BIDIRECTIONAL INTEGRATED NEURAL INTERFACE
FOR ADAPTIVE CORtical STIMULATION

by

Ruslana Shulyzki

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto

Copyright © 2009 by Ruslana Shulyzki
Abstract

Bidirectional Integrated Neural Interface for Adaptive Cortical Stimulation

Ruslana Shulyzki
Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto
2009

This thesis presents the VLSI implementation and characterization of a 256-channel bidirectional integrated neural interface for adaptive cortical stimulation.

The microsystem consists of 64 stimulation and 256 recording channels, implemented in a 0.35\(\mu\)m CMOS technology with a cell pitch of 200\(\mu\)m and total die size of 3.5mm \(\times\) 3.65mm. The stimulator is a current driver with an output current range of 20\(\mu\)A - 250\(\mu\)A. The current memory in every stimulator allows for simultaneous stimulation on multiple active channels. Circuit reuse in the stimulator and utilization of a single DAC yields a compact and low-power implementation. The recording channel has two stages of signal amplification and conditioning and a single-slope ADC. The measured input-referred noise is 7.99\(\mu\)V\(_{rms}\) over a 5kHz bandwidth. The total power consumption is 13.3mW.

A new approach to CMOS-microelectrode hybrid integration by on-chip \(Au\) multi-stud-bumping is also presented. It is validated by in vitro experimental measurements.
Acknowledgements

I would like to thank my supervisor, Professor Roman Genov, for his guidance throughout this project. My defense committee: Professor Chan Carusone, Professor Ng and Professor Hum for their feedback and time. CMC for fabrication access and bonding services. Dr. Carlen for allowing us to test in his lab in the Toronto Western Hospital. Demitre Serletis for his help with the biological experiments. And finally the people from ecehelp, for all their assistance and especially for saving our tapeout during the great network disaster of fall 2007.

I have to thank my parents, Ludmila and Aleksey Shulyzki and my sister, Anna Shulyzki for their unlimited love and support, and for making me laugh like nobody else can. Ora Gendler and Rika Valdman for always being there for me despite the geographical distance and constantly varying time zones.

I am grateful to all my friends at the University of Toronto. Farzaneh Shahrokhi for being the perfect partner in crime. Tina Tahmoureszadeh for all rain related activities. Azadeh Khaleghi for my first trip to Ottawa. Mike Bichan for the best morning conversations on every imaginable subject. Kentaro Yamamoto for being endlessly generous with his knowledge. Jospeh Aziz for introducing me to this project and his assistance. Karim Abdelhalim and Hamed Mazhab-Jafari for comic relief, dry-runs, hospital trips, bikes and birds and ice cream. Rituraj Singh, Tony Kao, Alireza Nilchi, Derek Ho, Dustin Dunwell, Bert Leesti, Trevor Caldwell, Dave Halupka, Aleksey Tyshchenko, Ricardo Aroca and Shahriar Shahramian. It would not have been the same without you. And everybody else who stopped in the hallway to say hi and talk for a bit. Thank you.
Contents

List of Tables vi

List of Figures vii

List of Acronyms ix

1 Introduction 1

1.1 Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

1.2 System Requirements and Specifications . . . . . . . . . . . . . . .. . 2

1.2.1 Neural Stimulation . . . . . . . . . . . . . . . . . . . . . . . . 3

1.2.2 Neural Recording . . . . . . . . . . . . . . . . . . . . . . . . . 5

1.2.3 Stimulation Artifact . . . . . . . . . . . . . . . . . . . . . . . . 6

1.3 Electrode-Tissue Interface . . . . . . . . . . . . . . . . . . . . . . . . 6

1.4 Bidirectional Integrated Neural Interfaces . . . . . . . . . . . . . . . . 8

1.5 Thesis Organization . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8

2 256-Channel Bidirectional Integrated Neural Interface 10

2.1 System Architecture . . . . . . . . . . . . . . . . . . . . . . . . . 10

2.2 Neural Stimulator . . . . . . . . . . . . . . . . . . . . . . . . . . . 12

2.2.1 Existing Current Driver Circuits . . . . . . . . . . . . . . . . . . . 13

2.2.2 Stimulator Architecture . . . . . . . . . . . . . . . . . . . . . . 14

2.2.3 Current Driver . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

2.2.4 Operational Transconductance Amplifier . . . . . . . . . . . . . . . . 18
2.2.5 Stimulator Experimental Characterization .......................... 20

2.3 Neural Recording ......................................................... 22
  2.3.1 Recording Channel Architecture .................................. 22
  2.3.2 Neural Amplifier ...................................................... 24
  2.3.3 Sample-and-Hold Circuit and ADC ................................. 32
  2.3.4 Recording Channel Characterization ............................... 35

2.4 Experimental Validation .................................................. 38

3 CMOS-Microelectrode Hybrid Integration ............................... 41
  3.1 Multi-bump Au Microelectrode ........................................ 42
    3.1.1 Layout Techniques ................................................. 43
    3.1.2 Microelectrode-Chip Integration ................................. 44
  3.2 CMOS-UEA Microelectrode Array Integration ....................... 46

4 Conclusions ........................................................................ 49
  4.1 Thesis Contributions ...................................................... 49
  4.2 Future Work ............................................................... 50

Appendix .............................................................................. 51

A Supplementary Hardware and Software Documentation .............. 51
  A.1 Board Design ............................................................. 51
  A.2 MATLAB Interface ....................................................... 53
  A.3 FPGA Programming ...................................................... 54

References .......................................................................... 56
# List of Tables

2.1 Stimulator switch network configuration .................................. 14
2.2 Stimulator OTA transistor sizing .......................................... 20
2.3 Comparative analysis of current mode neural stimulation arrays .... 22
2.4 Telescopic OTA transistor sizing ............................................. 26
2.5 Simulated telescopic OTA electrical characteristics ..................... 27
2.6 Folded cascode OTA transistor sizing ....................................... 27
2.7 Simulated folded cascode OTA electrical characteristics ............... 28
2.8 Comparative analysis of fully differential neural recording channels . 31
2.9 Comparator transistor sizing .................................................. 35
2.10 Recording channel experimental characteristics ......................... 36
2.11 System-level experimental characteristics .................................. 38
2.12 Comparative analysis of neural stimulation and recording systems . 40
List of Figures

1.1 A vision of an implantable adaptive neural stimulation system. . . . . . . 2
1.2 A block diagram of an adaptive neural stimulation microsystem with multiple bidirectional integrated neural interfaces. . . . . . . . . . . . . 3
1.3 Equivalent circuit model of the electrode-tissue interface. . . . . . . 7

2.1 Micrograph of the integrated neural stimulation and recording interface implemented in a standard 0.35\(\mu\)m double-poly CMOS technology. The die dimensions are 3.5mm x 3.65mm . . . . . . . . . . . . . . 11
2.2 A general implementation of biphasic stimulation. . . . . . . . . . . . . . 12
2.3 Common current driver designs. . . . . . . . . . . . . . . . . . . . . . 15
2.4 (a) Stimulator architecture and (b) timing diagram for implementing a biphasic stimulus. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
2.5 Current sink configured in (a) storage mode and (b) stimulation mode. . 18
2.6 Folded-cascode OTA in the stimulator current sink. . . . . . . . . . . . 19
2.7 Transfer characteristic of the current driver for 10k\(\Omega\) load. . . . . . . 21
2.8 Output of the current driver under variable load conditions. . . . . . . . 21
2.9 Neural recording channel architecture. . . . . . . . . . . . . . . . . . 23
2.10 Telescopic amplifier in the first stage of the recording channel. . . . . 25
2.11 Folded cascode amplifier in the second stage of the recording channel. . 28
2.12 Experimentally measured amplitude frequency response of the full channel with the gain set to minimum setting of 53dB. . . . . . . . . . . . . . 29
2.13 Experimentally measured noise of the full channel with the gain set to 54dB. .................................................. 30

2.14 In-channel single-slope ADC block diagram. ............................................. 32

2.15 Timing diagram of one sample period. ..................................................... 33

2.16 Comparator in the ADC. ................................................................. 34

2.17 Simulated comparator performance for worst case input. ......................... 34

2.18 Micrograph of a single recording channel. ............................................. 36

2.19 Measured output spectrum for the full channel including the ADC. ........ 37

2.20 Input referred output from the channel in response to a 200µVpp sinusoidal input. .......................................................... 37

2.21 (a) Original recordings of epileptic seizure spikes, and (b) the data recorded with this system. ............................................... 39

3.1 Cross section view of the stud-bump gold microelectrodes with a brain slice. ........................................................................ 42

3.2 (a) A single [1] and (b) a quad stud-bump gold microelectrodes. ......... 43

3.3 Cross section view of the metal and poly layers used in the layout of the electrode bonding pad. .................................................. 44

3.4 Micrograph of a neural recording interface implemented in a 0.35µm CMOS technology [2]. .............................................................. 45

3.5 Quad bump microelectrodes fabricated on the neural recording microchip surface. ................................................................. 45

3.6 Epileptic seizure neural activity recorded through the quad-bump electrodes by the neural amplifier chip shown in Figure 3.4. ............... 47

3.7 (a) The UEA. (b) Top view of the UEA attached to the die. (c) UEA on the packaged die with insulated bonding wires. (d) Top view of the UEA in the package with insulated bonding wires. .................. 48

A.1 The PCB designed for characterizing the adaptive neural stimulation interface. ................................................................. 52
List of Acronyms

ADC analog-to-digital converter
NEF noise efficiency factor
SC switched-capacitor
S/H sample-and-hold
EEG Electroencephalograph
PSNR peak signal-to-noise ratio
SEM scanning electron microscopy
OTA operational transconductance amplifier
UEA Utah electrode array
HPF high-pass filter
LPF low-pass filter
BPF band-pass filter
ACSF artificial cerebrospinal fluid
CPE constant phase element
$V_{DS}$ drain-to-source voltage
SAR  successive approximation

CMFB  common-mode feedback

CMRR  common-mode rejection ratio

PSRR  power supply rejection ratio

CDS  correlated double sampling

SNR  signal-to-noise ratio

VLSI  very-large-scale integration

DAC  digital-to-analog converter

OTA  operational transconductance amplifier

TIA  transimpedance amplifier

FOM  figure of merit

RMS  root mean square

IC  integrated circuit

HDL  hardware description language

CAL  common analog line

SOI  silicon on insulator
Chapter 1

Introduction

1.1 Motivation

Adaptive neural stimulators with closed loop feedback are a promising new technology in neurology and neural engineering. It has been shown that applying a stimulus in response to a neural event can be useful in treating neurological disorders such as epilepsy [3]. A system that acquires neural activity, processes the data and then, when it detects a certain neural behaviour, responds with a stimulus of a certain type, can be utilized in the treatment of such neurological disorders. Electrical stimulation is already being used for treatment of neural disorders such as Parkinson’s disease and depression [4, 5]. Monitoring the neural response of patients often requires them to be confined to a hospital bed. A minimally invasive implantable system that can perform both neural stimulation and real-time monitoring can make the therapy less burdensome.

Figure 1.1 shows the main components of the envisioned cortical adaptive neural stimulation system. The integrated neural interface modules are implanted in various locations of the cortex. They communicate with a signal processing and data commu-
1.2 System Requirements and Specifications

The high density of neural networks in biological tissue requires a large number of electrodes to obtain the most accurate representation of neural activity, and for better control over the location of the stimulation sites [6]. Several microelectrode technologies are currently available for use with integrated systems for neural recording and stimulation. The Utah electrode array (UEA) consists of an array of 1 to 1.5mm long silicon electrodes with a pitch of 400µm. Michigan probe [7] is another silicon based microelectrode assembly. Its long shanks incorporate multiple recording and stimula-
1.2. System Requirements and Specifications

1.2.1 Neural Stimulation

Stimulation arrays have been extensively researched and developed for use in cochlear and retinal prostheses [11–13]. Electrical stimulation of excitable neural tissue occurs whenever charge is delivered to the tissue through an electrode. Safety is a major concern in the design of implantable neural stimulators. Adverse chemical reactions such
as electrolysis, pH changes and tissue and electrode damage can occur whenever the
tissue is exposed to prolonged DC currents or anything else that results in charge ac-
cumulation in the tissue. It is therefore essential that the stimulator provides maximum
control over the charge delivered and retracted to and from the tissue, respectively. The
issue of charge balancing has been under extensive research and various schemes and
approaches have been used towards this goal as summarized in [14].

Neural stimulators usually employ one of three approaches to transfer charge to neu-
ral tissue. Constant current stimulation establishes a well controlled current between
pairs of electrodes for a short amount of time. Constant voltage stimulation establishes
current flow by controlling the voltage at the electrode site. And charge based stimula-
tion utilizes switched capacitor networks to deliver charge directly to the tissue.

Voltage controlled stimulation has the advantage of higher power efficiency [15].
The main drawback of voltage controlled stimulation however is the lack of control
over the charge delivered to the tissue. Tissue impedance varies from a nominal value
of 10kΩ due to cellular reactions [16]. Also, since the electrode tissue interface has
a capacitive component, the instantaneous current between electrodes is not well con-
trolled. Current controlled stimulation offers direct control over the charge delivered
to the tissue due to the linear relationship between charge and current. Currents in the
range of 10µA - 1mA are sufficient to invoke a neural response [17]. Charge controlled
stimulation is still in its early stages of development [18].

Voltage compliance of neural stimulators refers to the amount of voltage swing
(headroom) available at the output stage. Higher voltage compliance will allow the
current driver to supply wider range of currents to a given load. Maximizing the output
impedance will allow the current driver to maintain constant current under variable load
conditions.

It is desirable to have maximum control over the stimulation waveform. Stimula-
tors usually contain a digital-to-analog converter (DAC) to control the output current amplitude. Some designs use current mode DACs as the output stage of the stimulator [13,19,20], others use a voltage mode DAC and interface it with a voltage to current converter circuit [21]. For arbitrary waveform generation, which provides the highest functional flexibility, high resolution is required. The resolution is limited by the area allocated for the stimulator, since a DAC is commonly incorporated in every channel. A stimulator array that does not require a dedicated DAC at every output stage would reduce the area and the power consumption of a neural stimulation and recording system.

For maximum functional flexibility the stimulator array has to stimulate simultaneously at multiple electrode sites. This capability requires the incorporation of memory into every stimulator block. Stimulators that are integrated with DACs will need digital memory with an area that is proportional to the resolution of the DAC. Sharing a DAC among all the stimulators in the array requires an accurate method to store the current information in analog form.

### 1.2.2 Neural Recording

Extracellular neural action potentials vary in their amplitude approximately between 20µV and 500µV [22,23]. For example, epileptic neural activity is a result of synchronized firing of populations of neurons with an amplitude of several millivolts [24,25]. These signals occupy a frequency band of 0.1Hz - 5kHz. A neural recording channel has to detect and amplify these low amplitude signals, while keeping the input referred noise of the channel below the background noise at the recording site, which is typically 5-10µV [6]. Substrate noise due to the digital circuits on the same chip can degrade the quality of the recorded signal, requiring a fully differential architecture. DC offset and drift are a common problem at the electrode site and can get as large as 1V [6,22].
To be able to interface with a wireless data communication link and neural data processors the data has to be digitized. There are different approaches to incorporating an analog-to-digital converter (ADC) on a chip with a large number of recording channels. Single-ADC and column-parallel architectures don’t lend themselves well to scaling and require a redesign when the array size changes as their sampling rate depends on the number of channels. Designs with an ADC in each channel are scalable and easily adjust for various applications.

### 1.2.3 Stimulation Artifact

When stimulation and recording functions are combined in the same system the issue of stimulation artifact has to be addressed. The stimulation artifact refers to the temporary disturbance caused to the neural recording channel by the charge that accumulates on the electrode-tissue interface as a consequence of neural stimulation. During stimulation the tissue electrode interface is exposed to signals that are far larger (millivolts and higher) than the neural signals the recording channels were designed for (microvolts to millivolts). This can saturate the recording channels and create a delay between the application of stimulus and normal recording operation. The recording channel has to be able to return to normal functionality as fast as possible after stimulation.

### 1.3 Electrode-Tissue Interface

The tissue-electrode interface acts as a load during stimulation and a signal source during recording. Therefore the characterization of its impedance is of particular importance. The first-order-model for the tissue-electrode interface is shown in Figure 1.3 [26, 27]. $Z_{CPE}$ is a constant phase element that represents the interface capacitance of the electrode. It is given by
1.3. Electrode-Tissue Interface

Figure 1.3: Equivalent circuit model of the electrode-tissue interface.

\[ Z_{CPE}(w) = \frac{1}{(jwQ)^n}, \quad (1.1) \]

where \( Q \) is the measure of the magnitude of \( Z_{CPE} \), \( n \) is a constant between 0 and 1. \( Z_{CPE} \) is a purely capacitive impedance if \( n = 1 \). For commonly used stimulation microelectrodes the capacitance value is typically 0.16nF - 16nF. \( R_{ct} \) is the transfer resistance that is linearly related to the oxidation and reduction currents that flow across the electrode-tissue interface at equilibrium. Its value is typically in the order of 100MΩ - 10GΩ. The series resistance, \( R_S \), and the tissue resistance, \( R_{tissue} \), when lumped together are typically around 10kΩ. \( V_{sig} \) models the low amplitude neural signals. \( V_{os} \) models the DC offset at the electrode tissue interface, which can be as large as 1V [6, 22].

While this first-order-model is a good approximation for design purposes, real electrode-tissue interface is non-linear and time variant. For this reason the circuits directly interfacing with the electrodes should be able to tolerate some deviation from the nominal parameters [16].
1.4 Bidirectional Integrated Neural Interfaces

Several designs focusing on integrated neural stimulation and recording systems have been reported [28–31]. The design in [28] is a 128-channel integrated neural recording and stimulation interface implemented in a 0.6µm CMOS technology. The recording channel is fully differential, with a column-parallel ADC architecture. The power consumption for the recording channel is 787µW. The stimulator is a class-AB buffer consuming 150µW standby power. The area per channel is 0.33mm². The voltage-mode stimulators utilize a single on-chip DAC and require multiplexing to stimulate on multiple channels. The design in [29] has 8 recording channels and 64 stimulation channels implemented separately on the chip. Implemented in a 0.18µm CMOS technology it occupies an area of 4.48mm². The recording channels share a single ADC and consume a total power of 182µW. The current-mode stimulators incorporate a DAC in every channel and consume 89µW. The design in [30] is implemented in a 0.35µm CMOS technology, and includes 16 recording and stimulation channels, each occupying an area of 0.045mm². The recording channel is a single-ended amplifier/filter consuming 145µW. The output of the recording channel is analog. The stimulator is a buffer with analog input. The design in [31] is a 128-channel system with a fully differential channel with a column-parallel ADC, and a class-AB buffer as the stimulator. Implemented in a 0.35µm CMOS technology, it consumes 12.75µW in its recording channel and 51µW in the stimulation buffer. The voltage-mode stimulator incorporates memory and is capable of stimulating on all channels simultaneously.

1.5 Thesis Organization

This thesis presents a 0.35µm CMOS VLSI implementation of a bidirectional neural recording and stimulation microsystem with 64 neural stimulation circuits, and 256
fully-differential neural recording channels, each with an in-channel ADC. The stimulators are fully programmable with square or arbitrary waveforms. The memory in every stimulator allows for simultaneous stimulation on all active channels, which cannot be achieved with conventional multiplexing. The stimulators provide a wide range of functionality and occupy only $0.02 \text{mm}^2$ per channel. This is achieved by re-using the opamp in the stimulator in two different configurations, and sharing a single DAC (off-chip). These techniques also result in low quiescent power consumption of $2.76\mu\text{W}$ per stimulator. The recording channel is fully differential, occupying an area of $0.03 \text{mm}^2$ and consuming $51.9\mu\text{W}$. The sampling and quantization is performed simultaneously on all channels. This design targets hybrid integration with the UEA with custom Iridium tips, as needed for neural stimulation with lower electrode interface impedance.

The remainder of this thesis is organized as follows:

- Chapter 2 discusses the system architecture and very-large-scale integration (VLSI) circuit implementation of the bidirectional integrated neural interface for adaptive cortical stimulation. The experimental results obtained from a $0.35\mu\text{m}$ standard CMOS prototype of the neural interface are also presented.

- Chapter 3 presents a low-cost technique for fabricating microelectrode arrays on the surface of the chip. *In vitro* experimental measurements obtained through these electrodes with an existing neural interface are also shown.

- Chapter 4 concludes the thesis and outlines future research directions.
Chapter 2

256-Channel Bidirectional Integrated Neural Interface

This chapter presents the architecture, VLSI circuit implementation and experimental results of the bidirectional integrated neural interface.

2.1 System Architecture

The neural recording and stimulation microsystem was fabricated in a standard 0.35\(\mu\)m double-poly CMOS technology. Figure 3.4 shows the micrograph of the die. The die dimensions are 3.5mm x 3.65mm. The microsystem consists of an array of 8\(\times\)8 neural stimulation circuits, and 16\(\times\)16 fully differential neural recording channels with a single-slope ADC in each channel.

Top metal electrode bonding pads are arranged in a 16\(\times\)16 array with a 200\(\mu\)m pitch. Each electrode is interfaced with a recording channel. One of every four electrodes is connected to a programmable current-mode stimulator. This results in a 400\(\mu\)m pitch of stimulation electrodes. This pitch is chosen to match that of the UEA for direct on-chip microelectrode hybrid integration.
Figure 2.1: Micrograph of the integrated neural stimulation and recording interface implemented in a standard 0.35\( \mu \)m double-poly CMOS technology. The die dimensions are 3.5mm x 3.65mm
2.2 Neural Stimulator

In this design current controlled stimulation was implemented due to the control it provides over the charge delivered to the tissue. The stimulators are individually addressable and any subset of them can be disabled or enabled. The enabled stimulators are capable of truly simultaneous stimulation without multiplexing. The stimulator output stage is designed to supply biphasic stimulation current to the tissue with impedance $Z_L$ as illustrated in Figure 2.2.

Each electrode site that is connected to a stimulator can be connected either to the current driver or the supply. During the anodic phase, $\phi_1$ is high and the current is flowing from site $A$ to site $B$. During the cathodic phase $\phi_2$ is high and current is flowing from site $B$ to site $A$. The amplitude and duration of the current pulses during the two phases need not be the same. As long as the area under the two waveforms is the same the charge that was delivered to the tissue is retracted from the tissue. The stimulator is fully programmable and the amplitude and the duration of each pulse can

Figure 2.2: A general implementation of biphasic stimulation.
be set independently.

Precise current control is essential in the current driver design. High accuracy and linearity are important design considerations. In addition, due to the variable nature of the load impedance the output impedance of the current driver has to be maximized in order to deliver constant current under a wide range of conditions.

In addition to stimulating with a precisely controlled biphasic waveform other steps are taken to ensure charge balancing. To prevent DC currents, all stimulating sites that are not in use are configured as a high impedance node. During stimulation all sites can be connected to a common analog line (CAL) in order to redistribute any charge that has accumulated at the electrode-tissue interface.

### 2.2.1 Existing Current Driver Circuits

There are several approaches to setting the output current of a stimulation current driver. One way is shown in Figure 2.3(a) [29]. A current mode DAC is used to set the current which is then copied through a current mirror to the output. This design suffers from high power consumption, which can be reduced by moving the DAC to the output path, as shown in Figure 2.3(b) [12]. The output impedance of this topology is of the order of $r_o$ of the output transistor.

It is desirable to increase the output impedance of the current driver to reduce the sensitivity of the output current to load impedance. Introducing another transistor to the output path with active feedback significantly increases the output impedance of the current driver. Figure 2.3(c) demonstrates this technique. The output current is set by the voltage-mode DAC that drives the gate of transistor $M_1$. If $V_{ref}$ is sufficiently small transistor $M_1$ will be biased in triode and used as a voltage controlled resistor [21]. The output impedance is now increased significantly. The linearity of the transistor in triode region is compromised by the effect of mobility degradation, making the drain cur-
rent not linear with the overdrive voltage. The compensation circuit introduced in [21] increases the power consumption of the stimulator. One way to prevent the linearity problem is to reintroduce the current mode DAC to the output current path, as shown in Figure 2.3(d) [19].

None of the designs described above incorporate memory in the stimulator. This means that truly simultaneous stimulation cannot be performed on several channels. Another disadvantage is the necessity to include a DAC with every stimulator circuit, which increases the area of the stimulator.

### 2.2.2 Stimulator Architecture

The choice of the stimulator architecture was dictated by the low area, low power and high output impedance design constraints. The programmable stimulator consists of three major blocks: digital memory, digital control and a current driver. This is illustrated in Figure 2.4(a).

The code \((m_0, m_1)\) stored in the digital SRAM memory acts as the input to the control logic block and along with the \(\text{Clock}\) and \(\text{Phase}\) inputs configures the switch network to put each site into its desired operation mode. The different states and the corresponding switch configurations are described in Table 2.1.

The stimulator operates in two phases: setup phase and active phase. This is illustrated in Figure 2.4(b) for two stimulation sites, \(A\) and \(B\). During the setup phase the

<table>
<thead>
<tr>
<th>State</th>
<th>ON Switches</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Impedance</td>
<td>-</td>
<td>Setup/Active</td>
</tr>
<tr>
<td>CAL</td>
<td>(S_5)</td>
<td>Active</td>
</tr>
<tr>
<td>VDD</td>
<td>(S_4)</td>
<td>Active</td>
</tr>
<tr>
<td>Current Storage</td>
<td>(S_1, S_3)</td>
<td>Setup</td>
</tr>
<tr>
<td>Current Delivery</td>
<td>(S_2, S_3)</td>
<td>Active</td>
</tr>
</tbody>
</table>
Figure 2.3: Common current driver designs.
2.2. NEURAL STIMULATOR

![Diagram of digital memory, digital control, and current driver](image)

(a)

(b)

Figure 2.4: (a) Stimulator architecture and (b) timing diagram for implementing a biphasic stimulus.
sites are sequentially addressed and the digital memory is programmed with a 2-bit code which represents the state the output stage of the stimulator will be in during the active phase. Incorporation of memory in the stimulator ensures truly simultaneous operation and prevents unwanted current flow due to timing mismatches.

### 2.2.3 Current Driver

The current driver presented here achieves the task of storing the stimulation current and stimulating through the same path, without requiring a DAC in every stimulator circuit. This provides an advantage in terms of accurate current copying, small area and low power consumption.

The main components of the current driver are an operational transconductance amplifier (OTA), and two transistors \( M_1 \) (4×5.5\( \mu \)m/0.5\( \mu \)m) and \( M_2 \) (4×8\( \mu \)m/0.5\( \mu \)m). These components are configured with switches into two different circuits for the purposes of current storage (during setup phase) and current driving (during active phase). The two circuits can be seen in Figure 2.5.

Figure 2.5(a) shows the current driver configured for current storage. In this state the circuit is a constant \( V_{DS} \) current copier [32]. The desired current is flowing through transistor \( M_1 \). The drain voltage of transistor \( M_1 \) is held constant at the potential of \( V_{in} \) through the OTA feedback, which compensates for the effect of channel length modulation. The gate voltage of \( M_1 \) is driven by the OTA to the proper level and is stored in the capacitor \( C_{mem} \).

Figure 2.5(b) shows the current driver during the active phase, configured as a typical current sink. The negative feedback around \( M_2 \) serves two purposes. First it boosts the output impedance of the current sink. This is a well established gain boosting technique that by increasing the output impedance of the current driver makes it less sensitive to the load impedance [19, 21]. Second, it forces the drain voltage of \( M_1 \) to the
potential $V_{in}$. This together with the gate voltage that is set by $C_{mem}$ ensures that the current that flows through it is as close as possible to the current during the storage phase.

Capacitor $C_c$ is a compensation capacitor and it is used to stabilize the node $V_d$ during the switching from one configuration to the other.

The incorporation of analog memory in each current driver has two benefits. First, area is saved because the DAC can be placed outside of the cell and shared by all the channels. Second, stimulation can be performed on all enabled sites simultaneously.

### 2.2.4 Operational Transconductance Amplifier

To maximize the output impedance of the current sink the OTA needs to have large gain as seen in
where $A$ is the gain of the OTA, $g_{m2}$ is the transconductance of transistor $M_2$, $r_{o2}$ is the output impedance of transistor $M_2$ and $r_{o1}$ is the output resistance of transistor $M_1$.

The OTA has a PMOS input folded-cascode configuration, as shown in Figure 2.6. It provides a gain of 93dB at a common mode input of $V_{in} = 300mV$. The OTA consumes a bias current of $0.88\mu A$. The transistor sizes are given in Table 2.2.
Table 2.2: Stimulator OTA transistor sizing

<table>
<thead>
<tr>
<th>Transistor</th>
<th>W/L (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{1,2}$</td>
<td>2/4</td>
</tr>
<tr>
<td>$M_{3,4}$</td>
<td>1/5</td>
</tr>
<tr>
<td>$M_{5,6}$</td>
<td>1/6</td>
</tr>
<tr>
<td>$M_{7,8}$</td>
<td>1/4</td>
</tr>
<tr>
<td>$M_{9,10}$</td>
<td>1/4</td>
</tr>
<tr>
<td>$M_p$</td>
<td>1/5</td>
</tr>
</tbody>
</table>

### 2.2.5 Stimulator Experimental Characterization

The stimulator was experimentally characterized by loading the stimulation current through an external resistor. Figure 2.7 compares the measured and simulated input-output characteristics of the current driver. The output current is lower than the input current due to the effects of charge injection at the gate of transistor $M_1$. Due to the systematic nature of this error the mismatch can be compensated for in software.

Figure 2.8 shows the simulated and measured performance of the current driver under variable load conditions. The plot clearly shows the operating region of the current driver. The voltage at the electrode site can be as low as 700mV, resulting in a voltage compliance of 2.6V if the supply voltage is set to 3.3V. Currents above 250µA will drive the circuit out of the linear region even for the nominal load of 10kΩ.

In Table 2.3 this design is compared with other current mode stimulators published in [19], [33], [29], and [34]. Previously published multi-channel stimulators all require multiplexing and cannot stimulate on many channels simultaneously. The larger currents up to 1mA are achieved by having a power supply higher than that in [29]. The designs in [19, 34] also utilize active feedback in order to maximize the output impedance. This design has the lowest reported area.
2.2. Neural Stimulator

Figure 2.7: Transfer characteristic of the current driver for 10kΩ load.

Figure 2.8: Output of the current driver under variable load conditions.
### Table 2.3: Comparative analysis of current mode neural stimulation arrays

<table>
<thead>
<tr>
<th>Technology [µm]</th>
<th>Channels</th>
<th>Channel Area [mm²]</th>
<th>DAC</th>
<th>Supply Voltage [V]</th>
<th>Output Current ($Z_L$=10kΩ) [µA]</th>
<th>Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>[19] 1</td>
<td>4</td>
<td>0.38</td>
<td>per-channel</td>
<td>6</td>
<td>200-1000</td>
<td>No</td>
</tr>
<tr>
<td>[33] 1.5</td>
<td>8</td>
<td>0.6</td>
<td>shared</td>
<td>±6</td>
<td>0-600</td>
<td>No</td>
</tr>
<tr>
<td>[29] 0.18</td>
<td>64</td>
<td>-</td>
<td>per-channel</td>
<td>1.8</td>
<td>0-10</td>
<td>No</td>
</tr>
<tr>
<td>[34] 1.5</td>
<td>32</td>
<td>0.1</td>
<td>per-channel</td>
<td>5</td>
<td>0-270</td>
<td>No</td>
</tr>
<tr>
<td>This work 0.35</td>
<td>64</td>
<td>0.02</td>
<td>shared</td>
<td>3.3</td>
<td>20-250</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### 2.3 Neural Recording

This design includes an array of fully differential digital neural recording channels in order to adapt stimulation. Each recording channel is interfaced with a microelectrode site. The non-inverting input of the recording channel is connected to the local electrode. The inverting input of the recording channel has an option to either be connected to an adjacent electrode or a common reference electrode, which can be off-chip. These two configurations allow for two types of neural recordings: a localized differential recording, and a global recording obtaining signals from a wider spatial area with respect to a common reference.

#### 2.3.1 Recording Channel Architecture

The architecture of the recording channel is depicted in Figure 2.9. Each recording channel consists of two stages of filtering and amplification, a sample-and-hold circuit and a single slope ADC.
2.3. NEURAL RECORDING

Figure 2.9: Neural recording channel architecture.

The channel is fully differential to reduce the common mode noise that results from having both digital and analog circuits on the same substrate. Distributing the overall gain over two stages helps achieve large gain while maintaining smaller capacitor sizes and higher linearity.

The differential signal transfer function of a single stage is

\[ T(s) = \frac{s \frac{C_{\text{in}}}{C_f}}{s + \frac{1}{R_f C_f}}, \]

(2.2)

where \( R_f \) refers to the resistance of the feedback. The absolute sizes of the capacitors were limited by the available area, whereas the relative sizes are chosen to meet the channel specifications.

The closed loop gain is determined by the product \( \frac{C_{\text{in}}}{C_f} \times \frac{C'_{\text{in}}}{C'_f} \). In the first stage \( C_{\text{in}} = 4\text{pF} \) and \( C_f = 100\text{fF} \) which results in a gain of 32dB. In the second stage a variable gain amplifier is implemented by using a programmable bank of capacitors in the feedback. The input capacitor is \( C'_{\text{in}} = 2.5\text{pF} \). The capacitance in the feedback can be
programmed from 25fF to 200fF. This results in eight gain modes from 22dB to 40dB. The overall gain of the two stages is therefore adjustable from 54dB to 72dB. The input capacitors also play a role in DC offset rejection at the input terminals.

The high pass corner frequency is determined by \( \frac{1}{2\pi R_f C_f} \), where \( R_f \) is the feedback resistance. Since \( C_f \) had to be small to ensure large gain, \( R_f \) has to be maximized to set the high pass corner frequency as low as possible. A good way to implement large resistance in CMOS technology without sacrificing area is to use a MOS device in subthreshold region [35]. In this design four PMOS transistors are combined in series to maximize the resistance. The HP cut-off frequency is tunable from 0.1Hz to 10Hz by adjusting the control voltage \( V_{res} \). The first stage low pass cut-off frequency is set to 5kHz by the combined effect of the load capacitor \( C_{load} \) and the bias current of the amplifier. By changing the bias current of the amplifier it can be set from 5kHz to 10kHz.

Periodic reset may be necessary every several minutes to compensate for DC drift caused by junction leakage [28]. Resetting the feedback resistors is also used in quickening the recovery from saturation due to the stimulation artifact. By setting \( V_{res}=0V \) the amplifiers are configured in buffer mode which allows for faster recovery from saturation.

### 2.3.2 Neural Amplifier

The main considerations in the design of the neural amplifier are noise and power consumption, one coming at the expense of the other. Various OTA topologies have been used for the implementation of neural amplifiers: folded cascode, current mirror and two stage [27, 28, 36, 37]. The common factor with these topologies is the large output swing. Large output swing is important at the output stage of the amplifier, but comes at a cost of higher power consumption due to the large number of current branches.
Since in this design the gain is distributed over two stages a wide swing topology for the first stage OTA is not needed. The telescopic OTA saves power in the first stage while maintaining low input referred noise level.

The fully differential telescopic OTA of the first stage and the common-mode feedback (CMFB) circuit are shown in Figure 2.10. Transistor sizes in this amplifier are listed in Table 2.4.

Two types of noise have to be accounted for in the design of the telescopic OTA. Both can be modeled as voltage sources in series with the input. The input referred thermal noise per unit bandwidth is

\[
\overline{V_n^2} = 2 \left[ \frac{4kT}{g_{m1}} \left( \frac{2}{3} \right) \left( 1 + \frac{g_{m7}}{g_{m1}} \right) \right]. \tag{2.3}
\]
Table 2.4: Telescopic OTA transistor sizing

<table>
<thead>
<tr>
<th>Transistor</th>
<th>W/L (µm)</th>
<th>( g_{m}(\mu A/V) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( M_{1,2} )</td>
<td>( 4 \times 5.35/7 )</td>
<td>9.97</td>
</tr>
<tr>
<td>( M_{3,4} )</td>
<td>( 2 \times 3/12 )</td>
<td>5.78</td>
</tr>
<tr>
<td>( M_{5,6} )</td>
<td>1.5/10</td>
<td>5.36</td>
</tr>
<tr>
<td>( M_{7,8} )</td>
<td>1/20</td>
<td>2.98</td>
</tr>
<tr>
<td>( M_p )</td>
<td>2/6</td>
<td>6.65</td>
</tr>
<tr>
<td>( M_{9,10} )</td>
<td>1/3.25</td>
<td>2.38</td>
</tr>
<tr>
<td>( M_{11,12,13,14} )</td>
<td>1/20</td>
<td>1.22</td>
</tr>
<tr>
<td>( M_{15,16} )</td>
<td>1/20</td>
<td>3.01</td>
</tr>
</tbody>
</table>

To minimize the thermal noise contribution, transistors \( M_{1,2,7,8} \) have to be biased such that \( g_{m7,8} \ll g_{m1,2} \). For a fixed bias current that means that the sizes should be such that \( W/L_{7,8} \ll W/L_{1,2} \), which puts \( M_{1,2} \) in weak inversion and \( M_{7,8} \) in strong inversion. Increasing the bias current will further increase \( g_{m1,2} \) resulting in additional reduction of the thermal noise component at the expense of higher power dissipation.

The input referred flicker \( 1/f \) noise per unit bandwidth is

\[
\overline{V_n^2} = 2 \left[ \frac{K_P}{C_{ox}W_1L_1} + \frac{K_N}{C_{ox}W_7L_7} \left( \frac{g_{m7}}{g_{m1}} \right)^2 \right] \frac{1}{f} .
\] (2.4)

The major contribution of \( 1/f \) noise comes from transistors \( M_{1,2} \). Reduction in \( 1/f \) noise is achieved by choosing the input transistors to be of the PMOS type which have a lower \( 1/f \) noise coefficient than NMOS transistors \( (K_P < K_N) \), and maximizing the gate area of these transistors.

The CMFB circuit [38] is designed with the same bias current as the amplifier. The output swing of the CMFB circuit is maximized by increasing the overdrive voltage of the input transistors \( M_{11-14} \). For a given bias current this is achieved by sizing \( M_{11-14} \) to have the minimal \( W/L \) ratio. The simulated characteristics of the first stage OTA are outlined in Table 2.5.
In the design of the second stage OTA a higher noise level can be tolerated due to signal amplification in the first gain stage. The focus is on minimizing power dissipation and having adequate output swing. A folded cascode OTA was chosen for the second stage with a total bias current of $1.5\mu A$, which is almost half of the first stage bias current. The CMFB circuit in the folded cascode OTA is the same as the one in the telescopic OTA. The schematic of the folded cascode OTA and the CMFB circuit is shown in Figure 2.11. The sizes of the transistors in this implementation are listed in Table 2.6. The simulated characteristics of the folded cascode OTA are presented in Table 2.7.

The experimentally measured amplitude frequency response of the two filtering and amplification stages is shown in Figure 2.12. The gain was set to the minimum and the high pass corner frequency was adjusted by changing the bias voltage of the feedback network.
2.3. NEURAL RECORDING

Figure 2.11: Folded cascode amplifier in the second stage of the recording channel.

Table 2.7: Simulated folded cascode OTA electrical characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC Gain</td>
<td>91dB</td>
</tr>
<tr>
<td>Unity Gain Frequency (250fF)</td>
<td>1.2MHz</td>
</tr>
<tr>
<td>Output Voltage Swing</td>
<td>$1.7V_{pp}$</td>
</tr>
<tr>
<td>Total Bias Current (with CMFB)</td>
<td>$1.5\mu A$</td>
</tr>
<tr>
<td>Supply Voltage</td>
<td>3.0V</td>
</tr>
</tbody>
</table>

resistors.

The plot of the experimentally measured input-referred noise of the two stages is shown in Figure 2.13. Integrating the noise over the bandwidth of 10Hz-5kHz results in the overall input-referred noise of $7.99\mu V_{rms}$.

Table 2.8 compares the recording channel presented here to the neural recording channels published in [37], [28], and [31]. The design in [31] reports the lowest noise efficiency factor (NEF). It however only characterizes the first stage of the recording channel which results in lower gain. The low input-referred noise reported in [28] comes at the expense of increased power consumption.
Figure 2.12: Experimentally measured amplitude frequency response of the full channel with the gain set to minimum setting of 53dB.
Figure 2.13: Experimentally measured noise of the full channel with the gain set to 54dB.
Table 2.8: Comparative analysis of fully differential neural recording channels

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[37] Two-stage</td>
<td>0.5</td>
<td>7.8</td>
<td>0.1 - 10k</td>
<td>114.8</td>
<td>3</td>
<td>18.7</td>
<td>40</td>
<td>0.107</td>
</tr>
<tr>
<td>[28] Folded cascode</td>
<td>0.6</td>
<td>5.9</td>
<td>10 - 100k</td>
<td>160</td>
<td>5</td>
<td>8.1</td>
<td>20</td>
<td>0.062</td>
</tr>
<tr>
<td>[31] Telescopic</td>
<td>0.35</td>
<td>6.08</td>
<td>10 - 5k</td>
<td>8.4</td>
<td>3</td>
<td>5.55</td>
<td>33</td>
<td>0.02</td>
</tr>
<tr>
<td>This work</td>
<td>0.35</td>
<td>7.99</td>
<td>10 - 5k</td>
<td>12.9</td>
<td>3</td>
<td>8.9</td>
<td>53</td>
<td>0.03</td>
</tr>
</tbody>
</table>
2.3. Sample-and-Hold Circuit and ADC

The architecture of the in-cell fully differential single-slope ADC is shown in Figure 2.14. Considering the amplitude of the neural signals and the background noise at the recording site, an ADC with 8-bit resolution is sufficient for this application [22]. The quantization of the recorded sample is performed simultaneously on all channels during the read phase. This allows for rapid parallel evaluation of the entire frame.

The in-channel ADC consists of a fully differential comparator and an 8-bit SRAM block. The differential ramp for the comparator and the 8-bit data is generated off-chip by a ramp generator and a counter. Both the ramp generator and the counter are common to all ADCs as quantization is taking place simultaneously on all channels. In future implementations the ramp generator and the counter can be implemented on chip as one of the peripheral circuits without significant area and power overhead if used for a large number of channels.

To properly quantize the neural data one needs to sample at a Nyquist rate of twice the low-pass cut-off frequency with an additional margin to allow for the roll-off of the filter. This results in a sampling rate of 15kHz per channel, or 65µs sample time. All channels are sampled simultaneously with the sample being stored in the analog memory as described in [39]. This takes 32µs. Then the quantization begins on all
channels simultaneously and takes $20 \mu s$, which is the time it takes the counter to count to 256 with a clock of 12.5MHz. The data is then read out by addressing each channel sequentially with 4-bit row/column-select signals passed through decoders. The readout rate is 20MHz, limited by the instrumentation data rate. It takes $13 \mu s$ to read out all 256 channels. This is represented in Figure 2.15.

The comparator [40] is designed to consume $13 \mu A$ from a 3.0V supply. The ADC (comparator and SRAM) occupies an area of $0.00725 \text{mm}^2$. The schematic of the comparator in the ADC is shown in Figure 2.16. The transistor sizes of the comparator are in Table 2.9.

Figure 2.17 shows the simulated output of the comparator for two inputs close to the common mode level. When the differential input is one LSB below common mode (-5mV) the comparator output is zero. When it is one LSB above the common mode level (5mV), it takes the comparator output $70 \mu s$ to rise to VDD. The comparator was clocked at the maximum rate of 12.5MHz.
2.3. NEURAL RECORDING

Figure 2.16: Comparator in the ADC.

Figure 2.17: Simulated comparator performance for worst case input.
2.3.4 Recording Channel Characterization

The micrograph of the recording channel is shown in Figure 2.18. The majority of the area is occupied by the input capacitors of both the first and the second stages.

The full channel was characterized experimentally. A 50Hz sine wave was applied differentially to the input of the first stage and passed through the two stages of amplification and the ADC. The spectrum of the output signal is shown in Figure 2.19. A summary of the experimental characteristics of the full recording channel including the ADC is given in Table 2.10.

The channel was tested for low amplitude inputs. The output of the channel to a 200µV sinusoidal input with a frequency of 30Hz is shown in Figure 2.20. The signal was low-pass filtered in software to eliminate the high frequency noise.

The functionality of the recording channel was also verified with pre-recorded neural activity of epileptic seizure events. Seizure-like events were induced in an intact hippocampus from Wilstar rats (5 - 25 days old) by immersing it in a low-Mg2+ solution. During a seizure neural populations synchronize and fire at the same time, which results in spikes of large amplitudes. The signals were recorded with a conventional bench-top low-noise amplifier and then programmed onto a HP33120A arbitrary waveform generator and applied to the non-inverting terminal of one recording channel. The

<table>
<thead>
<tr>
<th>Transistor</th>
<th>W/L (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1,2,3,4</td>
<td>2 × 2/1</td>
</tr>
<tr>
<td>M5,6</td>
<td>1/2</td>
</tr>
<tr>
<td>M7,8</td>
<td>0.8/0.35</td>
</tr>
<tr>
<td>M9,10</td>
<td>2 × 1/1</td>
</tr>
<tr>
<td>M11,12</td>
<td>1/0.35</td>
</tr>
<tr>
<td>M13,14</td>
<td>1/4</td>
</tr>
<tr>
<td>M15,16</td>
<td>1/8</td>
</tr>
<tr>
<td>M17,18</td>
<td>1/0.35</td>
</tr>
<tr>
<td>M19</td>
<td>4 × 1/0.35</td>
</tr>
</tbody>
</table>
Figure 2.18: Micrograph of a single recording channel.

Table 2.10: Recording channel experimental characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural Amplifier</td>
<td></td>
</tr>
<tr>
<td>Programmable Gain</td>
<td>53-72 dB</td>
</tr>
<tr>
<td>Low Frequency Cut-off</td>
<td>0.5-50Hz</td>
</tr>
<tr>
<td>High Frequency Cut-off</td>
<td>500Hz-10kHz</td>
</tr>
<tr>
<td>Input-Reflected Noise (10Hz-5kHz)</td>
<td>7.99 µV</td>
</tr>
<tr>
<td>NEF</td>
<td>8.9</td>
</tr>
<tr>
<td>CMRR (200Hz)</td>
<td>60 dB</td>
</tr>
<tr>
<td>THD (1mVpp)</td>
<td>0.8%</td>
</tr>
<tr>
<td>Full Channel (including ADC)</td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td>8-bit</td>
</tr>
<tr>
<td>Power Dissipation</td>
<td>52 µW</td>
</tr>
<tr>
<td>SNDR</td>
<td>32.18 dB</td>
</tr>
<tr>
<td>ENOB</td>
<td>5.05</td>
</tr>
<tr>
<td>SFDR</td>
<td>38 dB</td>
</tr>
</tbody>
</table>
Figure 2.19: Measured output spectrum for the full channel including the ADC.

Figure 2.20: Input referred output from the channel in response to a $200\mu V_{pp}$ sinusoidal input.
inverting terminal was grounded. The original pre-recorded signal and the signal acquired with the recording channel are depicted in Figure 2.21.

### 2.4 Experimental Validation

The experimental results summary for the full system are presented in Table 2.11.

Table 2.12 compares the neural stimulation and recording system presented here to other bidirectional integrated neural interfaces published in [29], [28], [30], and [31]. The design in [30] does not have an on-chip ADC. In [29] the low overall power consumption is in part due to the low number of recording channels, which only require a single ADC. This design incorporates the largest number of channels while consuming relatively low power and area.
Figure 2.21: (a) Original recordings of epileptic seizure spikes, and (b) the data recorded with this system.
### Table 2.12: Comparative analysis of neural stimulation and recording systems

<table>
<thead>
<tr>
<th></th>
<th>[28]</th>
<th>[29]</th>
<th>[30]</th>
<th>[31]</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year</strong></td>
<td>2006</td>
<td>2008</td>
<td>2008</td>
<td>2009</td>
<td>2009</td>
</tr>
<tr>
<td><strong>Technology [(\mu m)]</strong></td>
<td>0.6</td>
<td>0.18</td>
<td>0.35</td>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>Area ([mm^2])</strong></td>
<td>6.5×6.5</td>
<td>1.8×1.5</td>
<td>0.3×2.4</td>
<td>3.4×2.5</td>
<td>3.5×3.65</td>
</tr>
</tbody>
</table>

**Stimulation**
- **No. of channels**: 128, 64, 16, 128, 64
- **Mode**: Voltage, Current, Current/Voltage, Voltage, Current
- **Output current**: up to 10mA, 0-10µA, 0-100µA, -, 20-250µA
- **Power/Ch. \([\mu W]\)**: 150, 7.4, -, 51, 2.6

**Recording**
- **No. of channels**: 128, 8, 16, 128, 256
- **Fully differential**: Yes, No, No, Yes, Yes
- **ADC Architecture**: Column parallel, Single, -, Column parallel, In-channel
- **Power/Ch. \([\mu W]\)**: 780, 21.25, -, 12.75, 52
- **NEF**: 8.1, -, 17.6 - 46.1, 5.6, 8.9
- **Parallel operation**: No, No, No, Yes, Yes
- **On-chip 3D electrodes**: No, No, No, Yes, Yes
Chapter 3

CMOS-Microelectrode Hybrid Integration

Integrated neural interfaces are receiving wide attention in the integrated circuits research community. Microchips have recently been reported with one or several of the following capabilities: neural recording, neural signal processing, neural stimulation and wireless communication and power for implantability [2, 12, 28, 31, 41]. Many of the designed systems have a planar channel arrangement effectively constituting two-dimensional electronic circuit arrays. As off-chip multiplexing is prohibitive and noise and interference requirements are stringent, for in vitro and in vivo animal studies such electronic microchips have to be integrated directly with microelectrode arrays.

This chapter presents two approaches for CMOS-microelectrode integration. Section 3.1 presents a method of fabricating microelectrode arrays for in vitro neural tissue electrophysiological studies that is low-cost and has a fast turn around time. Building on the approach taken in [1], it utilizes standard wire bonding tools to create arrays of stacked gold stud-bump electrodes. A cross section of the resulting structure is shown in Figure 3.1.
Figure 3.1: Cross section view of the stud-bump gold microelectrodes with a brain slice.

Section 3.2 shows a second option for CMOS-microelectrode integration that utilizes a commercially available UEA. The UEA was attached to the bidirectional neural interface presented in Chapter 2.

3.1 Multi-bump Au Microelectrode

Gold stud bumping is performed by utilizing standard wire bonding technology. When a die bonding pads are bonded to the package the wire bonding tool presses a melted gold ball against a bonding pad on the die and then stretches the wire to connect it to the package. For gold stud bumping the wire bonding tool snaps the wire after pressing the gold ball against the electrode pad on the surface of the die. This is done for all electrode pads that serve as inputs to the neural amplifiers and results in an array of single stud bump electrodes as shown in Figure 3.2(a).

To create multi-stack bump electrodes the wire bonding tool coins all golden studs and repeats this process several times as required. An electrode made by stacking four gold bumps is shown in Figure 3.2(b). The resulting electrodes are about 180µm tall. By varying the number of golden studs different penetration depths can be achieved.
3.1 Multi-bump Au Microelectrode

This results in three-dimensional access to the live tissue. These microelectrodes were manufactured by a third party packaging service provider.

3.1.1 Layout Techniques

During the process of pressing the gold stud onto the electrode pad a significant amount of pressure is applied to the die. If the layout of the recording channel is not carefully considered the circuits under the bonding pad can be damaged and become dysfunctional.

To conserve the integrity of the recording channel after stud bumping placement of active devices under the electrode bonding pad should be avoided. Since a significant fraction of the channel layout area is consumed by the double-poly capacitors, placing them directly under the bonding pad will minimize the mechanical stress on the active devices in the channel. The capacitors are then connected to a stack of metals through arrays of vias. The dimension of the metals will depend on the signal routing over the
channel. Cut-outs have to be made in the support metals in order to avoid contact with the signal routing that uses the same metal layer. This layout technique is illustrated in Figure 3.3.

### 3.1.2 Microelectrode-Chip Integration

An array of quad stud-bump electrodes was fabricated on the surface of a 256-channel neural amplifier chip first reported in [2]. The micrograph of the chip is shown in Figure 3.4. The bonding pads have a 200µm pitch. The scanning electron microscopy (SEM) images of the quad stud-bump electrode array fabricated directly on the surface of the CMOS chip are shown in Figures 3.5(a) and 3.5(b).

The integrated neural interface consisting of the multi-channel neural amplifier microchip and the gold electrodes was used to record epileptic neural activity in an intact mouse hippocampus. The hippocampus was placed on the surface of the packaged
Figure 3.4: Micrograph of a neural recording interface implemented in a 0.35μm CMOS technology [2].

Figure 3.5: Quad bump microelectrodes fabricated on the neural recording microchip surface.
microchip inside a fluidic chamber specifically designed for this purpose. It was submerged in a heated artificial cerebrospinal fluid (ACSF) to keep it alive and perfused with a low-Mg2+ solution to stimulate epileptic activity. The neural spikes recorded through the electrodes on one of the channels are shown in Figure 3.6(a). A zoomed-in version of the middle spike is shown in Figure 3.6(b). The recorded neural activity is typical for this epilepsy animal model.

3.2 CMOS-UEA Microelectrode Array Integration

For this project the bidirectional neural interface was integrated with the UEA. Figure 3.7(a) shows the image of the array. Figure 3.7(b) shows the top view of the array on the die. The bonding pads can be seen on the top and bottom of the structure. The packaged die with the UEA and the insulated bonding wires is shown in Figure 3.7(c). Figure 3.7(d) shows the top view of the UEA on the die with the insulated bonding wires.

The size of the array is $8 \times 8$ with $400\mu m$ pitch. The electrodes are made of Iridium for biocompatibility and lower interface impedance. The testing of the bidirectional neural interface with the electrode array is included in the future work section.
Figure 3.6: Epileptic seizure neural activity recorded through the quad-bump electrodes by the neural amplifier chip shown in Figure 3.4.
3.2. CMOS-UEA Microelectrode Array Integration

![Figure 3.7](image)

**Figure 3.7:** (a) The UEA. (b) Top view of the UEA attached to the die. (c) UEA on the packaged die with insulated bonding wires. (d) Top view of the UEA in the package with insulated bonding wires.
Chapter 4

Conclusions

4.1 Thesis Contributions

We presented a bidirectional integrated neural interface with 64 neural stimulation circuits, and 256 fully-differential neural recording channels, each with an in-channel ADC. The interface was implemented in a 0.35μm CMOS technology. The fully programmable stimulators include memory which allows for truly simultaneous stimulation without multiplexing. Re-use of the opamp in the stimulator for two different configurations, and elimination of the DAC from every stimulator channel result in a compact design. The recording channels perform simultaneous sampling and quantization due to the incorporation of an ADC in every cell. Low noise design techniques are used for the design of the neural amplifier which results in a NEF of 8.9. The total power dissipation of the neural interface is 13.3mW.

The validity of planar neural interfaces has been verified by obtaining neural recording through microelectrodes that were fabricated on an existing neural recording microchip.
4.2 Future Work

- Two prototypes of the adaptive neural stimulation interface were recently integrated with the UEA. Future work will involve testing their stimulation and recording capabilities in vitro in mice hippocampi.

- Future research could focus on the development of adaptive algorithms that can be implemented on-chip for local signal processing, such as data compression or epileptic seizure detection.
Appendix A

Supplementary Hardware and Software Documentation

A.1 Board Design

The 4-layer PCB that was used to test the chip was designed using the Altium Designer 6 software. The manufactured board with components is shown in Figure A.1. The main on-board components are:

- XCV400E Virtex-E 2.5 V Xilinx FPGA for generating the digital input signals to the chip and acquiring the digital output signals from the chip.

- XC18V04, 44-pin PLCC, 3.3 V configuration IC PROM for programming the FPGA.

- 2.5V and 3.3V ultralow-noise, high-PSRR, low-dropout linear voltage regulators for creating the supply voltages.

- Two 12-bit DACs (AD5328) to generate the bias and reference voltages of the chip.
Figure A.1: The PCB designed for characterizing the adaptive neural stimulation interface.

- Three 3.3V, 1MHz Operational Amplifiers (AD8544) to buffer the bias and reference voltages.

- Two Dual Channel, 20MHz, 10-bit, CMOS ADCs (AD9201) to quantize the output of the test channels and the array off-chip.

- Six bias currents which are controlled by two (one fine and one coarse) potentiometers.

- Ten variable voltage regulators (TPS79501) generating the VDDs for the various blocks on the chip and the ICs on the board.

- A current DAC (AD9742) and two opamps (AD8652) to generate the differential ramp input to the ADC.

- Two SMA connectors to interface the analog output of the chip with external equipment.

- Debug pins to monitor the digital signals generated by the FPGA and the chip on the logic analyzer.
• 25 MHz clock oscillator.

### A.2 MATLAB Interface

MATLAB (version 7) scripts are used to provide an automated interface between the FPGA and the user through the I/O card. The I/O card is the PCI-bus compatible, National Instruments PCI-6534 card. The interface performs the following tasks:

• Generating the bias voltages for the chip by programming the corresponding DACs.

• Recording and displaying data from the chip.

• Generating the setup commands for the stimulator.

MATLAB functions, making use of mex files provided by the I/O board Universal Library, transmit and receive proper hand-shake signals to establish the communication between the FPGA and I/O Board. A list of MATLAB routines is as follows:

• $sendCommand(command)$: Sends a command to the FPGA that indicates a desired mode of operation.

• $initBias$: Sends the required bit sequence to the on-board DACs through FPGA for programming.

• $startCalib$: Re-sends the bias information for a specific bias voltage that is responsible for the feedback transistor. This is used to put the channel in buffer mode for a short time to quicken recovery from saturation.

• $setAddress(row,col)$: Requests to change the address of the row-select and column-select signals, such that a particular channel is enabled.
• **channelDisplay\_realtime**: Displays live acquired data on as many channels as are specified.

• **channelDisplay\_offline**: Displays the recorded videos off-line.

• **runStimulator**: Addresses the stimulators sequentially and programs them with the appropriate stimulation parameters.

• **patternStimulator**: Enables a pre-programmed stimulation pattern.

### A.3 FPGA Programming

Xilinx ISE Tools (version 9.2) are used to program the FPGA in order to generate the required signals and configure the chip in different modes of operation. A number of Verilog hardware description language (HDL) scripts perform this configuration task. Here is a brief description of these Verilog modules:

• **sTop.v**: The top-level Verilog module which instantiates all of the required sub-modules.

• **In.v**: Responsible for the handshaking protocol during communication of instruction from the computer.

• **cmdUnit.v**: Establishes the current state of operation based on the command sent from the computer and managed by In.v.

• **biasDAC.v**: Programs the DACs to generate the appropriate bias voltages based on the data sent from the computer.

• **addrCntrl.v**: Responsible for advancing the address of the row and the column in either direction based on input from various modules.
• *scanArray.v*: Controls the timing of the sample-and-hold clocks and ADC clocks during the sampling and quantization of the data.

• *stimulatorCtrl.v*: Controls the stimulator clock inputs based on the commands from the computer.
References


