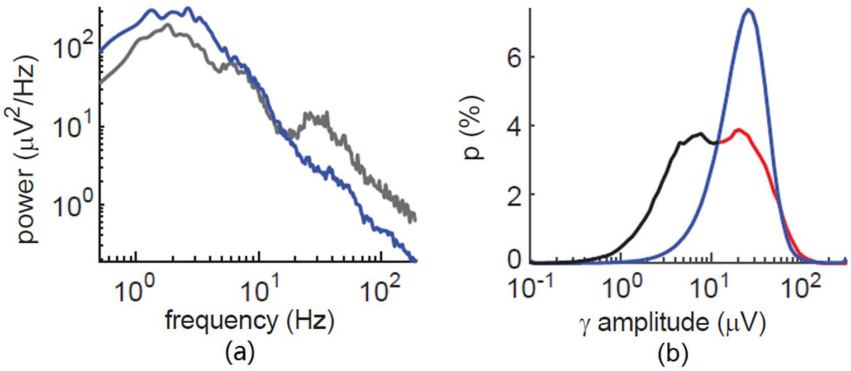


Figure 6.58: (a) The power spectral density of the hippocampal recordings during recovery (gray) and awake (blue) states in monkey F. (b) Distribution of gamma amplitude across sedated (red, black) and awake (blue) recording sessions in monkey F.

the sedation distribution (red) aligned with the awake gamma amplitude distribution (blue), while the lower mode (black) was not presented in the awake distribution.

In summary, the custom developed BMI system has enabled our collaborators from the Perelman School of Medicine, to perform key experiments in freely moving monkeys. The study above shows that the ketamine-dexmedetomidine sedation in primates produces phase-amplitude coupling of gamma and slow oscillations in the PHG-HIPP network. This work presents the first study to directly compare the macaques HIPP field potentials in sleep to sedation. The study suggests that future investigations of the SO in primates would best be conducted during natural slow-wave sleep rather than sedation.

To conclude, a comparison with the recently reported designs of the bidirectional neural interfaces is listed in Table 6.5. Compared with the state-of-the-art design, this work is the first reported wireless bidirectional, closed-loop BMI system used for

long-term freely behaving animal experiments and investigation. This design shows a promising and practical solution for the future development of animal experiments based on primate models.

Table 6.5: Comparison with the State-of-the-Art Bidirectional Neural Interface Designs

Reference	Rhew [11]	Cong [65]	Limnuson [62]	Chen [10]	Shulyzki [69]	Biederman [72]	This work
Publication	2014 JSSC	2014 ESSCIRC	2014 CICC	2014 JSSC	2015 TBioCAS	2015 JSSC	-
CMOS technology	180nm	0.25um&90nm	0.35um	180nm	0.35um	65nm	180nm
ch # of rec./stim.	4/8	32/16	1/1	8/1	256/64	64/8	16/16
AFE noise	6.3uVrms	100nV/rtHz	3.42uVrms	5.23uVrms	7.99uVrms	7.5uVrms	4.57uVrms
AFE NEF	-	-	2.75	1.77	8.9	3.6	4.77
Bandwidth (Hz)	0.64 - 6k	0.5 - 1.7k	0.1 - 12.3k	0.1 - 10k	10 - 5k	10 - 8k	0.3 - 7k
ADC ENOB	5.6	12	10	9.57	-	8.2	9.1
Feature extraction	Custom DSP	Custom CPU	Custom DSP	Custom DSP	-	Custom DSP	Parallel analog
Neural features	Energy	FFT	-	Spectrum, Entropy	-	Spikes	Energy, Spikes
Feedback control	PI	-	-	-	-	-	PID
Stim. mode	Monopolar	Monopolar	Monopolar	Bipolar	Monopolar	Bipolar	Mono/bipolar
Stim. artifact rej.	LP filter	LP filter	Subtraction	-	Blanking	-	Pole-shifting
Ouput current	4410uA	12mA	100uA	30uA	20-250uA	900uA	4mA

Chapter 7

Conclusion and Future Works

7.1 Summary of the Work

This work has presented the analysis and design of a BMI system. To the best of my knowledge, this is the first thesis dedicated to studying the bidirectional closed-loop BMI system. The main motivation of this work is the fact that many significant meaningful neuroscience experiments, especially in freely behaving animals, cannot be conducted without custom designed electronics. With the close collaboration of neuroscientists and engineers, this work was able to identify and address several practical and important issues of the BMI system design. The developed system has been successfully used in several animal experiments often resulting in significant new observations.

The main work and key contributions of this thesis are summarized as follows. In the first chapter, a brief historical review of BMI development was given. A comprehensive survey and review with a focus on the BMIs with the bidirectional neural

interface were performed. Consecutively, an overview of the system architecture of the BMI system was presented. Design considerations and key specifications were summarized. The configurations for the closed-loop operations were illustrated.

In the following three chapters, the design, analysis, and experimental results of the main building blocks of a BMI system, namely the neural signal recording module, the neural feature extraction module, and the neural stimulator module. In Chapter 2, the design of a general purpose low-noise instrumentation amplifier and a low-power ADC were discussed. In addition, a novel pre-whitening neural amplifier was proposed to increase the equivalent dynamic range of the front-end. The pre-whitening processing takes advantage of the neural signal characteristics, significantly relaxing the ADC design without sacrificing the signal quality. The compressive sensing technique was also used to reduce the wireless data rate of the recording front-end.

In Chapter 3, commonly used neural features for closed-loop operations were summarized. The circuit implementation for energy efficient feature extraction of both local field potential and action potential were presented. A natural logarithmic domain neural energy extraction circuit was proposed to provide a sufficient frequency resolution for low-frequency brain oscillations with a minimum number of tuning steps. A low-power action potential discriminator was designed and implemented in current-mode circuits. A matched filter was proposed to extract phase-amplitude coupled neural features. The performance of the matched filter was further improved by employing the proposed pre-whitening filter.

Chapter 4 presented the analysis and design of the electrical neural stimulator. The background and mechanisms of neurostimulation were first reviewed, followed

by a description of physical and electrical models of the electrode and electrolyte interface. Then, an overview of electrical stimulator design was given, including an analysis of methods for stimuli generation, stimulation waveform, and electrode configuration. The methods for achieving the charge balance was also discussed. Next, a general purpose neural stimulator was designed, followed by a novel net-zero charge neural stimulator design. Instead of focusing on circuit matching and residue charge removal, this work attempts to achieve a net-zero charge by employing feedback. The developed chip has been validated in both *in-vitro* and *in-vivo* experiments.

Chapter 5 discussed the design of a bidirectional closed-loop BMI from two important perspectives, the bidirectional neural interface and the on-chip closed-loop control and operation. The stimulation artifact is a known issue in simultaneous neural stimulation and recording. The long lasting stimulation artifact blanks the recording front-end and corrupts the signal. Previous designs proposed different methods to address this problem, however, all have constraints in their applications. In this work, a study of the stimulation artifacts in different recording and stimulator configurations were studied, and conclusions and design recommendations were given. In addition, the mechanisms of the closed-loop operation of the BMI were reviewed and summarized. Followed by the design of a general-purpose programmable PID controller. The implementation of the closed-loop controller is important for both neuroscience research and neuroprosthetic development.

Chapter 6 presented the system integration and animal experiments. The design of a general-purpose experimental platform was first described. Custom communication protocol and user-friendly interface were developed and used extensively in most animal experiments in this work. Then, a watermaze experiment was presented. A

complete experimental system was designed, including a wearable waterproofed stimulator designed to be worn by a rat, an animal location tracking system, and a computer based control interface. Different stimulation parameters and versatile neuromodulation algorithms can be configured in the system. Moreover, the design of a bidirectional neural interface device for the operation in a freely behaving monkey was presented. Long-term experiments of neural stimulation and recording during awake, sedated and sleeping monkeys were given. A study in the hippocampal gamma-slow oscillation coupling using the developed system was also described. The design shows a promising and practical solution for the future experiments using non-human primate models.

7.2 Future Work

A decade ago, Mikhail A. Lebedev and Miguel A.L. Nicolelis predicted that the future BMIs will have sensory feedback directly delivered to the cortical or subcortical somatosensory area, and the closed-loop BMIs would be the ideal tool to restore motor functions [264]. This prediction has inspired this thesis. With the efforts of scientists and engineers around the world, we are stepping into the future of BMI at an incredible pace. Nevertheless, several bottlenecks still need to be overcome.

(1) **Interfacing**: The direct interface between the neuron and electronics are still a challenge, preventing the long-term safe neural stimulation and recording. Novel interfacing material and electronics still need to be developed;

(2) **Wireless Communication**: Although a lot of work has been done in developing wireless neural recorders, a more reliable solution for real-time streaming of multiple-channel signal for recording single neuron activities is still highly desirable. The ideal solution would fully consider the trade-offs between the bandwidth and power consumption, with a minimum data corruption;

(3) **On-chip Processing**: On-chip processing is important for reducing the wireless data rate, and more importantly, to support the real-time closed-loop operation, which is the ultimate goal for the development of most BMI devices. The on-chip operation is usually much more reliable than streaming the data through the a wireless link and rely on an external processing station. However, the limited power budget and on-chip resources place a significant challenge on the on-chip neural signal processing design. With the help of artificial intelligence

and deep learning techniques, the on-chip neural signal processing is one of the most promising research areas in the next few years;

(4) **Power Consumption**: Low-power is always an important design consideration for BMI devices, for both extending battery life and minimizing the tissue damage caused by the generated heat. Developing low-power circuit design techniques as well as exploring energy harvesting opportunities would be the path to overcome this power challenge;

(5) **Packaging**: Biocompatible packaging is critical in the developing of implantable BMI devices. The ideal implantable BMI device would be fully sealed with only wireless interfaces for communication, programming, and battery recharging.

It should be noticed that the aforementioned challenges and opportunities are mainly from the electrical engineering perspective. In addition, the development of fundamental neuroscience and neural engineering innovations have always been the main driving force in the BMI research. So, the biggest opportunity is the close collaboration between the neuroscientists and electrical engineers, as well as scientists and engineers in all related fields. With further improvements in performance, reliability, and range of applications, the BMI technology would benefit a larger and larger population, revolutionize our way of interacting with the external world, and fundamentally help us understand ourselves.

Appendix A

Acronyms

AP	Action Potential
ASIC	Application-Specific Integrated Circuit
BMI	Brain Machine Interface
CMRR	Common Mode Rejection Ratio
CN	Cuneate Nucleus
CR	Compression Ratio
CS	Compressed Sensing
DAC	Digital to Analog Converter
DBS	Deep brain stimulation
DNL	Differential Non-Linearity
ECoG	Electrocorticography
EEG	Electroencephalogram
ENOB	Effective Number of Bits
FES	Functional Electrical Stimulation
FoM	Figure of Merit

INL	Integral Non-Linearity
LFP	Local Field Potential
LNA	Low Noise Amplifier
NEF	Noise Efficiency Factor
NI	Neural Interface
OTA	Operational Transconductance Amplifier
PEF	Power Efficiency Factor
PGA	Programmable Gain Amplifier
PID	Proportional-Integral-Derivative
PSD	Power Spectrum Density
SAR	Successive Approximation Register
SFDR	Spurious-Free Dynamic Range
SNR	Signal to Noise Ratio
SoC	System on Chip
TIA	Transimpedance Amplifier

Appendix B

Propositions

During my five years Ph.D. study and research, I made the following propositions and conclusions, which I think are useful and even defendable. I hope these words would be helpful for the readers from another perspective. This can also be considered as a short version of the thesis without involving the details of electrical and neural engineering.

- When people talk about the brain, it is the brains talk about themselves. Coincidently, we often misunderstand ourselves, and we often have no idea what we are talking about.
- The electrode mismatch causes an amplifier with a perfect CMRR to fail, and the charge diffusion causes a stimulator with a perfect current matching to fail. Despite the complex and elegant circuit design one may have, the biggest challenge is to identify the real problem.

- Our nervous system and electrical circuits rely on negative feedback, while our social relationship rely on the opposite.
- When life gives you a stimulus, don't be afraid and figure out the path. (Learnt from the rat in the watermaze experiment)
- Make a system work may be easy, make it work on a monkey is usually a different story.
- Conventional low-noise amplifier designs optimized for a noise efficiency factor often ignore the fact that the information contained in the signal is what really matters. Techniques like pre-whitening filter and compressive sensing, which significantly reduce the system power while preserving the dynamic range and bandwidth of the signal, are the keys for the next generation neural recorder design.
- When you are doing Ph.D. research, trivial problems will almost always dominate your time. But never, never let trivial problems dominate your mind.

Bibliography

[1] European Committee for Electrotechnical Standardization. Medical electrical equipment - part 2-47: Particular requirements for the safety, including essential performance, of ambulatory electrocardiographic systems. *ICE60601-2-47*, 2015.

[2] F. Yazicioglu. Low-power interface circuits for bio-potential and physiological signal acquisition. *ISSCC Short Course*, Feb 2014.

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

[4] E. Bharucha, H. Sepehrian, and B. Gosselin. A survey of neural front end amplifiers and their requirements toward practical neural interfaces. *J. Low Power Electron. Appl.*, 4:268–291, Appl 2014.

[5] Wikipedia and the Free Encyclopedia. Pid controller. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/PID_controller.

[6] S. Zanos, A. G. Richardson, L. Shupe, F. P. Miles, and E. E. Fetz. The neurochip-2: an autonomous head-fixed computer for recording and stimulating

in freely behaving monkeys. *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, 19(4):427–35, August 2011.

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

[8] S. Stanslaski, P. Afshar, P. Cong, J. Giftakis, P. Stypulkowski, D. Carlson, D. Linde, D. Ullestad, A.-T. Avestruz, and T. Denison. Design and validation of a fully implantable, chronic, closed-loop neuromodulation device with concurrent sensing and stimulation. *IEEE Trans Neural Syst Rehabil Eng*, 20(4): 410–21, Jul 2012.

[9] M. Azin, D.J. Guggenmos, S. Barbay, R.J. Nudo, and P. Mohseni. A battery-powered activity-dependent intracortical microstimulation ic for brain-machine-brain interface. *Journal of Solid State Circuits*, 46(4), April 2011.

[10] Wei-Ming Chen and et al. A fully integrated 8-channel closed-loop neural-prosthetic CMOS soc for real-time epileptic seizure control. *Journal of Solid State Circuits*, 49(1):232–247, Jan 2014.

[11] H. Rhew, J. Jeong, J. Fredenburg, and S. Dodani. A fully self-contained logarithmic closed-loop deep brain stimulation soc with wireless telemetry and wireless power management. *IEEE J. Solid-State Circuits*, 49(10):2213–2227, Oct 2014.

[12] M. Yin, D. Borton, J. Komar, N. Agha, and Y. Lu. Wireless neurosensor for full-spectrum electrophysiology recordings during free behavior. *Neuron*, 84: 1170–1182, Dec 2014.

[13] D. A. Schwarz, M. A. Lebedev, T. L. Hanson, D. F. Dimitrov, G. Lehew, J. Meloy, S. Rajangam, V. Subramanian, P. J. Ifft, Z. Li, A. Ramakrishnan, A. Tate, K. Z. Zhuang, and M. a L. Nicolelis. Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. *Nat. Methods*, 11 (6):670–6, 2014.

[14] E. C. Leuthardt, G. Schalk, D. Ph, J. Roland, and D. W. Moran. Evolution of brain-computer interfaces: going beyond classic motor physiology. *J. Neural Eng.*, 27(1):1–21, 2010.

[15] C. M. Lopez. An implantable 455-active-electrode 52-channel CMOS neural probe. *2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pages 288–289, Feb 2013.

[16] R. R. Harrison and C. Charles. A low-power low-noise CMOS amplifier for neural recording applications. *IEEE J. Solid-State Circuits*, 38(6):958–965, 2003.

[17] C. Toumazou. Novel current-mode instrumentation amplifier. *IEEE Electronics Letter*, 25(3), 1989.

[18] C. Liu, S. Chang, G. Huang, and Y. Lin. A 10-bit 50-ms/s SAR ADC with a monotonic capacitor switching procedure. *IEEE J. Solid-State Circuits*, 45(4): 731–740, 2010.

[19] I. M. Filanovsky and H. Baltes. CMOS schmitt trigger design. *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.*, 41(1):46–49, 1994.

[20] D. R. Merrill, M. Bikson, and J. G. R. Jefferys. Electrical stimulation of excitable tissue: design of efficacious and safe protocols. *Journal of Neuroscience Methods*, 141(2):171–198, Feb 2005.

[21] M. Ghovanloo and et al. A compact large voltage-compliance high output-impedance programmable current source for implantable microstimulators. *IEEE Tranactions on Biomedical Engineering*, Jan 2005.

[22] X. Liu, A. Demosthenous, and N. Donaldson. An integrated implantable stimulator that is fail-safe without off-chip blocking-capacitors. *IEEE Tran. on Biomeidcal Circuits and Systems*, 2(3):231–244, 2008.

[23] Neuron. Brain-computer interface. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Neuron.

[24] Wikipedia and the Free Encyclopedia. Star. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Star.

[25] E. Adrian and B. Matthews. The interpretation of potential waves in the cortex. *Journal of Physiology*, 81:440–471, 1934.

[26] Wikipedia and the Free Encyclopedia. Brain-computer interface. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Brain-computer_interface.

[27] J. Vidal. Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering*, 2(1):157–80, 1973.

[28] P. R. Kennedy. The cone electrode: a long-term electrode that records from neurites grown onto its recording surface. *Journal of Neurosciencce Methods*, 29:181–193, 1989.

[29] S. Bozinovski, M. Sestakov, and L. Bozinovska. Using eeg alpha rhythm to control a mobile robot. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 3:1515–1516, 1988.

[30] Garrett B. Stanley, Fei Li, and Yang Dan. Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *Journal of Neuroscience*, 19(18):8036–8042, 1999.

[31] John K. Chapin and et al. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nature neuroscience*, 2(7):664–670, 1999.

[32] Wessberg, Johan, and et al. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, 408(6810):361–365, 2000.

[33] Schalk, Gerwin, and et al. Bci2000: a general-purpose brain-computer interface (bci) system. *IEEE Transactions on biomedical engineering*, 51(6):1034–1043, 2004.

[34] Leigh R. Hochberg and et al. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398):372–375, 2012.

[35] Sharlene N. Flesher and et al. Intracortical microstimulation of human somatosensory cortex. *Science Translational Medicine*, 8(361), 2016.

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

[37] J. C. Rothwell, M. M. Traub, B. L. Day, J. A. Obeso, P. K. Thomas, and C. D. Marsden. Manual motor performance in a deafferented man. *Brain*, 105(3): 515–42, Sep 1982.

[38] S. J. Bensmaia and L. E. Miller. Restoring sensorimotor function through intracortical interfaces: progress and looming challenges. *Nat Rev Neurosci*, 15: 313–25, May 2014.

[39] Wikipedia and the Free Encyclopedia. Control theory. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Control_theory.

[40] Hammond C.; Ammari R; Bioulac B; Garcia L. Latest view on the mechanism of action of deep brain stimulation. *Mov Disord.*, 23(15):2111–21, 2008.

[41] B. Rosin, M. Slovik, R. Mitelman, M. Rivlin-Etzion, S. N. Haber, Z. Israel, E. Vaadia, and H. Bergman. Closed-loop deep brain stimulation is superior in ameliorating parkinsonism. *Neuron*, 72(2):370–384, 2011.

[42] A. Berenyi and et al. Closed-loop control of epilepsy by transcranial electrical stimulation. *Science*, 337(6095):735–737, 2012.

[43] T. Jeanne and et al. Closed-loop optogenetic control of thalamus as a tool for interrupting seizures after cortical injury. *Nature neuroscience*, 16(1):64–70, 2013.

[44] U. Rutishauser, A. Kotowicz, and G. Laurent. A method for closed-loop p-resentation of sensory stimuli conditional on the internal brain-state of awake animals. *J. Neurosci. Methods*, 215(1):139–55, Apr. 2013.

[45] Eberhard E. Fetz. Restoring motor function with bidirectional neural interfaces. *Progress in Brain Research*, 218:241–252, 2015.

[46] C. T. Moritz, S. I. Perlmutter, and E. E. Fetz. Direct control of paralysed muscles by cortical neurons. *Nature*, 456:639–42, Dec 2008.

[47] Timothy H. Lucas and Eberhard E. Fetz. Myo-cortical crossed feedback reorganizes primate motor cortex output. *Journal of Neuroscience*, 33(12):5261–5274, Mar. 2013.

[48] Ryan W. Eaton, Tyler Libey, and Eberhard E. Fetz. Operant conditioning of neural activity in freely behaving monkeys with intracranial reinforcement. *Journal of Neurophysiology*, 117(3):1112–1125, Mar. 2017.

[49] Scopus. Retrieved Jan 2017. URL https://www.scopus.com.

[50] J. Mavoor and et al. An autonomous implantable computer for neural recording and stimulation in unrestrained primates. *Journal of Neuroscience Methods*, 148 (1):71–77, Oct 2005.

[51] F. Heer, S. Hafizovic, and W. Franks. CMOS microelectrode array for bidirectional interaction with neuronal networks. *IEEE J. Solid-State Circuits*, 41(7): 1620–1629, Jul. 2006.

[52] R. Blum, J. Ross, E. A. Brown, and S. P. DeWeerth. An integrated system for simultaneous, multichannel neuronal stimulation and recording. *IEEE Transactions on Circuits and System-I: Regular Papers*, 54(12):2608–2618, 2007.

[53] S. Venkatraman, K. Elkabany, J. D. Long, Y. Yao, and J. M. Carmena. A system for neural recording and closed-loop intracortical microstimulation in awake rodents. *IEEE Transactions on Biomedical Engineering*, 56(1):15–22, Jan 2009.

[54] J. Rolston, R. Gross, and S. Potter. Neurorighter: Closed-loop multielectrode stimulation and recording for freely moving animals and cell cultures. *IEEE Eng. in Med. Bio. Society (EMBC)*, pages 6489–6492, 2009.

[55] J. Lee and H. Rhew. A 64 channel programmable closed-loop neurostimulator with 8 channel neural amplifier and logarithmic ADC. *IEEE J. Solid-State Circuits*, 45(9):1935–1945, 2010.

[56] F. Shahrokhi and K. Abdelhalim. The 128-channel fully differential digital integrated neural recording and stimulation interface. *IEEE Transactions on Biomedical Circuits and Systems*, 4(3), June 2010.

[57] S. F. Liang, F. Z. Shaw, C. P. Young, D. W. Chang, and Y. C. Liao. A closed-loop brain computer interface for real-time seizure detection and control. *IEEE Eng. in Med. Bio. Society (EMBC)*, pages 4950–4953, 2010.

[58] D. loi and et al. Peripheral neural activity recording and stimulation system. *IEEE Transactions on Biomedical Circuits and Systems*, 5(4):368–379, Aug. 2011.

[59] U. Bihr, T. Ungru, H. Xu, J. Anders, J. Becker, and M. Ortmanns. A bidirectional neural interface with a hv stimulator and a lv neural amplifier. *IEEE Int. Symp. Circuits Syst. (ISCAS)*, pages 401–404, May 2013.

[60] K. Abdelhalim and et al. 64-channel uwb wireless neural vector analyzer soc with a closed-loop phase synchrony-triggered neurostimulator. *IEEE J. Solid-State Circuits*, 2013.

[61] X. Liu, B. Subei, M. Zhang, A. G. Richardson, T. H. Lucas, and J. Van der Spiegel. The pennbmbi: A general purpose wireless brain-machine-brain interface system for unrestrained animals. *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 650–653, May 2014. URL http://ieeexplore.ieee.org/document/6865219/.

[62] K. Limnuson and et al. A bidirectional neural interface soc with an integrated spike recorder, microstimulator, and low-power processor for real-time stimulus artifact rejection. *CICC*, 2014.

[63] G. Angotzi and et al. A programmable closed-loop recording and stimulating wireless system for behaving small laboratory animals. *Scientific Reports*, 4, 2014.

[64] M. Ballini, J. Muller, P. Livi, Y. Chen, U. Frey, A. Stettler, A. Shadmani, V. Viswam, I. L. Jones, D. Jackel, M. Radivojevic, M. K. Lewandowska, W. Gong, M. Fiscella, D. J. Bakkum, F. Heer, A. Hierlemann, J. Muller, P. Livi, Y. Chen, A. Stettler, A. Shadmani, V. Viswam, I. L. Jones, D. Jackel, M. Radivojevic, M. K. Lewandowska, W. Gong, M. Fiscella, D. J. Bakkum, and A. Hierlemann. A 1024-channel CMOS microelectrode array with 26,400 electrodes for recording and stimulation of electrogenic cells in vitro. *IEEE J. Solid-State Circuits*, 49(11):2705–2719, 2014.

[65] P. Cong and et al. A 32-channel modular bi-directional neural interface system with embedded dsp for closed-loop operation. *ESSCIRC*, 2014.

[66] T. K. Nguyen and et al. Closed-loop optical neural stimulation based on a 32-channel low-noise recording system with online spike sorting. *Journal of neural engineering*, 11(4), 2014.

[67] X. Liu, B. Subei, M. Zhang, A. G. Richardson, T. H. Lucas, and J. Van der Spiegel. The pennbmbi : Design of a general purpose wireless brain-machine-brain interface system. *IEEE Transactions on Biomedical Circuits and Systems*, 9(2):248–258, 2015. URL http://ieeexplore.ieee.org/document/7055376/.

[68] X. Liu, M. Zhang, A. G. Richardson, T. H. Lucas, and J. Van der Spiegel. A 12-channel bidirectional neural interface chip with integrated channel-level feature extraction and pid con-troller. *IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2015.

[69] R. Shulyzki, K. Abdelhalim, A. Bagheri, M. T. Salam, C. M. Florez, J. L. P. Velazquez, P. L. Carlen, and R. Genov. 320-channel active probe for high-resolution neuromonitoring and responsive neurostimulation. *IEEE Trans. Biomed. Circuits Syst.*, 9(1), 2015.

[70] M. A. Bin Altaf, C. Zhang, and J. Yoo. A 16-channel patient-specific seizure onset and termination detection soc with impedance-adaptive transcranial electrical stimulator. *IEEE J. Solid-State Circuits*, 50(11):2728–2740, 2015.

[71] Adam E. Mendrela and et al. Enabling closed-loop neural interface: A bi-directional interface circuit with stimulation artifact cancellation and cross-channel cm noise suppression. *VLSI Circuits (VLSI Circuits), 2015 Symposium on*, 2015.

[72] W. Biederman, D. J. Yeager, N. Narevsky, J. Leverett, R. Neely, J. M. Carmena, E. Alon, and J. M. Rabaey. A 4.78mm2 fully-integrated neuromodulation soc combining 64 acquisition channels with digital compression and simultaneous dual stimulation. *IEEE J. Solid-State Circuits*, 50(4):1038–1047, 2015.

[73] X. Liu, M. Zhang, A. G. Richardson, T. H. Lucas, and J. Van der Spiegel. Design of a closed-loop, bi-directional brain machine interface system with energy efficient neural feature ex-traction and PID control. *IEEE Transactions on Biomedical Circuits and Systems*, 2016.

[74] A. Abdi and H.-K. Cha. A bidirectional neural interface CMOS analog front-end ic with embedded isolation switch for implantable devices. *Microelectronics J.*, 58:70–75, April 2016.

[75] Adam E Mendrela, Jihyun Cho, Jeffrey A Fredenburg, Vivek Nagaraj, Theoden I Netoff, Michael P Flynn, and Euisik Yoon. A bidirectional neural interface circuit with active stimulation artifact cancellation and cross-channel common-mode noise suppression. *IEEE J. Solid-State Circuits*, 51(4), 2016.

[76] Shoaran, Mahsa, Shahshahani, Masoud, Farivar, Masoud, Almajano, Joyel, Shahshahani, Amirhossein, Schmid, Alexandre, Bragin, Anatol, Leblebici, Yusuf, Emami, and Azit. A 16-channel 1.1mm2 implantable seizure control soc with sub- μw / channel consumption and closed-loop stimulation in 0.18μm CMOS. *VLSI Symposium on Circuits*, pages 256–257, 2016.

[77] Y. Su, S. Routhu, K. Moon, S. Q. Lee, W. Youm, and Y. Ozturk. A wireless 32-channel implantable bidirectional brain machine interface. *Sensors*, pages 1–19, 2016.

[78] Haas, Michael, Jens Anders, and Maurits Ortmanns. A bidirectional neural interface featuring a tunable recorder and electrode impedance estimation. *Biomedical Circuits and Systems Conference (BioCAS)*, 2016.

[79] X. Liu, H. Zhu, M. Zhang, A. G. Richardson, S. Y. Sritharan, D. Ge, Y. Shu, T. H. Lucas, and J. Van der Spiegel. A fully integrated wireless sensor-brain interface system to restore finger sensation. *Circuits and Systems, IEEE International Symposium on (ISCAS)*, May 2017.

[80] G. Santhanam, S. I. Ryu, B. M. Yu, A. Afshar, and K. V. Shenoy. A high-performance brain-computer interface. *Nature*, 442:195–8, Jul 2006.

[81] C. A. Chestek and et al. Hermesc : Low-power wireless neural recording system for freely moving primates. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 17(4):330–338, Aug 2009.

[82] H. Miranda, V. Gilja, C. A. Chestek, K. V. Shenoy, and T. H. Meng. Hermesd: A high-rate long-range wireless transmission system for simultaneous multi-channel neural recording applications. *IEEE Tans. Biomed. Circuits Syst.*, 4 (3):181–191, Jun 2010.

[83] H. Gao and et al. Hermese: A 96-channel full data rate direct neural interface in 0.13um CMOS. *IEEE J. Solid-State Circuits*, 47(4):1043–1055, April 2012.

[84] T. Jochum, T. Denison, and P. Wolf. Integrated circuit amplifiers for multi-electrode intracortical recording. *J. Neural Eng.*, 6(1):12001–26, Feb 2009.

[85] I. H. Stevenson and K. P. Kording. How advances in neural recording affect data analysis. *Nat. Neurosci.*, 14(2):139–142, 2011.

[86] Human brain project framework partnership agreement proposal. June 2014.

[87] K. A. Ng, E. Greenwald, Y. P. Xu, and N. V. Thakor. Implantable neurotechnologies: a review of integrated circuit neural amplifiers. *Med. Bio. Eng. Comput.*, 33(4):395–401, Jan 2015.

[88] R. Yazicioglu and P. Merken. A 60 mu w 60 nv/root hz readout front-end for portable biopotential acquisition systems. *IEEE J. Solid-State Circuits*, 42(5): 1100–1110, 2007.

[89] N. Van Helleputte, S. Kim, H. Kim, J. P. Kim, C. Van Hoof, and R. F. Yazicioglu. A 160 ua biopotential acquisition ic with fully integrated ia and motion artifact suppression. *IEEE Tans. Biomed. Circuits Syst.*, 6(6):552–61, Dec 2012.

[90] R. Harrison and P. Watkins. A low-power integrated circuit for a wireless 100-electrode neural recording system. *IEEE J. Solid-State Circuits*, 42(1):123–133, Jan 2007.

[91] W. Wattanapanitch, M. Fee, and R. SARpeshkar. An energy-efficient micropower neural recording amplifier. *IEEE Tans. Biomed. Circuits Syst.*, 1(2):136–47, Jun 2007.

[92] T. Denison, K. Consoer, W. Santa, A. Avestruz, J. Cooley, and A. Kelly. A 2 uw 100 nv/rthz chopper-stabilized instrumentation amplifier for chronic measurement of neural field potentials. *IEEE J. Solid-State Circuits*, 42(12):2934–2945, 2007.

[93] Q. Fan and et al. A 1.8μw 60 nv/rthz capacitively-coupled chopper instrumentation amplifier in 65 nm CMOS for wireless sensor nodes. *IEEE J. Solid-State Circuits*, 1(2):136–47, Jun 2007.

[94] X. Zou, L. Liu, J. Cheong, and L. Yao. A 100-channel 1-mw implantable neural recording ic. *IEEE Trans. on Circuits and Syst.-I: Regular Papers*, 60(10): 2584–2596, Oct 2013.

[95] M. S. Chae, Z. Yang, M. R. Yuce, L. Hoang, and W. Liu. A 128-channel 6 mw wireless neural recording ic with spike feature extraction and uwb transmitter. *IEEE Trans. Neural Syst. Rehab. Eng.*, 17(4):312–21, Aug 2009.

[96] R. Muller, H.-P. Le, W. Li, P. Ledochowitsch, S. Gambini, T. Bjorninen, and J. M. Rabaey. A minimally invasive 64-channel wireless uecog implant. *IEEE J. Solid-State Circuits*, 50(1):344–359, Aprl. 2015.

[97] R. R. Harrison, R. J. Kier, C. A. Chestek, and V. Gilja. Wireless neural recording with single low-power integrated circuit. *IEEE Trans. Neural Syst. Rehab. Eng.*, 17(4):322–329, Aug 2009.

[98] A. Borna and K. Najafi. A low power light weight wireless multichannel microsystem for reliable neural recording. *IEEE J. Solid-State Circuits*, 49(2): 439–451, Feb 2014.

[99] P. Mohseni, K. Najafi, S. J. Eliades, and X. Wang. Wireless multichannel biopotential recording using an integrated fm telemetry circuit. *IEEE Trans. Neural Syst. Rehab. Eng.*, 13(3):263–271, Sep 2005.

[100] N. M. Neihart and R. R. Harrison. Micropower circuits for bidirectional wireless telemetry in neural recording applications. *IEEE Trans. on Biomed. Eng.*, 52 (11):1950–1959, Nov. 2005.

[101] W. Biederman and D. Yeager. A fully-integrated, miniaturized (0.125 mm2) 10.5 uw wireless neural sensor. *IEEE J. Solid-State Circuits*, 48(4):960–970, Aprl 2013.

[102] D. J. Yeager, J. Holleman, R. Prasad, J. R. Smith, and B. P. Otis. Neuralwisp: A wirelessly powered neural interface with 1-m range. *IEEE Tans. Biomed. Circuits Syst.*, 3(6):379–87, Dec 2009.

[103] R. Olsson, Wise, and K. A three-dimensional neural recording microsystem with implantable data compression circuitry. *IEEE J. Solid-State Circuits*, 40 (12):2796–2804, Dec 2005.

[104] M. Steyaert, W. Sansen, and C. Zhongyuan. A micropower low-noise monolithic instrumentation amplifier for medical purposes. *Journal of Solid-State Circuit*, 22(6):1163–1168, 1987.

[105] J. Holleman. Design considerations for neural amplifiers. *IEEE Int. Symp. Circuits Syst. (ISCAS)*, pages 6331–6334, 2016.

[106] Muller, R.; Gambini, S.; Rabaey, and J.M. A 0.013 mm 5 mu w, dc-coupled neural signal acquisition ic with 0.5 v supply. *IEEE J. Solid-State Circuits*, (47):232–243, 2012.

[107] D. Han, Y. Zheng, R. Rajkumar, G. S. Dawe, and M. Je. A 0.45 v 100-channel neural-recording ic with sub-uw/channel consumption in 0.18um CMOS. *IEEE Trans. Biomed. Circuits Syst.*, 7(6):735–746, 2013.

[108] Gosselin, B.; Sawan, M.; Chapman, and C.A. A low-power integrated bioamplifier with active low- frequency suppression. *IEEE Trans. Biomed. Circuits Syst.*, 1(184-192), 2007.

[109] R. Rieger, M. Schuettler, D. Pal, C. Clarke, P. Langlois, J. Taylor, and N. Donaldson. Very low-noise eng amplifier system using CMOS technology. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 14(4):427–437, 2006.

[110] J. Holleman and B. Otis. A sub-microwatt low-noise amplifier for neural recording. *IEEE Eng. in Med. Bio. Society (EMBC)*, Aug 2007.

[111] W. A. Smith, B. J. Mogen, E. E. Fetz, V. S. Sathe, and B. P. Otis. Exploiting electrocorticographic spectral characteristics for optimized signal chain design: A 1.08uw analog front end with reduced ADC resolution requirements. *IEEE Trans. Biomed. Circuits Syst.*, pages 1–10, 2016.

[112] C. J. Deepu, X. Zhang, W. Liew, D. Liang, T. Wong, and Y. Lian. An ecg-on-chip with 535 nw / channel integrated lossless data compressor for wireless sensors. *IEEE J. Solid-State Circuits*, 49(11):2435–2448, 2014.

[113] N. Verma, A. Shoeb, J. Bohorquez, J. Dawson, J. Guttag, and A. P. Chandrakasan. A micro-power eeg acquisition soc with integrated feature extraction processor for a chronic seizure detection system. *IEEE J. Solid-State Circuits*, 45(4):804–816, Apr. 2010.

[114] C. H. Chan, J. Wills, J. LaCoss, J. J. Granacki, and J. Choma. A micro-power low-noise auto-zeroing CMOS amplifier for cortical neural prostheses. *IEEE 2006 Biomed. Circuits Syst. Conf. Healthc. Technol.*, pages 214–217, 2006.

[115] A. Bagheri; M. T. Salam; J. L. Perez Velazquez; R. Genov. Low-frequency noise and offset rejection in dc-coupled neural amplifiers: A review and digitally-assisted design tutorial. *IEEE Transactions on Biomedical Circuits and Systems*, 2016.

[116] Y. Chen, D. Jeon, Y. Lee, Y. Kim, Z. Foo, I. Lee, N. B. Langhals, G. Kruger, H. Oral, O. Berenfeld, Z. Zhang, D. Blaauw, and D. Sylvester. An injectable 64 nw ecg mixed-signal soc in 65 nm for arrhythmia monitoring. *IEEE J. Solid-State Circuits*, 50(1):375–390, 2015.

[117] J. Yoo, L. Yan, D. El-Damak, M. A. Bin Altaf, A. H. Shoeb, and A. P. Chandrakasan. An 8-channel scalable eeg acquisition soc with patient-specific seizure classification and recording processor. *IEEE J. Solid-State Circuits*, 48(1):214–228, 2013.

[118] R. Sarpeshkar2010. Ultra low power bioelectronics. *Cambridge University Press*, 2010.

[119] K.J. Miller, S. Zanos. E.E. Fetz, M. den Nijs, and J.G. Ojemann. Decoupling the cortical power spectrum reveals real-time representation of individual finger movements in humans. *J. Neurosci.*, 29:3132–3137, 2009.

[120] Kai J. Miller, Larry B. Sorensen, Jeffrey G. Ojemann, and Marcel den Nijs. Power-law scaling in the brain surface electric potential. *PLoS Comput Biol.*, 5 (12), Dec 2009.

[121] B. Razavi. Design of analog CMOS integrated circuits. *McGraw-Hill Companies, Inc.*, 2001.

[122] T. Jang, M. Choi, S. Jeong, S. Bang, D. Sylvester, and D. Blaauw. A 4.7nw 13.8ppm/c self-biased wakeup timer using a switched-resistor scheme. *ISSCC Dig. Tech. Papers*, pages 102–103, Feb 2016.

[123] J. Venesty. J. Chen, Y. Huang, and I. Cohen. Pearson correlation coefficient. *Noise Reduction in Speech Processing*, March 2009.

[124] J. Fredenburg and M. P. Flynn. ADC trends and impact on SAR ADC architecture and analysis. *Proc. Cust. Integr. Circuits Conf.*, pages 1–8, Nov 2015.

[125] T. Rabuske and J. Femandes. Charge-sharing SAR ADCs for low-voltage low-power applications. *Analog Circuits and Signal Processing*, 2017.

[126] S. Roshani and S. Roshani. Low power mixed-mode circuit design of SAR ADC. *Lambert Academic Publishing*, 2013.

[127] A. Agnes, E. Bonizzoni, P. Malcovati, and F. Maloberti. A 9.4-enob 1v 3.8uw 100ks/s SAR ADC with time-domain comparator. *Solid-State Circuit Conference (ISSCC), IEEE International*, pages 246–610, Feb 2008.

[128] A. Agnes, E. Bonizzoni, P. Malcovati, and F. Maloberti. An ultra-low power successive approximation a/d converter with time-domain comparator. *Analog Integr. Circuits Signal Process.*, 64(2):183–190, 2010.

[129] C. H. Kuo and C. E. Hsieh. A high energy-efficiency SARADC based on partial floating capacitor switching technique. *Proc. IEEE ESS- CIRC*, pages 475–478, 2011.

[130] M. Van Elzakker and E. Van Tuijl. A 10-bit charge-redistribution ADC consuming 1.9 uw at 1 ms/s. *IEEE J. Solid-State Circuits*, 45(5):1007–1015, 2010.

[131] Y. Zhu, C.-H. Chan, U-F. Chio, S.-W. Sin, S.-P. U, R. P. Martins, and F. Maloberti. A 10-bit 100-ms/s reference-free SAR ADC in 90 nm CMOS. *IEEE J. Solid-State Circuits*, 45:1111–1121, Jun 2010.

[132] S. W. M. Chen and R. W. Brodersen. A 6-bit 600-ms/s 5.3-mw asynchronous ADC in 0.13-um CMOS. *IEEE J. Solid-State Circuits*, 41(12):2669–2680, Dec 2006.

[133] S. O. Driscoll, K. V. Shenoy, and T. H. Meng. Adaptive resolution ADC array for an implantable neural sensor. *IEEE Trans. Biomed. Circuits Syst.*, 5(2): 120–130, 2011.

[134] B. Malki, T. Yamamoto, B. Verbruggen, P. Wambacq, and J. Craninckx. A 70 db dr 10 b 0-to-80 ms/s current-integrating SAR ADC with adaptive dynamic range. *IEEE J. Solid-State Circuits*, 49(5):1173–1183, 2014.

[135] V. Chaturvedi, T. Anand, and B. Amrutur. An 8-to-1 bit 1-ms/s SAR ADC with VGA and integrated data compression for neural recording. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 21(11):2034–2044, 2013.

[136] W. Tang, A. Osman, B. Kim, D.and Goldstein, C. Huang, B. Martini, V. Pieribone, and E. Culurciello and. Continuous time level crossing sampling ADC for bio-potential recording systems. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(6):1407–1418, Jun 2013.

[137] Y. Lyu and C. Wu. A low power 10bit 500ks/s delta-modulated SAR ADC (dmSAR ADC) for implantable medical devices. *IEEE Int. Symp. Circuits Syst. (ISCAS)*, pages 2046–2049, May 2013.

[138] G. Y. Huang, S. J. Chang, C. C. Liu, and Y. Z. Lin. A 1-uw 10-bit 200-ks/s SAR ADC with a bypass window for biomedical applications. *IEEE Int. Symp. Circuits Syst. (ISCAS)*, 47(11):2783–2795, May 2012.

[139] F. M. Yaul and A. P. Chandrakasan. A 10b SAR ADC with data-dependent energy savings using lsb-first successive approximation. *IEEE J. Solid-State Circuits*, 57(12):198–199, 2014.

[140] P. Harpe, H. Gao, R. Van Dommele, E. Cantatore, and A. Van Roermund. A 3nw signal-acquisition ic integrating an amplifier with 2.1 nef and a 1.5fj/conv-step ADC. *Solid-State Circuit Conference (ISSCC), IEEE International*, 58: 382–383, Feb 2015.

[141] B. Razavi. The bootstrapped switch. *IEEE Solid-State Circuits Magazine*, 7 (3):12–15, Summer 2015.

[142] F. Chen, A. P. Chandrakasan, and V. M. Stojanovi. Design and analysis of a hardware-efficient compressed sensing architecture for data compression in wireless sensors. *IEEE Journal of Solid-State Circuits*, 47(3):744–756, 2012.

[143] B. Gosselin and M. Sawan. An ultra low-power CMOS automatic action potential detector. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 17(4):346–353, Aug 2009.

[144] S. Narasimhan, H. J. Chiel, and S. Bhunia. Ultra-low-power and robust digital-signal-processing hardware for implantable neural interface microsystems. *IEEE Transaction of Biomedical Circuits and Systems*, 5(2):169–78, Apr. 2009.

[145] Karim G. Oweiss, A. Mason, Y. Suhail, A. M. Kamboh, and K. E. Thomson. A scalable wavelet transform VLSI architecture for real-time signal processing in high-density intra-cortical implants. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 54(6):1266–1278, June 2007.

[146] Y. Yang, A. M. Kamboh, and A. J. Mason. A configurable realtime dwt-based neural data compression and communication VLSI system for wireless implants. *Journal of Neuroscience Methods*, 227:140–150, Apr 2014.

[147] E. J. Candes and M. B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, Mar 2008.

[148] D. L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, April 2006.

[149] Y. Oike and A. El Gamal. CMOS image sensor with per-column sigma-delta ADC and programmable compressed sensing. *IEEE Journal of Solid-State Circuits*, 48(1):319–328, 2013.

[150] J. Zhang, Y. Suo, S. Mitra, S. Chin, S. Hsiao, R. F. Yazicioglu, T. D. Tran, and R. Etienne-Cummings. An efficient and compact compressed sensing microsystem for implantable neural recordings. *IEEE Tansactions on Biomedical Circuits and Systems*, 8(4):485–496, Aug 2014.

[151] X. Chen, Z. Yu, and S. Hoyos. A sub-nyquist rate sampling receiver exploiting compressive sensing. *IEEE Trans. on Circuits and Syst.-I: Regular Papers*, pages 507–520, 2011.

[152] S. Aviyente. Compressed sensing framework for eeg compression. *IEEE Statistical Signal Processing*, pages 181–184, 2007.

[153] T. Xiong, J. Zhang, Y. Suo, D. N. Tran, R. Etienne-Cummings, S. Chin, and T. D. Tran. A dictionary learning algorithm for multi-channel neural recordings. *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 9–12, 2014.

[154] T. Xiong, J. Zhang, Y. Suo, D. N. Tran, R. Etienne-Cummings, S. Chin, and T. D. Tran. An unsupervised dictionary learning algorithm for neural recordings. *IEEE International Symposium onCircuits and Systems (ISCAS)*, pages 1010–1013, 2015.

[155] R. Baraniuk. Compressive sensing. *IEEE Signal Processing Magazine*, pages 118–121, July 2007.

[156] M. Shoaib, N.K. Jha, and N. Verma. Signal processing with direct computations on compressively sensed data. *Very Large Scale Integration Systems, IEEE Trans. on*, 23(1):30–43, Jan 2015.

[157] Y. Lu and W. Ki. A 13.56 mhz CMOS active rectifier with switched-offset and compensated biasing for biomedical wireless power transfer systems. *IEEE Tans. Biomed. Circuits Syst.*, 8(3):334–344, 2014.

[158] X. Li, C. Tsui, and W. Ki. A 13.56mhz wireless power transfer system with reconfigurable resonant regulating rectifier and wireless power control for implantable medical devices. *IEEE J. Solid-State Circuits*, 50(4), Aprl. 2014.

[159] M. Kiani and M. Ghovanloo. A 13.56-mbps pulse delay modulation based transceiver for simultaneous near-field data and power transmission. *IEEE Trans. Biomed. Circuits Syst.*, 9(1), Feb 2015.

[160] D. Gangopadhyay, E. G. Allstot, a M. R. Dixon, K. Natarajan, S. Gupta, and D. J. Allstot. Compressed sensing analog front-end for bio-sensor applications. *IEEE J. Solid-State Circuits*, 49(2):426–438, 2014.

[161] J. Zhang, S. Mitra, Y. Suo, A. Cheng, T. Xiong, F. Michon, M. Welkenhuysen, F. Kloosterman, P. S. Chin, S. Hsiao, T. D. Tran, F. Yazicioglu, and R. Etienne-Cummings. A closed-loop compressive-sensing-based neural recording system. *J. Neural Eng.*, 12(3), 2015.

[162] Wikipedia and the Free Encyclopedia. Feature extraction. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Feature_extraction.

[163] J. Echauz. Feature extraction for brain computer interface. *Lecture notes from BE521, the University of Pennsylvania*, (3):1–2, Jan 2013.

[164] W. Chen, X. Liu, and B. Litt. Logistic-weighted regression improves decoding of finger flexion from electrocorticographic signals. *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2014.

[165] J. Donoghue, J. Sanes, N. Hatsopoulos, and G. Gaal. Neural discharge and local field potentials oscillations in primate motor cortex during voluntary movements. *J. Neurophysiol.*, 79(1):159–173, 1998.

[166] A. Shoeb, S. Schachter, D. Schomer, B. Bourgeois, and J. Guttag. Detecting seizure onset in the ambulatory setting: Demonstrating feasibility. *Proc. IEEE Eng. Med. Biol. Soc. Conf.*, pages 3546–3550, Sep 2005.

[167] A.-T. Avestruz, W. Santa, D. Carlson, R. Jensen, S. Stanslaski, A. Helfenstine, and T. Denison. A 5 uw/channel spectral analysis ic for chronic bidirectional brain-machine interfaces. *IEEE J. Solid-State Circuits*, 43(12):3006–3024, Dec 2008.

[168] S. Debener, A. Beauducei, D. Nessler, B. Brocke, H. Heilemann, and J. Kayser. Is resting anterior eeg alpha asymmetry a trait marker for depression. *Neuropsychobiology,*, 41:31–37, 2000.

[169] F. Zhang, A. Mishra, A. G. Richardson, and B. Otis. A low-power ecog / eeg processing ic with integrated multiband energy extractor. *The IEEE Transactions on Circuits and Systems - Part I: Regular Papers*, 58(9):2069–2082, 2011.

[170] R. R. Harrison, G. Santhanam, and K. V Shenoy. Local field potential measurement with low-power analog integrated circuit. *IEEE Eng. in Med. Bio. Society (EMBC)*, 6:4607–70, Jan 2004.

[171] Gagnon-Turcotte, G. Camaro, C. O Dufresne, Gosselin, and B. Comparison of low-power biopotential processors for on-the-fly spike detection. *Proceedings - IEEE International Symposium on Circuits and Systems*, pages 802–805, May 2015.

[172] V. Shalchyan, W. Jensen, and D. Farina. Spike detection and clustering with unsupervised wavelet optimization in extracellular neural recordings. *IEEE Trans. Biomed. Eng.*, 59(9):2576–2585, 2012.

[173] Wikipedia and the Free Encyclopedia. Principal component analysis. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Principal_component_analysis.

[174] Wikipedia and the Free Encyclopedia. Independent component analysis. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Independent_component_analysis.

[175] Wikipedia and the Free Encyclopedia. Genetic algorithm. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Genetic_algorithm.

[176] Ciaccio, E., Dunn, Akay, and M. Biosignal pattern recognition and interpretation systems. 2. methods for feature extraction and selection. *IEEE Eng. Med. Biol. Mag.*, 12:106–113, 1993.

[177] Stark E and Abeles M. Predicting movement from multiunit activity. *J Neurosci*, 27:8387–94, 2007.

[178] Ahmed B, Redissi A, and Tafreshi R. An automatic sleep spindle detector based on wavelets and the teager energy operator. *Conf Proc IEEE Eng Med Biol Soc*, pages 2596–9, 2009.

[179] Lin SC and Gervasoni D. Defining global brain states using multielectrode field potential recordings. *Methods for Neural Ensemble Recordings*, 2008.

[180] Ohara S, Mima T, Baba K, Ikeda A, and Kunieda T. Increased synchronization of cortical oscillatory activities between human supplementary motor and primary sensorimotor areas during voluntary movements. *J Neurosci*, 21:9377–86, 2001.

[181] Graef A, Hartmann M, Flamm C, Baumgartner C, Deistler M, and Kluge T. A novel method for the identification of synchronization effects in multichannel ecog with an application to epilepsy. *Biol Cybern*, 107:321–35, 2013.

[182] M. S. Lewicki. A review of methods for spike sorting: the detection and classification of neural action potentials. *Network: Comput. Neural Syst.*, 9(4):53–78, Nov 1998.

[183] Hiseni S, Sawigun C, Ngamkham W, and Serdijn WA. A compact, nano-power CMOS action potential detector. *Biomed. Circuits Syst. Conf.*, pages 97–100, 2009.

[184] Gibson S, Judy JW, and Markovic D. An fpga-based platform for accelerated offline spike sorting.

[185] Tam DC. Real-time estimation of predictive firing rate. *Neurocomputing*, 52 (4):637–41, 2003.

[186] Smith AC, Scalon JD, Wirth S, Yanike M, Suzuki WA, and Brown EN. State-space algorithms for estimating spike rate functions. *Comput Intell Neurosci*, page 426539, 2010.

[187] H. Scherberger, M. R. Jarvis, and R. A. Andersen. Cortical local field potential encodes movement intentions in the posterior parietal cortex. *Nature*, 46(2): 347–354, 2005.

[188] M. Spler, A. Walter, A. Ramos Murguialday, G. Naros, N. Birbaumer, A. Gharabaghi, W. Rosenstiel, and M. Bogdan. Decoding of motor intentions from epidural ecog recordings in severely paralyzed chronic stroke patients. *Journal of Neural Engineering*, 11(6), 2014.

[189] G. Buzsaki and A. Draguhn. Neuronal oscillations in cortical networks. *Science*, 304:1926–1930, 2004.

[190] M. Penttonen and G. Buzsaki. Natural logarithmic relationship between brain oscillators. *Thalamus & Related Systems*, 2(2):145–152, 2003.

[191] R. R. Harrison. The design of integrated circuits to observe brain activity. *Proc. IEEE*, 96(7):1203–1216, Jul 2008.

[192] P. M. Furth and H. a. Ommani. Low-voltage highly-linear transconductor design in subthreshold CMOS. *IEEE Int. Midwest Symp. Circuits Syst. (MWSCAS)*, 1(156-159), 1997.

[193] Rahul SARpeshkar. Ultra low power bioelectronics: Fundamentals, biomedical applications, and bio-inspired systems. *Cambridge University Press*, 2010.

[194] Willy M.C. Sansen. Analog design essentials. *Springer*, 2006.

[195] R. Torrance, T. Viswanathan, and J. Hanson. CMOS voltage to current transducers. *IEEE Trans. Circuits Syst.*, 32(11):1097–1104, 1985.

[196] J. Gak, M. R. Miguez, and A. Arnaud. Nanopower otas with improved linearity and low input offset using bulk degeneration. *IEEE Trans. Circuits Syst. - I*, 61(3):689–698, 2014.

[197] C. D. Salthouse and R. SARpeshkar. A practical micropower programmable bandpass filter for use in bionic ears. *IEEE J. Solid-State Circuits*, 38(1):63–70, Jan 2003.

[198] A. Kumar. A wide dynamic range high-q high-frequency bandpass filter with an automatic quality factor tuning scheme. *PhD Thesis, Georgia Institute of Technology*, 2008.

[199] Wikipedia and The Free Encyclopedia. Digital biquad filter. Retrieved Jan. 2017. URL https://en.wikipedia.org/wiki/Digital_biquad_filter.

[200] B. Gilbert. A precise four-quadrant multiplier with subnanosecond response. *IEEE J. Solid-State Circuits*, 3(4):365–373, Dec 1968.

[201] Gunhee Han and E. Sanchez-Sinencio. CMOS transconductance multipliers: a tutorial. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 45(12):1550–1563, Dec 1998.

[202] S-C. Qin and Randy L. Geiger. A 5V CMOS analog multiplier. *IEEE Journal of Solid-State Circuits*, 22(6):1143–1146, 1987.

[203] H. G. Rey, C. Pedreira, and R. Quian Quiroga and. Past, present and future of spike sorting techniques. *Brain Res. Bull.*, 119:106–117, 2015.

[204] R. Quian Quiroga, L. Reddy, C. Koch, and I. Fried. Decoding visual inputs from multiple neurons in the human temporal lobe. *J. Neurophysiol.*, 98(4): 1997–2007, Oct 2007.

[205] R. Q. Quiroga. Spike sorting. *Curr. Biol.*, 22(2):45–46, 2012.

[206] Peter Dayan and L. F. Abbott. Theoretical neuroscience computational and mathematical modeling of neural systems. *The MIT Press*, 2001.

[207] I. Obeid and P. D. Wolf. Evaluation of spike-detection algorithms for a brain-machine interface application. *IEEE Trans. Biomed. Eng.*, 51(6):905–911, 2004.

[208] J. Holleman and A. Mishra. A micro-power neural spike detector and feature extractor in 0.13um CMOS. 2008.

[209] Zumsteg, ZS, Kemere, and C. Power feasibility of implantable digital spike sorting circuits for neural prosthetic systems. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 13(3):4237–40, Sept. 2005.

[210] L. Chen, B. Shi, and C. Lu. A high speed/power ratio continuous-time CMOS current comparator. *7th IEEE Int. Conf. Electron. Circuits Syst.*, 2:883–886, 2000.

[211] S. Al-Ahdab, R. Lotfi, and W. A. Serdijn. A 1-v 225-nw 1ks/s current successive approximation ADC for pacemakers. *6th Conference on Ph.D. Research in Microelectronics & Electronics*, pages 1–4, 2010.

[212] Wikipedia and The Free Encyclopedia. Mached filter. Retrieved Jan. 2017. URL https://en.wikipedia.org/wiki/Matched_filter.

[213] Turin and G. An introduction to matched filters. *IRE Transactions on Information Theory*, 6:311–329, Jan 1960.

[214] Frank C. Robey and et al. A cfar adaptive matched filter detector. *IEEE Transactions on Aerospace and Electronic Systems*, 28(1):208–216, 1992.

[215] Ian G. Cumming and Frank H. Wong. Digital processing of synthetic aperture radar data. *Artech house*, 1(2), Mar 2005.

[216] Seto, Naoki, Seiji Kawamura, and Takashi Nakamura. Possibility of direct measurement of the acceleration of the universe using 0.1 hz band laser interferometer gravitational wave antenna in space. *Physical Review Letters*, 87(22): 221103, 2001.

[217] Al-Rawi, Mohammed, Munib Qutaishat, and Mohammed Arrar. An improved matched filter for blood vessel detection of digital retinal images. *Computers in Biology and Medicine*, 37(2):262–267, 2007.

[218] I. Rankman and R. Chandra. An algorithm for generating templates of neural waveforms. *Proceedings of the 15th Annual International Conference of the IEEE*, 1993.

[219] Wu, Shun Chi, A. Lee Swindlehurst, and Zoran Nenadic. Matched subspace detector based feature extraction for sorting of multi-sensor action potentials. *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011.

[220] Alfred M. Haas, Marc H. Cohen, and Pamela A. Abshires. Real-time variance based template matching spike sorting system. *Life Science Systems and Applications Workshop*, pages 7–10, 2007.

[221] D. J. Krusienski, G. Schalk, D. J. McFarland, and J. R. Wolpaw. A mu-rhythm matched filter for continuous control of a brain-computer interface. *IEEE Trans. Biomed. Eng.*, 54(2):273–280, 2007.

[222] Wikipedia and The Free Encyclopedia. Digital biquad filter. Retrieved Jan. 2017. URL https://en.wikipedia.org/wiki/Cauchy-Schwarz_inequality.

[223] T.H. Chao, A.M. Tai, M.S. Dymek, and F.T.S. Yu. Optimum correlation detection by prewhitening. *Applied Optics*, 19(14):2461–2464, 1980.

[224] A. B. Tort and et al. Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of neurophysiology*, 104(2):1195–1210, 2010.

[225] M. Ghovanloo and K. Iniewski. Integrated circuits for neural interfacing: Neural stimulation. *VLSI Circuits for Biomedical Applications*, pages 191–199, 2008.

[226] Wikipedia. Electrical brain stimulation. . URL https://en.wikipedia.org/wiki/Electrical_brain_stimulation.

[227] Haddad SAP, Houben RPM, and Serdijin WA. The evolution of pacemakers. *IEEE Eng. in Med and Bio Magazine*, 25(3):38–48, May-June 2006.

[228] A. Demosthenous, I. F. Triantis, and X. Liu. Circuits for implantable neural recording and stimulation. pages 207–240, 2008.

[229] Wikipedia. Neurostimulation. . URL https://en.wikipedia.org/wiki/Neurostimulation.

[230] Z. B. Kagan, A. K. RamRakhyani, G. Lazzi, R. A. Normann, and D. J. Warren. In vivo magnetic stimulation of rat sciatic nerve with centimeter- and millimeter-scale solenoid coils. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(11):1138–1147, Nov 2016.

[231] S. W. Lee and S. Fried. Enhanced control of cortical pyramidal neurons with micro-magnetic stimulation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, (99), 2016.

[232] R. D. Meyer, S. F. Cogan, T. H. Nguyen, and R. D. Rauh. Electrodeposited iridium oxide for neural stimulation and recording electrodes. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 9(1):2–11, Mar 2001.

[233] J. Chen, K. D. Wise, J. F. Hetke, and S. C. Bledsoe. A multichannel neural probe for selective chemical delivery at the cellular level. *IEEE Transactions on Biomedical Engineering*, 44(8):760–769, Aug. 1997.

[234] K. Paralikar, P. Cong, W. Santa, D. Dinsmoor, B. Hocken, G. Munns, J. Giftakis, and T. Denison. An implantable 5mw/channel dual-wavelength optogenetic stimulator for therapeutic neuromodulation research. *ISSCC Dig. Tech. Papers*, pages 238–240, Feb 2010.

[235] K. Paralikar, P. Cong, O. Yizhar, L. E. Fenno, W. Santa, C. Nielsen, D. Dinsmoor, B. Hocken, G. O. Munns, J. Giftakis, K. Deisseroth, and T. Denison. An implantable optical stimulation delivery system for actuating an excitable biosubstrate. *J. of Solid-State Circuits*, 46(1):321–332, Jan 2011.

[236] Krzysztof Iniewski. VLSI circuits for biomedical applications. *Artech house Inc.*, 2008.

[237] S. F. Cogan. Neural stimulation and recording electrodes. *Annu. Rev. Biomed. Eng.*, 10:275–309, Jan 2008.

[238] D. A. Borkholder. Cell based sensors using microelectrodes. *Ph. D. Dissertation, Stanford University*, Nov. 1998.

[239] Richard A. Blum, James D. Ross, Samir K. Das, Edgar A. Brown, and Stephen P. DeWeerth. Models of stimulation artifacts applied to integrated circuit design. *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, Sep 2004.

[240] J. Sit and R. SARpeshkar. A low-power blocking-capacitor-free with less than 6 na dc error for 1-ma full-scale stimulation. *IEEE Tran. on Biomeidcal Circuits and Systems*, 1(3):172–183, 2007.

[241] S. Kelly and J. Wyatt. A power-efficient neural tissue stimulator with energy recovery. *IEEE Tansactions on Biomedical Circuits and Systems*, 5(1):20–29, 2011.

[242] L. Wong, S. Hossain, A. T. Jorgen Edivinsson, D. Rivas, and H. Haas. A very low-power CMOS mixed-signal ic for implantable pacemaker applications. *Journal of Solid-State Circuit*, 39(12):2446–2456, 2004.

[243] B. Thurgood, D. Warren, N. M. Ledbetter, G. a. Clark, and R. R. Harrison. A wireless integrated circuit for 100-channel charge-balanced neural stimulation. *IEEE Tran. on Biomeidcal Circuits and Systems*, 3(6):405–414, Dec. 2009.

[244] M. Sivaprakasam, W. Liu, M. Humayun, and J. Weiland. A variable range bi-phasic current stimulus driver circuitry for an implantable retinal prosthetic device. *IEEE Journal of Solid-State Circuits*, 40(3):763–771, Mar 2005.

[245] M. Ghovanloo. Switched-capacitor based implantable low-power wireless microstimulating systems. *IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 2197–2200, 2006.

[246] H. Lee, K. Y. Kwon, and W. Li. A power-efficient switched-capacitor stimulating system for electrical / optical deep brain stimulation. 50(1):360–374, 2015.

[247] M. Sahin and Y. Tie. Non-rectangular waveforms for neural stimulation with practical electrodes. *J. Neural Eng.*, 4:227–233, 2007.

[248] C. M. Zierhofer, I. J. H.-Desoyer, and E. S. Hochmair. Electronic de- sign of a cochlear implant for multichannel high-rate pulsatile stimu- lation strategies. *IEEE Trans. Neural Syst. Rehabil. Eng*, 3(1):112–116, Mar 1995.

[249] M. Bugbee, N. de N.Donalsdon, A. Lickel, N. J. Rijkhoff, and J.Taylor. An implant for chronic selective stimulation of nerves. *Med. Eng. Phys*, 23(1): 29–36, Jan 2001.

[250] J. D. Techer, S. Bernard, Y. Bertrand G. Cathebras, and D. Guiraud. New implantable stimulator for the fes of paralyzed muscles. *in Proc. 30th Eur. Solid-State Circuits Conf. (ESSCIRC04)*, pages 455–458, 2004.

[251] K. Chen, Z. Yang, and L. Hoang. An integrated 256-channel epiretinal prosthesis. *Solid-State Circuits, IEEE Journal of*, 45(9):1946–1956, 2010.

[252] K. Song, H. Lee, and S. Hong. A sub-10 na dc-balanced adaptive stimulator ic with multi-modal sensor for compact. *IEEE Tran. on Circuits and Systems*, 6 (6):533–541, 2012.

[253] M. Monge, M. Raj, M. Honarvar-nazari, H. Chang, Y. Zhao, J. Weiland M. Humayun, Y. Tai, and A. Emami-neyestanak. A fully intraocular 0.0169 mm^2/pixel 512-channel self-calibrating epiretinal prosthesis in 65nm CMOS. *ISSCC Dig. Tech. Papers*, pages 296–298, Feb 2013.

[254] K. Sooksood, T. Stieglitz, and M. Ortmanns. An active approach for charge balancing in functional electrical stimulation. *IEEE Tran. on Biomeidcal Circuits and Systems*, 4(3):162–170, 2010.

[255] M. Ortmanns, A.Rocke, M.Gehrke, and H.J.A.T.H. J.Tiedtke. A 232-channel epiretinal stimulator asic. *IEEE J. Solid-State Circuits*, 42(12):2946–2959, Dec 2007.

[256] E. Noorsal, K. Sooksood, and H. Xu. A neural stimulator frontend with high-voltage compliance and programmable pulse shape for epiretinal implants. *Journal of Solid-State Circuits*, Jan 2012.

[257] R. J. Baker. CMOS circuit design, layout, and simulation, 3rd edition. *IEEE Press*, Oct 2010.

[258] S. Kelly and et al. A power-efficient voltage-based neural tissue stimulator with energy recovery. *ISSCC Dig. Tech. Papers*, 26(2):579–588, Feb 2004.

[259] J. Vidal and M. Ghovanloo. Towards a switched-capacitor based stimulator for efficient deep-brain stimulation. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2010.

[260] Lapicque L. Recherches. Quantitatives sur l'excitation electrique des nerfs traites comme une polarization. *J Physiol (Paris)*, (9):622–35, 1907.

[261] F. Boi, T. Moraitis, V. De Feo, F. Diotalevi, C. Bartolozzi, G. Indiveri, and A. Vato. A bidirectional brain-machine interface featuring a neuromorphic hardware decoder. *Front. Neurosci.*, 10(12):1–15, 2016.

[262] M. Lebedev. Augmentation of sensorimotor functions with neural prostheses. *Opera. Med. Physiol.*, 2(3):211–227, 2016.

[263] A. Wallach, D. Eytan, A. Gal, C. Zrenner, and S. Marom. Neuronal response clamp. *Front. Neuroeng.*, 4(4), Jan 2011.

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

[265] L. Rossi, G Foffani, S Marceglia, F Bracchi, S Barbieri, and A Priori. An electronic device for artefact suppression in human local field potential recordings during deep brain stimulation. 4(2):96–106, Mar 2007.

[266] G.a. DeMichele and P.R. Troyk. Stimulus-resistant neural recording amplifier. *IEEE Eng. in Med. Bio. Society (EMBC)*, Sep 2003.

[267] Edgar A. Brown, James D. Ross, Richard A. Blum, Yoonkey Nam, Bruce C. Wheeler, and Stephen P. DeWeerth. Stimulus-artifact elimination in a multi-electrode system. *IEEE Trans. Biomed. Circuits Syst.*, 2(1):10–21, Mar 2008.

[268] C. Ethier, E. R. Oby, M. J. Bauman, and L. E. Miller. Restoration of grasp following paralysis through brain-controlled stimulation of muscles. *Nature*, 485:368–371, May 2012.

[269] C. E. Bouton, A. Shaikhouni, N. V. Annetta, M. A. Bockbrader, D. A. Friedenberg, D. M. Nielson, and et al. Restoring cortical control of functional movement in a human with quadriplegia. *Nature*, 533:247–50, May 2016.

[270] S. Musallam, B. D. Corneil, B. Greger, H. Scherberger, and R. A. Andersen. Cognitive control signals for neural prosthetics. *Science*, 305:258–62, Jul. 2004.

[271] H.-V. V Ngo, T. Martinetz, J. Born, and M. Mlle. Auditory closed-loop stimulation of the sleep slow oscillation enhances memory. *Neuron*, 78(3):545–53, May 2013.

[272] Real-time multi-channel stimulus artifact suppression by local curve fitting. *Journal of Neuroscience Methods*, 120(2):113–120, May 2002.

[273] Toshiba. Toshiba photocoupler ingaas ired & photo-transistor. *Datasheet*, 10 2015. URL https://toshiba.semicon-storage.com/info/docget.jsp?did=15284&prodName=TLP292-4.

[274] Burr Brown. High speed fet-input instrumentation amplifier. *Datasheet*, March 1998. URL http://www.ti.com/lit/ds/symlink/ina111.pdf.

[275] Wikipedia and the Free Encyclopedia. Control theory. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Control_theory.

[276] Wikipedia and the Free Encyclopedia. Feedback. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Feedback.

[277] Wikipedia and the Free Encyclopedia. Frequency compensation. Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/Frequency_compensation.

[278] Astrom K. J. and Hagglund T. H. New tuning methods for pid controllers. *Proceedings of the 3rd European Control Conference*, 1995.

[279] J. Liu, H. K. Khalil, and K. G. Oweiss. Neural feedback for instantaneous spatiotemporal modulation of afferent pathways in bi-directional brain-machine interfaces. *IEEE Trans Neural Syst Rehabil Eng*, 19:521–33, Oct 2011.

[280] J. Daly, J. Liu, M. Aghagolzadeh, and K. Oweiss. Optimal space-time precoding of artificial sensory feedback through mutichannel microstimulation in bi-directional brain-machine interfaces. *J Neural Eng*, 9, Nov 2012.

[281] K. Kurosawa, R. Futami, T. Watanabe, and N. Hoshimiya. Joint angle control by fes using a feedback error learning controller. *IEEE Trans Neural Syst Rehabil Eng*, 13:359–71, Sep 2005.

[282] M. Arsiero, H.-R. Lscher, and M. Giugliano. Real-time closed-loop electrophysiology: towards new frontiers in in vitro investigations in the neurosciences. *Arch. Ital. Biol.*, 145(3-4):193–209, Nov 2007.

[283] Lita, Ioan, Daniel Alexandru Visan, and Ion Bogdan Cioc. Fpaa based pid controller with applications in the nuclear domain. *32nd International Spring Seminar on Electronics Technology*, 2009.

[284] V. Aggarwal, Meng Mao, and U. M. O'Reilly. A self-tuning analog proportional-integral-derivative (pid) controller. *First NASA/ESA Conference on Adaptive Hardware and Systems (AHS'06)*, pages 12–19, Dec. 2006.

[285] L. Samet, N. Masmoudi, M. W. Kharrat, and L. Kamoun. A digital pid controller for real time and multi loop control: a comparative study. *IEEE International Conference on Electronics, Circuits and Systems. Surfing the Waves of Science and Technology*, pages 291–296, 1998.

[286] L. H. Keel, J. I. Rego, and S. P. Bhattacharyya. A new approach to digital pid controller design. *IEEE Transactions on Automatic Control*, 48(4):687–692, April 2003.

[287] W. K. Ho, C. C. Hang, and J. H. Zhou. Performance and gain and phase margins of well-known pi tuning formulas. *IEEE Transactions on Control Systems Technology*, 3(2):245–248, Jun 1995.

[288] K. H. Ang, G. Chong, and Y. Li. Pid control system analysis, design, and technology. *IEEE Trans. on Control Systems Technology*, 13(4):559–576, Jul. 2005.

[289] V. Gilja and et al. A high-performance neural prosthesis enabled by control algorithm design. *Nature neuroscience*, 15(12):1752–1757, 2012.

[290] Atmel. 8/16-bit atmel xmega microcontroller. Sep 2014. URL http://www.atmel.com/devices/ATXMEGA128A4U.aspx.

[291] Taxes Instruments. Micropower, single-supply, CMOS instrumentation amplifier. *INA2321 datasheet*, 2000.

[292] Atmel. At32uc3c series complete. *Atmel Datasheet*, Oct 2000.

[293] Analog Devices. 3-axis, ±2g/±4g/±8g/±16g digtial accelerometer. June 2009.

[294] Nordic Semiconductor. nrf24l01+ single chip 2.4ghz transceiver product specification v1.0. September 2008.

[295] S. K. Talwar, S. Xu, E. S. Hawley, S. A. Weiss, K. A. Moxon, and J. K. Chapin. Rat navigation guided by remote control. *Nature*, 417:37–8, May 2002.

[296] E. E. Thomson, R. Carra, and M. A. Nicolelis. Perceiving invisible light through a somatosensory cortical prosthesis. *Nat Commun*, 4:1482, 2013.

[297] R. Morris. Developments of a water-maze procedure for studying spatial learning in the rat. *J Neurosci Methods*, 11:47–60, May 1984.

[298] Texas Instruments. Ultralow noise, high psrr, fast rf 100-ma low-dropout linear regulators. June 2008. URL http://www.ti.com/product/TPS791.

[299] Linear Technology. Dual micropower dc/dc converter with positive and negative outputs. *Datasheet*, . URL http://www.linear.com/product/LT1945.

[300] Linear Technology. 100ma regulated charge-pump inverters in thinsot. *Datasheet*, . URL http://www.linear.com/product/LTC1983.

[301] Texas Instruments. High-precision, low-noise, rail-to-rail output 11-mhz jfet op amp. Aug 2016. URL http://www.ti.com/product/OPA2140.

[302] Andrew G. Richardson, Pauline K. Weigand Srihari Y. Sritharan, and Timothy H. Lucas. A chronic neural interface to the macaque dorsal column nuclei. *Journal of Neurophysiology*, 115(5):2255–64, May 2016.

[303] Texas Instruments. Tiva tm4c123gh6pm microcontroller. *Datasheet*, 10 2007-2013. URL http://www.ti.com/product/TM4C123GH6PM.

[304] Open Source. Fatfs - generic fat file system module. Retrieved 2017. URL http://elm-chan.org/fsw/ff/00index_e.html.

[305] Wikipedia and the Free Encyclopedia. Iic (inter-integrated circuit). Retrieved Jan 2017. URL https://en.wikipedia.org/wiki/I%C2%B2C.

[306] Diekelmann S and Born J. The memory function of sleep. *Nat. Rev. Neurosci.*, 11(2):114–26, 2010.

[307] Huber R, Ghilardi MF, Massimini M, and Tononi G. Local sleep and learning. *Nature*, 430(6995):78–81, 2004.

[308] Marshall L, Helgadottir H, Molle M, and Born J. Boosting slow oscillations during sleep potentiates memory. *Nature*, 444(7119):610–3, 2006.