AN ULTRA LOW POWER AMPLIFICATION AND
DIGITIZATION SYSTEM FOR NEURAL
SIGNAL RECORDING APPLICATIONS

By

SITONG YUAN

Bachelor of Science in Applied Physics
South China University of Technology
Guangzhou, China
1999

Master of Science in Natural and Applied Sciences
Oklahoma State University
Stillwater, OK
2005

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2010
AN ULTRA LOW POWER AMPLIFICATION AND DIGITIZATION SYSTEM FOR NEURAL SIGNAL RECORDING APPLICATIONS

Dissertation Approved:

Dr. Louis G. Johnson

Dissertation Adviser

Dr. Chris Hutchens

Dr. Sohum Sohoni

Dr. Nohpill Park

Dr. A. Gordon Emslie

Dean of the Graduate College
ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Louis G. Johnson for his trust, patience and help. In the last four years he has been helping me through my PhD study and guiding me towards the completion of my degree. This work would not exist without his efforts.

I would also like to thank Dr. Chris Hutchens. His discussion with me about analog circuit design has proved to be very helpful and fruitful. I also appreciate Dr. Cheng Chi Liu, Zhe Yuan and other students in Dr. Hutchens’ mixed signal lab, for their friendship and encouragement. I am proud to be a member among them.

I appreciate the support from my parents, Hanguang Yuan and Ruifang Zhong, and the encouragement from my wife, Jia Feng. I owe them a big ‘thank you’ because I couldn’t even finish my PhD study without their support.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background and Motivation</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Objective</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Organization</td>
<td>4</td>
</tr>
<tr>
<td>2 REVIEW OF FULLY INTEGRATED CONTINUOUS TIME FILTERS</td>
<td>6</td>
</tr>
<tr>
<td>2.1 Gm-C Filters</td>
<td>6</td>
</tr>
<tr>
<td>2.1.1 First Order Gm-C Filters</td>
<td>7</td>
</tr>
<tr>
<td>2.1.2 Transconductance Linearization Techniques</td>
<td>9</td>
</tr>
<tr>
<td>2.1.3 Simulation Result for A Gm-C Filter</td>
<td>19</td>
</tr>
<tr>
<td>2.2 MOSFET-C Filters</td>
<td>21</td>
</tr>
<tr>
<td>2.2.1 Equivalent Resistance</td>
<td>21</td>
</tr>
<tr>
<td>2.2.2 MOSFET-C Linearization</td>
<td>23</td>
</tr>
<tr>
<td>3 TUNABLE LOW-POWER FULLY INTEGRATED FILTER DESIGN</td>
<td>27</td>
</tr>
<tr>
<td>3.1 Review of Low Current MOSFET Model</td>
<td>27</td>
</tr>
<tr>
<td>3.2 Filter Configurations - Inverting and Non-inverting</td>
<td>30</td>
</tr>
<tr>
<td>3.3 Subthreshold OTA Design</td>
<td>35</td>
</tr>
<tr>
<td>3.4 Voltage Biased Pseudo-Resistor and Current Biased Pseudo-Resistor</td>
<td>37</td>
</tr>
<tr>
<td>3.5 Characterization of Current Biased Pseudo-Resistor</td>
<td>39</td>
</tr>
<tr>
<td>3.5.1 Linearity</td>
<td>39</td>
</tr>
<tr>
<td>3.5.2 Tunability</td>
<td>47</td>
</tr>
<tr>
<td>3.5.3 Frequency Response of the Pseudo-Resistor Bias</td>
<td>48</td>
</tr>
<tr>
<td>3.5.4 Noise Analysis of the Pseudo-Resistor</td>
<td>50</td>
</tr>
<tr>
<td>3.6 Noise Analysis of the Neural Filter</td>
<td>51</td>
</tr>
<tr>
<td>3.7 Bias Current Generation</td>
<td>55</td>
</tr>
<tr>
<td>3.8 Layout Issues</td>
<td>57</td>
</tr>
<tr>
<td>3.9 Single Ended Configuration in a Half-Micron CMOS Bulk Process</td>
<td>59</td>
</tr>
<tr>
<td>3.10 Fully Differential Configuration in a Half-Micron SOS CMOS Process</td>
<td>65</td>
</tr>
<tr>
<td>3.11 Conclusions</td>
<td>67</td>
</tr>
</tbody>
</table>
Chapter ........................................................................................................................................................................... Page

4. LOW-POWER SIGMA-DELTA ADC DESIGN ................................................................. 69

  4.1 Review of Nyquist rate ADC’s and oversampling ADC’s .................................. 69
  4.2 Modulator Order Selection ............................................................................. 70
  4.3 Modulator Design .......................................................................................... 71
    4.3.1 Reduced Harmonic Distortion CIFB ...................................................... 71
    4.3.2 Correlated Double Sampling ................................................................... 73
    4.3.3 Building Block Design ............................................................................. 75
  4.4 Simulation Results .......................................................................................... 76
  4.5 Conclusions .................................................................................................... 77

5. Measurement Results ............................................................................................... 79

  5.1 Single-Ended Filter Test Results ................................................................. 79
    5.1.1 Dies for Testing ....................................................................................... 79
    5.1.2 Frequency Response of the OTA ........................................................... 80
    5.1.3 Frequency Response of the Filter ........................................................... 82
    5.1.4 Tunability Test ....................................................................................... 82
    5.1.5 Input Common Mode Voltage Sensitivity .............................................. 83
    5.1.6 Noise Test ............................................................................................. 84
  5.2 Fully-Differential Filter Test Results ............................................................... 90
    5.2.1 Dies for Testing ....................................................................................... 90
    5.2.2 OTA Test Results ................................................................................... 90
    5.2.3 Filter Test Results .................................................................................. 91

6. CONCLUSIONS ....................................................................................................... 94

BIBLIOGRAPHY ............................................................................................................. 96
LIST OF TABLES

Table                                               Page

2.1 BJT and MOSFET Transconductor Comparison                      20
2.2 Gm-C and MOSFET-C Filter Comparison                           26
3.1 Cutoff Frequencies at Different $I_B$                        65
3.2 Fully Differential Filter Characterization                   66
4.1 Modulator Comparison                                         78
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 A diode connected transconductor.</td>
<td>8</td>
</tr>
<tr>
<td>2.2 (a) Gm-C lowpass filter and (b) Gm-C highpass filter</td>
<td>8</td>
</tr>
<tr>
<td>2.3 Source degeneration techniques</td>
<td>10</td>
</tr>
<tr>
<td>2.4 Transconductors biased in weak inversion region</td>
<td>11</td>
</tr>
<tr>
<td>2.5 Drain currents of the input pairs M1 to M4 versus the input differential voltage before (a) and after (b) source degeneration</td>
<td>12</td>
</tr>
<tr>
<td>2.6 The effective transconductance of the transconductors in Fig. 2.5</td>
<td>13</td>
</tr>
<tr>
<td>2.7 Active source degeneration</td>
<td>14</td>
</tr>
<tr>
<td>2.8 Source degeneration technique combining Fig. 2.7(b) and (c)</td>
<td>15</td>
</tr>
<tr>
<td>2.9 Bump-linearization</td>
<td>16</td>
</tr>
<tr>
<td>2.10 A circuit implementation of nonlinear term cancellation technique</td>
<td>17</td>
</tr>
<tr>
<td>2.11 Transconductance linearization topologies using multiple differential pairs with input offset voltages: (a) BJT case, (b) MOSFET case in weak inversion region</td>
<td>18</td>
</tr>
<tr>
<td>2.12 Transconductor with source degeneration and bump linearization techniques</td>
<td>20</td>
</tr>
<tr>
<td>2.13 Frequency response of the gm-C high pass filter</td>
<td>21</td>
</tr>
<tr>
<td>2.14 Fully differential MOSFET-C integrator</td>
<td>25</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>2.15</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>30</td>
</tr>
<tr>
<td>3.2</td>
<td>31</td>
</tr>
<tr>
<td>3.3</td>
<td>34</td>
</tr>
<tr>
<td>3.4</td>
<td>38</td>
</tr>
<tr>
<td>3.5</td>
<td>40</td>
</tr>
<tr>
<td>3.6</td>
<td>40</td>
</tr>
<tr>
<td>3.7</td>
<td>41</td>
</tr>
<tr>
<td>3.8</td>
<td>42</td>
</tr>
<tr>
<td>3.9</td>
<td>46</td>
</tr>
<tr>
<td>3.10</td>
<td>48</td>
</tr>
<tr>
<td>3.11</td>
<td>49</td>
</tr>
<tr>
<td>3.12</td>
<td>52</td>
</tr>
<tr>
<td>3.13</td>
<td>54</td>
</tr>
<tr>
<td>3.14</td>
<td>55</td>
</tr>
<tr>
<td>3.15</td>
<td>59</td>
</tr>
<tr>
<td>3.16</td>
<td>61</td>
</tr>
<tr>
<td>3.17</td>
<td>61</td>
</tr>
<tr>
<td>3.18</td>
<td>62</td>
</tr>
<tr>
<td>3.19</td>
<td>63</td>
</tr>
<tr>
<td>3.20</td>
<td>63</td>
</tr>
</tbody>
</table>
Figure Page

3.21 Input and output noise spectrum densities for filter A and filter C ...............64
3.22 Fully-differential folded-cascode OTA..........................................................66
3.23 Frequency responses of the fully differential filters ......................................67
4.1 Ordinary second order topology (a) and reduced harmonic distortion topology (b) ..............................................................72
4.2 Integrator comparison: (a) ordinary integrator, (b) integrator with CDS technique, (c) integrator with CDS technique and slew-prevention capacitor ..73
4.3 (a) The clocked comparator (the 1-bit ADC) and (b) The 1-bit DAC used in this project ................................................................................................................76
4.4 Implementation of the second order sigma-delta modulator .............................77
4.5 Transient simulation results of the second order modulator ..............................78
5.1 Die structure for the single ended filter in a half micron CMOS process............80
5.2 (a) Die photo for the chip and (b) die photo for filter A.................................81
5.3 Frequency response of the OTA .....................................................................83
5.4 Frequency response test setup.........................................................................84
5.5 Frequency response of filter A ........................................................................85
5.6 Frequency response of filter B ........................................................................85
5.7 Frequency response of filter C ........................................................................86
5.8 Tunability test setup.........................................................................................86
5.9 Measured frequency responses for filter B for different $I_B$.............................87
5.10 Frequency response test setup for different input common mode voltages.....87
<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.11</td>
<td>88</td>
</tr>
<tr>
<td>5.12</td>
<td>88</td>
</tr>
<tr>
<td>5.13</td>
<td>89</td>
</tr>
<tr>
<td>5.14</td>
<td>89</td>
</tr>
<tr>
<td>5.15</td>
<td>91</td>
</tr>
<tr>
<td>5.16</td>
<td>92</td>
</tr>
<tr>
<td>5.17</td>
<td>92</td>
</tr>
<tr>
<td>5.18</td>
<td>93</td>
</tr>
<tr>
<td>5.19</td>
<td>93</td>
</tr>
<tr>
<td>6.1</td>
<td>95</td>
</tr>
</tbody>
</table>

5.11 Measured frequency responses at different common mode voltage for filter B

5.12 Noise test setup.

5.13 Measured noise spectrum densities for filter A

5.14 Measured noise spectrum densities for filter C

5.15 Filter structure and test setup

5.16 Die photo for the fully differential filters fabricated in a half micron SOS process

5.17 OTA test setup

5.18 OTA test results

5.19 Frequency responses of stage 1 filter

6.1 System diagram of the amplification and digitization front end
CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

Biopotentials reflect the biochemical processes ongoing in living tissues [1]. The importance of understanding biopotentials lies in the fact that they can aid clinicians to diagnose diseases and bioengineers to study the communication between living cells. However, different biopotentials contain different frequency components with weak amplitudes. For example, Electrocardiography (ECG) has a frequency range from 0.05 Hz to 100 Hz and an amplitude range from 1 mV to 5 mV; Electromyography (EMG) has a frequency range from 20 Hz to 2 KHz and an amplitude range from 1 mV to 10 mV; Extracellular Action Potential (EAP) has a frequency range from 0.1 Hz to 10 KHz and an amplitude range from 50 µV to 500µV [1;2]. As a result, amplifying and filtering circuitry is usually required before a useful biopotential is present for further study.

Among these biopotentials, EAP is especially useful for neuroprosthetic studies [2]. By observing the neural spikes generated by neurons in the brain, one is able to predict the instructions given by the brain intended for body actions. However, most of the neural spike energy falls in the frequency range from 300 Hz to 5 KHz while EAP contains the frequency components from 0.1 Hz up to 10 KHz. Hence, a filtering stage has to be implemented to enhance the signals within the interested band and attenuate those outside the band.
Biopotentials can be amplified and recorded externally with external circuitry, or internally with internal or implanted circuitry. Implanted circuitry has the advantages over its external counterpart such as lower noise, smaller physical size and lower risk of infection of the body [2]. Due to the low frequency characteristics of the biopotentials, fully integrated, implantable filters with accordingly low cutoff frequencies have to be utilized to accomplish the internally filtering and recording tasks. Such low frequencies usually require very large capacitance and resistance, which take unacceptable chip area. As a result, alternative techniques such as gm-C and MOSFET-C methods have been utilized to implement these filters.

As a popular technique to implement fully integrated filters, gm-C uses transconductors to replace resistors to realize the cutoff frequency [3]. In recent years transconductors with very small transconductance have been proposed by different research groups. These transconductors utilize attenuation, or such feedback technique as source degeneration to realize small transconductances with an extended linear range [4-6]. However, gm-C methods are probably not very area and power efficient because the area dedicated to a transconductor is comparable to that for an OTA, and a gm-C filter usually doesn’t provide a very high gain while performing the task of filtering.

MOSFET-C filters are another solution for fully integrated applications. Historically, gate voltage controlled MOSFET’s operated in triode region have been used as pseudo-resistors for a long time, and have been applied to neural signal recording applications by several research groups in recent years [7;8]. The advantages of using a MOSFET as a pseudo-resistor are that: 1) Very high resistance is achievable in a small area; and 2) The resistance is tunable. However, the disadvantage is also obvious, which is that when a
gate voltage-controlled MOSFET resistor is used, the resistance is vulnerable to process variations, and even to OTA output common mode voltage deviations when one side of the pseudo-resistor is connected to the output of an OTA, especially when the MOSFET resistor is biased in the subthreshold region in order to achieve a very high resistance, in which case the resistance is exponentially dependent upon the bias voltages. Significant changes in the resistance can be induced by small process variations or output common mode voltage shifts. Hence tuning is usually required for voltage-controlled pseudo-resistors to bring back the desired resistance value. In cases such as implantable applications when tuning is impossible, gate voltage controlled pseudo-resistors are not practical.

An alternative way to bias the pseudo-resistor is to use “current biasing”. As it is easier to achieve predictable currents in today’s semiconductor processes than to achieve predictable voltages, it is easier to achieve predictable resistance by biasing a pseudo-resistor with a current. In this work, we use current biasing technique to bias pseudo-resistors in the subthreshold region. Test results show that current biased pseudo-resistors have more predictable resistance and are insensitive to the OTA output common mode voltage. Current biased pseudo-resistors are good candidates for implantable filter applications.

1.2 Objective

The objective of this project is to develop a tunable low power fully integrated bandpass filter for neural signal amplifications. The bandwidth of this filter is from 500 Hz to 5 KHz, and the area should be as small as possible for implantable applications. The targeted power consumption is less than 2 µW for a gain of 40 dB. The lower cutoff
frequency is tunable from tens of Hz to hundreds of Hz. The filter will be fabricated in both a half micron 3-metal 2-poly CMOS bulk process for a power supply of 5 V and a half micron silicon-on-sapphire (SOS) CMOS process for a power supply of 1.2 V. Test results will be given.

A low power second order sigma-delta ADC modulator will be designed as the digitization circuit. The signal bandwidth for the modulator is 5 KHz and the resolution is 10 bits. The power consumption is about 13 \( \mu \text{W} \) for a power supply of 1.2 V. The ADC will be simulated in a half micron SOS CMOS process.

1.3 Organization

Chapter I introduces the background, motivation and objective of this work.

Chapter II is the literature review section, which reviews the techniques in gm-C and MOSFET-C filter design. Some of the circuits are also simulated to verify their performance.

Chapter III discusses the design of a tunable low power full integrated filter which uses current biased pseudo-resistors. This section starts with the review of low current MOS transistor models which will be used in pseudo-resistor characterization and subthreshold filter design. The performance of current biased pseudo-resistor is then analyzed in detail. This chapter analytically verifies that current biased pseudo-resistors are suitable for fully integrated biomedical applications. Layout issues concerning subthreshold circuit design are also addressed. The filters were designed and simulated in two CMOS processes: a half micron 3-metal 2-poly CMOS bulk process for the single ended filters and a half micron silicon-on-sapphire (SOS) CMOS process for the fully differential filters. Simulation results are given.
Chapter IV discusses the design of a 10-bit second order sigma-delta ADC. The OTA’s used in this ADC operate in the subthreshold region. Techniques for reducing harmonic distortion and OTA offsets are also discussed. Simulation results will be given.

Chapter V gives the measurement results for the bandpass filters, and experimentally proves the analysis about the current biased pseudo-resistor in chapter III.

Chapter VI summarizes the results of this work.
CHAPTER 2

REVIEW OF FULLY INTEGRATED CONTINUOUS-TIME FILTERS

Analog filters can be divided into two categories: switched-capacitor filters and continuous time filters. In order to avoid aliasing, switched-capacitor filters must sample the input signal at frequencies higher than Nyquist frequency, which results in OTAs with higher gain-bandwidth-product (GBP) and hence higher power consumptions than continuous time filters for the same signal bandwidth [3]. Considering the low power requirement for this project, continuous time filters are selected.

This section reviews the techniques to implement continuous-time filters, some of which have been utilized to the neural amplifier of this project.

2.1 Gm-C Filters

A popular technique to implement fully integrated continuous time filters is gm-C filters. Gm-C filters use transconductors and capacitors as their key building blocks. Transconductors are actually voltage controlled current sources with a linear relationship between the input voltage and the output current. An additional requirement placed on gm-C filters for neural recording applications is a very small transconductance since neural recording gm-C filters normally realizes very low cutoff frequencies. This section reviews some basic techniques and configurations to implement first order gm-C filters and to realize transconductance reduction and linearization.
2.1.1 First Order Gm-C Filters

Fig. 2.1 illustrates a diode connected transconductor which is a building block for first order high-pass or low-pass gm-C filters [9]. Assuming the input voltage and the input current are $V_{in}$ and $I_{in}$ respectively, the equivalent input impedance can be expressed as:

$$Z_{in} = \frac{V_{in}}{I_{in}}. \quad (2.1)$$

Since the transconductor is a voltage controlled current source, the relationship between the input voltage (as showing in Fig. 2.1) and the output current is:

$$I_{out} = -g_m V_{in}. \quad (2.2)$$

Observing $I_{out} = -I_{in}$, the input impedance can be finally expressed as:

$$Z_{in} = \frac{1}{g_m}. \quad (2.3)$$

The advantage of this structure is that its $g_m$ is variable, i.e. it’s tunable. For fully integrated neural filters, when large resistance is not realizable by a real resistor, a transconductor with extremely small transconductance can be an alternative. Techniques for minimizing the transconductance will be reviewed in section 2.1.2.

Fig. 2.2 shows the structures of a first order lowpass filter (Fig. 2.2 (a)) and a first order highpass filter (Fig. 2.2 (b)) by replacing the real resistors with the diode connected transconductor mentioned above. The transfer function for both filters can be obtained using voltage divider rule. By referring to (2.3), the transfer function for the low-pass filter is:

$$\frac{V_{out}}{V_{in}} = \frac{g_m}{sC + g_m}. \quad (2.4)$$
And for the high-pass filter, the transfer function is found to be:

\[
\frac{V_{out}}{V_{in}} = \frac{sC}{sC + g_m}.
\]  

(2.5)

Figure 2.1, A diode connected transconductor.

Both filters in Fig. 2.2 filter the input signals without offering any gain (not very power and area efficient), however the cutoff frequencies are adjustable, and can be made very small without using large resistors. The positions of the poles for both cases are \( g_m/C \).

Another disadvantage associated with these filters is that the transconductors introduce more noise than a single MOSFET pseudo-resistor.

Figure 2.2, (a) Gm-C lowpass filter and (b) Gm-C highpass filter.
The structure showing in Fig. 2.2 (b) has been used for neural signal recording applications to realize an adjustable large time constant [2]. By cascading it after a high gain stage, its noise contribution can be ignored.

2.1.2 Transconductance Linearization Techniques

As mentioned above, transconductors are voltage controlled current sources with linear transconductance. Ordinary OTAs’ transconductance is not linear because it is a function of the overdrive voltages and the drain currents of the input transistors. In order to resolve this problem, different techniques have been proposed by different authors, such as source degeneration, bump linearization, nonlinear term cancellation and attenuation [10;11].

Fig. 2.3 shows two source degeneration schemes. Source degeneration is actually current-voltage feedback in which cases the transconductance is given by [12]:

\[ g_{m,eff} = \frac{g_m}{1 + g_m R_S}, \]

where \( g_m \) is the transconductance of the input transistors. Each case in Fig. 2.3 has its advantages and disadvantages. Fig. 2.3(a) has a better noise performance than Fig. 2.3(b) because the noise generated by the current sink appears as common mode noise at the output, which is in contrast to the differential noise in Fig. 2.3 (b) [11]. However, the bias current flows through the resistors \( R_S \) and reduces the input an output common mode ranges in Fig. 2.3 (a), which is in contrast to the circuit in Fig. 2.3 (b) where the bias current doesn’t flow through \( R_S \) and the input and output common mode ranges are not reduced. Fig. 2.3 (b) is more suitable for low voltage applications, such as RFID chips.
The resistor $R_S$ can also be implemented by diode connected MOSFETs when a large value is required, as showing in Fig. 2.4(b). This is actually corresponding to the case in Fig. 2.3 (a) where the real resistors are replaced by active elements. In Fig. 2.4(b) the transistors $M_5$ and $M_6$ act as two degenerating resistors whose equivalent resistance is $1/g_{ms}$, and (2.6) can be expressed as:

$$g_{m,\text{eff}} = \frac{g_m}{1 + g_m / g_{ms}}. \quad (2.7)$$

![Diagram](image)

Figure 2.3, Source degeneration techniques.

When a transistor is biased in the subthreshold region, its transconductance is proportional to its drain current, i.e. $g_m = I_D/nU_T$ [13]. In Fig. 2.4(b), $M_3$, $M_4$, $M_5$ and $M_6$ all have the same transconductance because they are all biased in weak inversion region at the same bias current, i.e. $g_m = g_{ms}$ in (2.7), and the effective transconductance can be simplified as:

$$g_{m,\text{eff}} = \frac{g_m}{2}. \quad (2.8)$$
It is interesting that the transconductance of a transistor in weak inversion region is not related to its geometry, so it is impossible to change the effective transconductance of M3 to M6 by altering their geometries.

The simulation results of the transconductance of the transconductors in Fig. 2.4(a) and Fig. 2.4(b) are compared in Fig. 2.5 and Fig. 2.6. The simulation results were obtained from Cadence analog environment in a half micron three-metal two-poly CMOS process. The sizes of M1 to M4 are 4@6µm/6µm and the sizes of M5 and M6 are both 2@6µm/6µm. All transistors are biased at 30 nA to guarantee a weak inversion operation.

![Figure 2.4, Transconductors biased in weak inversion region.](image)

Fig. 2.5 was obtained by fixing one of the inputs at the common mode voltage of 0 V (with a power supply of ±2.5 V) while sweeping the other input from -600 mV to 600 mV. It is observed that the linear range of the transconductance is expanded with a reduced magnitude. The effective transconductance of the transconductors in Fig. 2.4(a) and Fig. 2.4(b) is plotted in Fig. 2.6, labeled by \( g_{ma} \) and \( g_{mb} \) respectively. It is found that the change of \( g_{mb} \) is more gradual than \( g_{ma} \), and the maximal value of \( g_{mb} \) is one-half of
For the same percentage error, $g_{mb}$ expands to a wider input voltage range. The disadvantages of this source degeneration technique are obvious: 1) the improvement in linearity is limited due to the different transconductance characteristic in weak inversion region and; 2) the two diode connected transistors reduce the input common mode range of the circuit.

![Graph](image)

Figure 2.5, Drain currents of the input pairs M1 to M4 versus the input differential voltage before (a) and after (b) source degeneration.

The circuits in Fig. 2.7 correspond to the case in Fig. 2.3(b) with active elements replacing the linear resistor. Czarnul et al and Krummenacher et al proposed the circuits in Fig. 2.7(a) and (b) with transistors biased in triode region [14;15], respectively; and Torrance et al proposed the circuit in Fig. 2.7(c) with transistors biased in saturation region [16]. The effective transconductance of the circuit in Fig. 2.7(a) is identical to that in Fig. 2.4(b). The transconductance in Fig. 2.7(a) is tunable by adjusting the gate voltage of $M_3 (V_{G3})$; however, the gate voltage has to be large enough to bias $M_3$ in triode region.
for better linearity. The disadvantages of the circuit in Fig. 2.7(a) are that: 1) it is sensitive to the common mode input voltage because $V_{GS3}$ can be altered by changing its source voltage $V_{S3}$ which is affected by the common mode input voltage [11] and; 2) it is sensitive to the offset voltage which is a result of device mismatches. The latter disadvantage makes it unsuitable for fully integrated applications. An improved version of Fig. 2.7(a) is shown in Fig. 2.7(b), in which case the gates of $M_3$ and $M_4$ are tied to the input terminal. This topology alleviates the first disadvantage mentioned above because the gate voltages of $M_3$ and $M_4$ track the input voltages. The effective transconductance is given by the authors as ($M_1 = M_2$, $M_3 = M_4$):

$$g_{mef} = \frac{g_{m1}}{1 + \frac{\beta_1}{4\beta_3}}. \quad (2.9)$$

![Graph](image)

**Figure 2.6.** The effective transconductance of the transconductors in Fig. 2.5.
The circuit in Fig. 2.7(c) uses transistors in saturation to perform source degenerating, whose effective transconductance is also identical to (2.12). This circuit is not sensitive to input common mode voltage but the linearity improvement is limited.

Figure 2.7, Active source degeneration.

Silva-Martinez et al. combined the circuits in Fig. 2.7(b) and Fig. 2.7(c) together to achieve better linearization performance and smaller transconductance [17], as showing in Fig. 2.8.

Another linearization technique, called “bump-linearization” by the authors, was proposed by Sarpeshkar et al. [5]. This circuit is from the “bump circuit” invented by Delbruck [10]. In this circuit, an additional current path consisting of M3 and M4 is introduced into the transconductor, as shown in Fig. 2.9. It is called “bump-linearization” because the current \( I_3 \) reaches its maximum when \( V_+ = V_- \), which is like a “bump shape.” The underlying principle is explained by the authors in [5] with the aid of (2.10) to (2.13), if the transistors are biased in the subthreshold region:

\[
I_1 - I_2 = \frac{\sinh x}{\beta + \cosh x},
\]  

(2.10)
where $w = W/L$ is the aspect ratio of the “bump transistors” $M_3$ and $M_4$, $n$ is the subthreshold slope factor and $U_T$ is the thermal voltage. When $w = 2$ and hence $\beta = 2$, (2.10) can be expressed as:

$$\frac{\sinh x}{2 + \cosh x} = \frac{x}{3} - \frac{x^5}{540} + \frac{x^7}{4536} - \frac{x^9}{77760} + \ldots$$

Attenuation techniques achieve linear range extension by attenuating the input signals. Attenuation can be achieved by using either passive elements (a voltage divider consisting of two linear resistors in series) or active elements (bulk driven transistors [18] or two transistors in series [19]). However, attenuation technique increases the input referred noise, which disqualifies this technique from being applied to neural signal recording applications when the noise performance is important.
Equation (2.13) indicates that the 2nd and 3rd order terms don’t exist and the 5th order term dominates the distortion, however, this term is small enough to be ignored.

Wang and Guggenbuhl described a nonlinear term cancellation technique in [20]. This technique uses two cross-coupled input pairs, one of which has an offset bias voltage $V_B$, as showing in Fig. 2.10. Assuming all transistors are biased in the square law region, the currents $I_1$ and $I_2$ can be expressed as:

$$I_1 = \frac{\beta}{2} (V_{i+} - V_x - V_{m})^2 + \frac{\beta}{2} (V_{i-} - V_B - V_x - V_{m})^2,$$  \hfill (2.14)

$$I_2 = \frac{\beta}{2} (V_{i-} - V_x - V_{m})^2 + \frac{\beta}{2} (V_{i+} - V_B - V_x - V_{m})^2.$$  \hfill (2.15)

The output current, $I_1 - I_2$, is then given by:

$$I_1 - I_2 = \beta V_B (V_{i+} - V_{i-}).$$  \hfill (2.16)

This circuit requires that all transistors are in the square law region. When the circuit is biased in weak inversion region for low power applications, alternative topologies have to be implemented.
Figure 2.10, A circuit implementation of nonlinear term cancellation technique.

Since MOSFETs biased in weak inversion region behave like bipolar junction transistors (BJTs), the transconductance linearization techniques developed for BJT circuits can be transplanted to MOSFET cases in weak inversion region. Fig. 2.11 shows a linearization technique using multiple input pairs with offset input voltages for BJT cases (Fig. 2.11 (a)) and MOSFET cases (Fig. 2.11 (b)) in weak inversion region.

Calder and Tanimoto et al described this technique in [21] and [22] for BJT transconductors. When the offset voltage $V_i$ is chosen properly, the combinational transconductance contributed by $Q_1$ and $Q_3$ (or $Q_2$ and $Q_4$) will be maximally linearized. When there are only two differential pairs in parallel (as in Fig. 2.11 (a)), the equivalent transconductance is given by taking the derivative of the differential currents with refer to the differential input voltages:
Figure 2.11, Transconductance linearization topologies using multiple differential pairs with input offset voltages: (a) BJT case, (b) MOSFET case in weak inversion region.

\[ G(V_{in}) = \frac{dI_{diff12}}{d(V_{in} - V_1)} + \frac{dI_{diff34}}{d(V_{in} + V_1)}, \]  

(2.17)

where

\[ V_{in} = V_{i+} - V_{i-}, \]  

(2.18)

\[ I_{diff12} = I_{c1} - I_{c2} = I_{EE} \tanh \left( \frac{V_{in} - V_1}{2U_T} \right), \]  

(2.19)
\[ I_{\text{diff}34} = i_{C3} - i_{C4} = I_{EE} \tanh \left( \frac{V_m + V_l}{2U_T} \right), \]  \hspace{1cm} (2.20)

It is found that when \( V_l = 1.317 \, U_T \), the derivatives of \( G \) at \( V_{in} = 0 \) are zero to the highest order, which indicates a maximally linearized transconductance. A practical realization of the offset voltage \( V_l \) is achieved by setting the aspect ratio between the two input transistors ((\( W/L \))\(_1\)/((\( W/L \))\(_2\)) to be \( \exp(V_l/U_T) \). In the case when \( \exp(1.317) = 3.73 \), the ratio is set to 4.

For MOSFETs in weak inversion region, the thermal voltage \( U_T \) is replaced by \( nU_T \) in (2.19) and (2.20) for the same results, where \( n \) is the subthreshold slope factor. A comparison for these two cases is given in Table 2.1.

2.1.3 Simulation Result for a Gm-C Filter

The first order high pass filter in Fig. 2.2 (b) was simulated with Cadence Virtuoso Analog Environment in a half micron three-metal two-poly CMOS process. The schematic of the transconductor used in this simulation is showing in Fig. 2.12, which is a modified version of the one described in [5]. The difference of this transconductor is that it doesn’t use bulk driven transistors to achieve transconductance linearization and attenuation as in [5], since bulk driven transistors are not very well modeled by the simulator. The transconductor was biased in the subthreshold region with a bias current \( I_B = 1.3 \) nA, and the capacitor in Fig. 2.2(b) was 2 pF. The simulation result is showing in Fig. 2.13, in which the cutoff frequency is 480 Hz. The simulation result indicates that even though the transconductor is biased at a very low current, a relatively large capacitor is still required for low frequency applications. Given the fact that this filter doesn’t provide any gain, it is not an efficient way to implement low frequency filters.
### Table 2.1, BJT and MOSFET Transconductor Comparison

<table>
<thead>
<tr>
<th>Transconductance of Q1 and M1</th>
<th>BJT (Fig. 2.11 (a))</th>
<th>MOSFET (Fig. 2.11(b))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transconductance of Q1 and M1</td>
<td>( g_{m1} = \frac{I_{C1}}{U_T} )</td>
<td>( g_{m1} = \frac{I_{D1}}{nU_T} )</td>
</tr>
<tr>
<td>Differential Output Currents</td>
<td>( I_{diff 12} = I_{EE} \tanh \left( \frac{V_{in} - V_1}{2U_T} \right) )</td>
<td>( I_{diff 12} = I_{SS} \tanh \left( \frac{V_{in} - V_1}{2nU_T} \right) )</td>
</tr>
<tr>
<td>Differential Output Currents</td>
<td>( I_{diff 34} = I_{EE} \tanh \left( \frac{V_{in} + V_1}{2U_T} \right) )</td>
<td>( I_{diff 34} = I_{SS} \tanh \left( \frac{V_{in} + V_1}{2nU_T} \right) )</td>
</tr>
<tr>
<td>Condition for Maximal Linearization</td>
<td>( V_1 = 1.317U_T )</td>
<td>( V_1 = 1.317nU_T )</td>
</tr>
</tbody>
</table>

Figure 2.12, Transconductor with source degeneration and bump linearization techniques.
Figure 2.13, Frequency response of the gm-C high pass filter.

2.2 MOSFET-C Filters

MOSFET-C filters are an alternative to gm-C filters for realization of fully-integrated continuous time filters. Due to their similarities to active-RC filters, filter theories about active-RC filters can be translated into MOSFET-C applications directly [3]. Since a MOSFET biased in triode region (called pseudo-resistor in this project) is a nonlinear resistor, linearization techniques have to be utilized to reduce output signal distortions.

2.2.1 Equivalent Resistance

Since MOSFET-C filters replace the linear resistors in RC filters with MOSFETs, it is necessary to study the equivalent resistance of a MOSFET in triode region. Nowadays
all the MOSFET-C filters use gate voltage controlled MOSFET’s as pseudo-resistors. If a MOSFET is biased in strong inversion region, its equivalent resistance is given by [23]:

\[
 r_{DS} = \frac{1}{\mu C_{ox} \left( \frac{W}{L} \right) \left[ (V_{GS} - V_t) - V_{DS} \right]}.
\]  (2.21)

It’s clear that the equivalent resistance is not linear and its value depends upon the drain-to-source voltage \( V_{DS} \). The effective resistance is defined for \( V_{DS} = 0 \):

\[
 r_{eff} = \frac{1}{\mu C_{ox} \left( \frac{W}{L} \right) \left( V_{GS} - V_t \right)}.
\]  (2.22)

If the MOSFET is biased in the weak inversion region, the equivalent resistance is [13]:

\[
 r_{DS} = \frac{U_T}{I_{D0} \exp \left( \frac{V_G - nV_S}{nU_T} \right) \exp \left( \frac{V_{DS}}{U_T} \right)}.
\]  (2.23)

\[
 I_{D0} = 2n \mu C_{ox} \frac{W}{L} U_T^2 \exp \left( \frac{-V_{T0}}{nU_T} \right).
\]  (2.24)

Equation (2.23) shows that the equivalent resistance for a subthreshold MOSFET is a strong function of the gate bias voltage, the threshold voltage and the drain-to-source voltage due to the exponential relationships. There would be magnitudes of difference for the equivalent resistance across process corners. The effective resistance is defined for \( V_{DS} = 0 \):

\[
 r_{eff} = \frac{U_T}{I_{D0} \exp \left( \frac{V_G - nV_S}{nU_T} \right)}.
\]  (2.25)
Since MOSFET pseudo-resistors are nonlinear, such technique as nonlinear terms cancellation have to be implemented to extend the linear range of the pseudo resistors. One popular technique is to use fully differential topology to cancel out the even order terms. For pseudo-resistors biased in the subthreshold region, the equivalent resistance is exponentially dependent on the gate bias voltage and the threshold voltage which is hard to control, and alternative bias schemes have to be utilized to establish a more predictable resistance.

2.2.2 MOSFET-C Linearization

One technique for MOSFET-C linearization is to use fully differential circuits to cancel out higher order distortion terms. Fig. 2.15 shows two integrator circuits to implement this technique. Banu and Tsividis described the circuit that is able to cancel out all even order distortion terms [24], which is showing in Fig. 2.15(a); and Czarnul modified Banu and Tsividis’ circuit and proposed a circuit which is able to cancel out both even order and odd order distortion terms [25], as showing in Fig. 2.15(b). Their techniques are based on the drain current model for MOSFETs in the triode region introduced by Penney and Lau in [26].

For the circuit in Fig. 2.15(a), the output voltage of the integrator can be written as:

\[
V_o = V_{o+} - V_{o-} = -\frac{1}{2C} \int_{-\infty}^{t'} (I_{D1} - I_{D2})dt',
\]

(2.29)

where \(I_{D1}\) and \(I_{D2}\) are the drain currents of the input pseudo-resistors. By expanding \(I_{D1}\) and \(I_{D2}\) in a Taylor series with refer to \(V_D\) (i.e. \(V_{i+}\) and \(V_{i-}\)), and observing that \(V_{i+}\) and \(V_{i-}\) are opposite in sign, all even order terms can be cancelled.

Fig. 2.15(b) is a modified version of Fig. 2.15(a) in which case:
\begin{align*}
V_o &= V_{o+} - V_{o-} = -\frac{1}{2C} \int_{-\infty}^{\infty} [(I_{D1} + I_{D2}) - (I_{D3} + I_{D4})] dt' \\
&= -\frac{1}{2C} \int_{-\infty}^{\infty} [(I_{D1} - I_{D3}) + (I_{D2} + I_{D4})] dt'. 
\end{align*}

(2.30)

By using the model suggested by Penney and Lau in [26], it is found that:

\begin{align*}
I_{D1} - I_{D3} &= 2K(V_{c1} - V_{c2})(V_{i+} - V_3), \quad (2.31) \\
I_{D2} - I_{D4} &= -2K(V_{c1} - V_{c2})(V_{i-} - V_3). \quad (2.32)
\end{align*}

And (2.30) can be expressed as:

\begin{equation}
V_o = V_{o+} - V_{o-} = -\frac{1}{2C} \int_{-\infty}^{\infty} 2K(V_{c1} - V_{c2})(V_{i+} - V_{i-}) dt'. 
\end{equation}

(2.33)

Equation (2.33) is a linear function of the input differential voltage, and the equivalent resistance is:

\begin{equation}
R = \frac{1}{K(V_{c1} - V_{c2})}. 
\end{equation}

(2.34)

A fully differential neural amplifier front end was also described by Mollazadeh et al in [27], which is showing in Fig. 2.15. This circuit is based on Harrison’s capacitive coupling high-pass filter described in [28]. By using a fully differential topology, even order distortion terms introduced by the pseudo-resistor are cancelled out. The authors report a total harmonic distortion less than 1% while the input peak-to-peak voltage is 10 mV and the gain is 39.6 dB.

In this project, a fully-differential topology is also implemented to cancel out the even order distortion terms coming from the nonlinear pseudo-resistor which is biased in the weak inversion region. Chapter III details the design of a current-biased pseudo-resistor and the techniques to improve its linearity. Table 2.2 summarizes the advantages and disadvantages associated with gm-C filters and MOSFET-C filters.
Figure 2.14, Fully differential MOSFET-C integrator.

Figure 2.15, Fully differential neural filter.
<table>
<thead>
<tr>
<th></th>
<th><strong>Gm-C</strong></th>
<th><strong>MOSFET-C</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
<td>1. Transconductance tunable;</td>
<td>1. Tunable;</td>
</tr>
<tr>
<td></td>
<td>2. Very small transconductance achievable – large time constant.</td>
<td>2. Large time constant, small area;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. High gain – power and area efficient, lower input referred noise;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Simple structure – closely relative to active RC filters.</td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
<td>1. Low gain – not power and area efficient, high input referred noise;</td>
<td>1. Nonlinear;</td>
</tr>
<tr>
<td></td>
<td>2. Complicated in structure;</td>
<td>2. Vulnerable to process variations – tuning required.</td>
</tr>
<tr>
<td></td>
<td>3. Linearization techniques required.</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.2, Gm-C and MOSFET-C Filter Comparison**
3.1 Review of Low Current MOSFET Model

The low current, weak inversion model proposed by Enz and Vittoz [13] is used in this project for circuit modeling. This section briefly reviews this model as it sets the fundamentals of this project.

When MOSFETs are biased in weak inversion region, the current is mainly due to diffusion mechanism, as opposed to the strong inversion case where drift current dominates. The advantage of this model is that it provides a smooth transition from triode region to saturation region, i.e. the same drain current expression can be used in either triode region or saturation region, as showing in (3.1):

\[
I_D = I_F - I_R = I_{D0} \cdot \exp \left( \frac{V_G}{n \cdot U_T} \right) \cdot \left\{ \exp \left( -\frac{V_s}{U_T} \right) - \exp \left( -\frac{V_D}{U_T} \right) \right\}, \quad (3.1)
\]

where \( I_{D0} = I_s \cdot \exp \left( -\frac{V_{T0}}{n \cdot U_T} \right) \). \( I_F \) and \( I_R \) are called “forward current” and “reverse current” by the authors, as expressed in (3.2) and (3.3):

\[
I_F = I_s \cdot \exp \left( \frac{V_p - V_S}{U_T} \right), \quad (3.2)
\]

\[
I_R = I_s \cdot \exp \left( \frac{V_p - V_D}{U_T} \right). \quad (3.3)
\]
In this model, the authors refer all the voltages to the substrate, and $V_{T0}$, $V_S$ and $V_D$ are the threshold voltage, the source voltage and the drain voltage respectively. $V_P$ and $I_S$ are defined as the “pinch-off voltage” and the “specific current” by the authors, where

$$V_P = \frac{V_G - V_{T0}}{n},$$

and $I_S$ is expressed as:

$$I_S = 2 \cdot n \cdot \beta \cdot U_T^2,$$  \hspace{1cm} (3.4)

$$\beta = \mu \cdot C_{ox} \cdot \frac{W_{eff}}{L_{eff}}.$$ \hspace{1cm} (3.5)

The specific current $I_S$ can be compared to the drain current to determine if the device is biased in strong inversion (for $I_D > I_S$) or weak inversion (for $I_D < I_S$). By factoring out the exponential term containing $V_S$, (3.1) can also be expressed as the following equivalent expression:

$$I_D = I_F \cdot \left(1 - \frac{I_R}{I_F}\right) = I_S \cdot \exp\left(\frac{V_P - V_S}{U_T}\right) \cdot \left(1 - \exp\left(-\frac{V_D - V_S}{U_T}\right)\right).$$ \hspace{1cm} (3.6)

Equation (3.6) suggests two important operation modes which are useful in this project: the saturation mode and the conduction mode.

The transistor can be considered to be saturated if $V_D - V_S > 5U_T$ in which case $1 - \exp[-(V_D - V_S)/U_T] > 99.33\%$. The last exponential term can be ignored and the drain current is independent of the drain voltage $V_D$. Increasing $V_D - V_S$ increases the accuracy of the drain current. In this case the transistor can be modeled as a current source which is use in OTA design.

Another case is when $V_D - V_S$ is small ($< 5U_T$) and the drain current is strongly dependant upon the drain voltage $V_D$. Under such a circumstance the transistor is
modeled as a pseudo-resistor. The current-biased pseudo-resistor in this project is intended to be operated in such conduction mode.

The authors of [13] also introduce the concept of “inversion factor” (or “inversion coefficient”) which is defined by normalizing the forward current to the specific current $I_S$:

$$i_f = I_F / I_S.$$  \hfill (3.10)

The purpose of introducing inversion factor is to determine the bias condition of the transistors. By looking into the inversion factor, one is able to know if the transistor is biased in weak inversion region ($i_f \ll 1$), moderate inversion region ($i_f = 1$) or strong inversion region ($i_f \gg 1$).

The expression of gate transconductance in weak inversion is very different from that in strong inversion region, as can be obtained by taking the derivative of $I_D$ with regard to $V_G$:

$$g_{mg} = \frac{\partial I_D}{\partial V_G} = \frac{I_D}{nU_T}.$$  \hfill (3.11)

Equation (3.11) suggests that the transconductance is independent of the geometry of the transistor, as long as it’s safely in weak inversion region. This expression will be used to calculate the bias currents for the subthreshold OTAs in the following sections.

By normalizing the transconductance of a transistor in saturation (in all operation regions including strong inversion, moderate inversion and weak inversion) to its value in weak inversion, one is able to obtain the relationship between the normalized transconductance and the inversion factor, as showing in (3.12) and Fig. 3.1. Notice that in forward saturation region, $I_D = I_F$. 

29
\[ G(i_f) = \frac{g_{m}}{I_F} \cdot n \cdot U_T = \frac{1 - \exp\left[\frac{-\sqrt{i_f}}{\sqrt{i_f}}\right]}{\sqrt{i_f}} \]  

(3.12)

In (3.12) and Fig. 3.1, \( G(i_f) \propto g_m/I_D \), and it shows that \( g_m/I_D \) value decreases as the bias point is moving from weak inversion region to strong inversion region, and it approaches its maximum value in weak inversion, in other words, transistors biased in weak inversion region are more power efficient.

![Graph showing transconductance from weak to strong inversion.](image)

**Figure 3.1, Transconductance from weak to strong inversion.**

### 3.2 Filter Configurations - Inverting and None-inverting

Fig. 3.2 shows two different configurations associated with filters in negative feedback networks – the inverting configuration and the non-inverting configuration. Both configurations are voltage-voltage (series-shunt) feedback. Assuming the second
pole (the first non-dominant pole) of the OTA is far away from the dominant pole, the OTA is modeled as a single pole system.

![ OTA schematic diagram ]

(a)  (b)

Figure 3.2, Inverting configuration (a) and non-inverting configuration (b).

For the inverting configuration in Fig. 3.2 (a), the following equations can be formulated:

\[
\frac{v_o - v_f}{Z_F} = \frac{v_f - v_i}{Z_I},
\]

(3.13)

\[-v_f A(s) = v_o. \]

(3.14)

Solving (3.13) and (3.14) gives the transfer function of the inverting filter:

\[
\frac{v_o}{v_i} = \frac{Z_F}{Z_I} \cdot \frac{1}{1 + \frac{1}{A(s)} + \frac{Z_F}{Z_I} \frac{1}{A(s)}}.
\]

(3.15)

Since the OTA is modeled as a single pole system, its transfer function can be expressed as \( A(s) = \frac{A}{1 + \frac{s}{\omega_{3dB}}} \) where \( A \) is the DC gain, \( \omega_{3dB} = 1/\tau \) is the 3-dB bandwidth and \( \text{GBP}_{\text{OTA}} = A/\tau \) is the gain-bandwidth product of the OTA. Substituting the transfer function of the OTA into (3.15), one can rewrite the transfer function of the filter as:
\[
\frac{v_o}{v_i} = -\frac{Z_F}{Z_I} \cdot \frac{A}{1 + A + \frac{Z_F}{Z_I}} \cdot \frac{1}{1 + \frac{Z_F}{Z_I} \tau s}.
\] (3.16)

For the non-inverting configuration in Fig. 3.2 (b):

\[(v_i - v_f)A(s) = v_o,\] (3.17)

\[v_f = \frac{Z_I v_o}{Z_I + Z_F},\] (3.18)

And from (3.17) and (3.18), the transfer function is found to be:

\[\frac{v_o}{v_i} = \frac{Z_F + Z_I}{Z_I} \cdot \frac{A}{A + 1 + \frac{Z_F}{Z_I}} \cdot \frac{1}{1 + \frac{Z_F + Z_I}{Z_I} \tau s}.
\] (3.19)

If both \(Z_I\) and \(Z_F\) are resistors \((R_I\) and \(R_F\)), i.e. not frequency dependent, the configurations showing in Fig. 3.2 (a) and (b) both have one pole and their 3-dB bandwidths can be expressed by (3.20) and (3.21) respectively, assuming \(A \gg 1\):

\[s_{3dB-inv} = -\frac{A}{\tau} \cdot \frac{1 + (1 + R_F / R_I)(1/A)}{1 + R_F / R_I} \approx -\frac{A}{\tau} \cdot \frac{1}{1 + R_F / R_I},\] (3.20)

\[s_{3dB-noninv} = -\left(\frac{R_F + R_I}{R_I} \cdot \frac{\tau}{A + 1 + R_F / R_I}\right)^{-1}.
\] (3.21)

And the closed-loop gains for both configurations are given by (3.22) (inverting) and (3.23) (non-inverting) respectively, assuming \(A \gg 1\):

\[A_{CL-inv} = -\frac{R_F}{R_I} \cdot \frac{A}{1 + A + R_F / R_I} \approx -\frac{R_F}{R_I},\] (3.22)

\[A_{CL-noninv} = \frac{R_F + R_I}{R_I} \cdot \frac{A}{1 + A + R_F / R_I} \approx \frac{R_F + R_I}{R_I}.
\] (3.23)
Synthesizing the results from (3.20) to (3.23), the gain-bandwidth-products of the inverting configuration (GBP_{inv}) and the non-inverting configuration (GBP_{non-inv}) are given by (3.24) and (3.25) respectively:

\[
\text{GBP}_{\text{inv}} = A_{\text{CL-inv}} s_{3\text{dB-inv}} = \frac{A}{\tau} \cdot \frac{R_F / R_I}{1 + R_F / R_I} = \text{GBP}_{\text{op-amp}} \frac{A_{\text{CL-inv}}}{1 + A_{\text{CL-inv}}}, \quad (3.24)
\]

\[
\text{GBP}_{\text{non-inv}} = A_{\text{CL-noninv}} s_{3\text{dB-noninv}} = \frac{A}{\tau} = \text{GBP}_{\text{op-amp}}. \quad (3.25)
\]

Equations (3.24) and (3.25) reveal the fact that inverting configuration has a reduced GBP (by the factor of \(A_{\text{CL-inv}}/(1+A_{\text{CL-inv}})\)) as compared to the non-inverting case, if they both use the same OTA.

In this project, the design of the neural amplifier filter is based on the capacitive coupling topology proposed by Harrison in [28], which is an inverting configuration as in Fig. 3.2 (a), with \(Z_I\) replaced by a capacitor and \(Z_F\) replaced by a parallel combination of a resistor and a capacitor, as shown in Fig. 3.3 (a). Harrison et al use a “MOS-bipolar pseudo-resistor” as \(R_2\).

The transfer function of the configuration in Fig. 3.3 (a) can be obtained by substituting the corresponding circuit components for \(Z_I\) and \(Z_F\) in (3-16), or uses the small signal model in Fig. 3.3 (b). For the small signal model in Fig. 3.3 (b), the following equations are formulated according to KCL:

\[
v_m s C_{gs} + (v_m - v_i) s C_1 = (v_o - v_m) \left( \frac{1}{R_2} + s C_2 \right), \quad (3.26)
\]

\[
g_m v_m + v_o \left( \frac{1}{r_o} + s C_L \right) = (v_o - v_m) \left( \frac{1}{R_2} + s C_2 \right). \quad (3.27)
\]

Solving (3.26) and (3.27) gives the transfer function of the filter:
\[
\frac{v_o}{v_i} = \frac{sC_1(sC_2 + G_2 - g_m)}{sC_1(sC_2 + G_2 - g_m) + sC_2(sC_L + sC_g + g_m) + sC_L(sC_g + G_2) + sC_g(G_2 + g_o) + G_2 g_m}, \tag{3.28}
\]

where \(G_2 = 1/R_2\) and \(g_o = 1/r_o\).

The numerator reveals that the system consists of two zeros, \(s_{01} = 0\) and \(s_{02} = (g_m - G_2)/C_2\) where \(s_{02}\) is at a relatively high frequency.

Observing that \(g_m > g_o\) and \(g_o > G_2\), the denominator can be simplified and the following equation can be formulated to solve for the poles:

\[
\frac{s^2\left[C_1(C_2 + C_L) + C_2(C_L + C_{gs}) + C_L C_{gs}\right]}{G_2 g_m} + \frac{sC_2 g_m}{G_2 g_m} + 1 = 0. \tag{3.29}
\]

And the poles are:

\[
s_1 = -\frac{G_2}{C_2}, \tag{3.30}
\]

\[
s_2 = -\frac{g_m}{K(C_2 + C_L) + C_L + C_{gs} + C_L C_{gs}/C_2}, \tag{3.31}
\]

where \(K = C_1/C_2\) is the closed-loop gain of the filter.

Figure 3.3, Capacitive coupling amplifier filter (a) and its small signal model (b).
From the locations of the zeros and poles, it is obvious that the configuration showing in Fig. 3.3 (a) is a band-pass filter whose lower cutoff frequency is set by \( s_1 \) and upper cutoff frequency is set by \( s_2 \). The topology showing in Fig. 3.3 (a) is actually a high pass configuration if the OTA has an infinite bandwidth; however, the limited bandwidth of a practical OTA makes it a bandpass filter. If \( R_2 = 1.59 \text{ G}\Omega \) and \( C_2 = 200 \text{ fF} \), the lower cutoff frequency can be calculated to be 500.5 Hz.

Assuming \( C_L \gg C_2 \gg C_{gs} \), the gain bandwidth product of the filter can be calculated as:

\[
\text{GBP}_{\text{filter}} = \frac{K_{gm}}{K(C_2 + C_L) + C_L + C_{gs} + C_L C_{gs}} \approx \frac{K}{K + 1} \frac{g_m}{C_L} = \frac{K}{K + 1} \cdot \text{GBP}_{\text{op-amp}}, \tag{3.32}
\]

which is consistent with the conclusion from (3.24). If \( C_{gs} \) is large enough to be comparable with \( C_2 \) or even \( C_L \), the gain bandwidth product is degenerated.

### 3.3 Subthreshold OTA Design

As discussed in section 3.1, transistors biased in weak inversion region have a high \( g_m/I_D \) value and hence are suitable for low power applications due to its power efficiency. In this work, all the OTAs are biased in weak inversion or moderate inversion region due to the strict power budget in this project.

Assuming the application bandwidth of the filter is \( BW \), and the closed-loop gain of the filter is \( A_{CL} \), the gain-bandwidth-product of the neural filter (GBP_{filter}) is defined as:

\[
\text{GBP}_{\text{filter}} = A_{CL} \cdot BW. \tag{3.33}
\]

And the relationship between the gain-bandwidth-product of the OTA (GBP_{OTA}) and the gain-bandwidth-product of the filter in an inverting configuration is expressed in (3.24).
Assuming a single stage OTA, its transconductance is determined by its GBP and the loading:

\[
\text{GBP}_{\text{OTA}} = \frac{g_m}{2\pi C_{\text{Leff}}},
\]

(3.34)

where \(C_{\text{Leff}}\) is the effective load capacitance to the OTA. For an OTA biased in weak inversion region, \(g_m = \frac{I_D}{nU_T}\) and the biased current is found to be:

\[
I_D = 2\pi \cdot n \cdot U_T \cdot C_{\text{Leff}} \cdot \text{GBP}_{\text{OTA}}.
\]

(3.35)

Or in terms of the application bandwidth:

\[
I_D = 2\pi \cdot n \cdot U_T \cdot C_{\text{Leff}} \cdot (A_{\text{CL}} + 1) \cdot \text{BW}.
\]

(3.36)

Equation (3.36) can be used to estimate the bias current for the OTA. Equations (3.35) and (3.36) also indicate that the gain-bandwidth-product of an OTA or a filter in weak inversion region is limited due to its very small bias current \((I_D)\) in subthreshold region. It is also interesting to notice that the transconductance is independent of device geometry in weak inversion region. However, the length of the transistor is limited by the transition frequency (or unity current gain frequency) \(f_T\) in weak inversion region [29]:

\[
f_T = \frac{\mu U_T}{2\pi L^2}.
\]

(3.37)

The output resistance of a transistor in weak inversion can be expressed as:

\[
r_o \cong \frac{1}{\lambda I_D},
\]

(3.38)

and hence its self-gain (dc gain) is:

\[
A_{\text{vo}} = \frac{1}{nU_T \lambda}.
\]

(3.39)
Another important characteristic of an OTA is its slew rate, which can be calculated as:

\[ \text{SR} = \frac{I_{SS}}{C_{\text{eff}}} = \frac{2I_D}{C_{\text{eff}}} = 2 \cdot \text{GBP}_{\text{OTA}} \cdot n \cdot U_T. \] (3.40)

Equation (3.40) assumes an OTA with differential inputs whose tail current is \( I_{SS} \).

3.4 Voltage Biased Pseudo-Resistor and Current Biased Pseudo-Resistor

For the capacitive coupling topology in Fig. 3.3 (a), if the linear resistor is replaced by the “MOS-bipolar pseudo-resistor” described in [28], the lower cutoff frequency is below 1 Hz due to the very high resistance of this type of pseudo-resistor. As discussed in chapter I, the lower cutoff frequencies for some neural signals could be in the range of tens of hertz or hundreds of hertz, and a pseudo-resistor biased in weak inversion region or moderate inversion region could be used to achieve desired cutoff frequencies. An additional benefit from a pseudo-resistor biased by external signals (either a bias voltage or a bias current) lies in its tunability.

Fig. 3.4 shows two different bias schemes to bias up the pseudo-resistor – gate voltage biasing and current biasing. In the voltage bias scheme, some researchers replace a single n-FET or a single p-FET by a series combination of an n-FET and a p-FET for a better linearity [7;8].

The disadvantage of the gate voltage biased pseudo-resistor is that its effective resistance is hard to control. As discussed in chapter 2, assuming the drain-to-source voltage equals zero, the effective resistance is inversely proportional to its overdrive voltage \( r_{\text{eff}} \propto (V_{GS} - V_t)^{-1} \) in strong inversion region and exponentially dependent on its
gate voltage in weak inversion region \( r_{\text{eff}} \propto \left[ \exp\left( \frac{V_G - V_t - nV_S}{nU_T} \right) \right]^{-1} \). Due to process variation, the threshold voltage \( V_t \) will change across process corners and the output common mode voltage \( V_B \) will also change due to the OTA offset. As a result, the gate voltage has to be tuned accordingly to achieve the desired effective resistance. The situation becomes even worse when the pseudo-resistor is pushed into weak inversion region when there is an exponential dependence on the gate-to-source voltage.

![Figure 3.4, Voltage biasing method (a) and current biasing method (b).](image)

Now consider the current bias method in Fig. 3.4 (b). The diode connected transistor \( M_B \) supports the bias voltage for \( M_R \), and the bias voltage is established by the current \( I_B \). If \( I_B \) is constant, the gate-to-source voltages \( V_{GB} \) for \( M_B \) and \( M_R \) are both constants. The gate-to-source voltage of the pseudo-resistor is also independent of the output common voltage of the OTA. If the output common mode voltage \( V_B \) shifts up or down, the gate...
voltage ($V_G$) follows it to maintain a constant gate-to-source voltage. This means that the effective resistance is not affected by the OTA offset or the output common mode voltage shift. The following sections will give a more detailed discussion of the characteristics of the current bias pseudo-resistor, and it will be shown that the effective resistance is independent of the gate bias voltage and the accuracy of the effective resistance is significantly improved.

3.5 Characterization of Current Biased Pseudo-Resistor

Current biased pseudo-resistors operated in the subthreshold region were first introduced in [30] for the calculation of harmonic mean. In this work, we introduce the idea of current biased pseudo-resistor into the design of MOSFET-C filters. This section details the linearity of the pseudo-resistor, the frequency response of the pseudo-resistor bias and the noise of the pseudo-resistor.

3.5.1 Linearity

Unlike a real resistor, pseudo-resistors are highly nonlinear and their equivalent resistance depend upon their drain-to-source voltages $V_{DS}$. This may be the worst disadvantage associated with pseudo-resistors. The test setup in Fig. 3.5 simulates the dependence of the equivalent resistance on $V_{DS}$ for a current biased pseudo-resistor. In this simulation a constant current $I_B$ is injected into the diode connected transistor $M_B$ and the drain voltage of $M_R$ ($V_{dc}$) is swept from -100 mV to 100 mV to observe the change of its drain current $I_D$. As illustrated in Fig. 3.6, the drain current of $M_R$ ($I_D$) changes significantly with $V_{dc}$. $I_D$ can also be changed by $I_B$. The simulation result is obtained for
two different bias currents ($I_B = 23$ pA and $I_B = 46$ pA) and the transistors are sized at $W/L = 4 \ \mu m / 6 \ \mu m$.

Pseudo-resistor nonlinearity introduces harmonic distortions into the filter and special techniques have to be implemented to reduce the effect of nonlinearity. The equivalent resistance as a function of the drain-to-source voltage is plotted in Fig. 3.7.

![Figure 3.5, Current biased pseudo-resistor with a swept drain-to-source voltage.](image)

Figure 3.5, Current biased pseudo-resistor with a swept drain-to-source voltage.

![Figure 3.6, $I_D$ vs. $V_{dc}$ plots for different bias currents $I_B$.](image)

Figure 3.6, $I_D$ vs. $V_{dc}$ plots for different bias currents $I_B$. 

40
Fig. 3.7 shows that the equivalent resistance of a pseudo-resistor is a strong function of the drain-to-source voltage; however, if a pseudo-resistor is introduced into a system such as a MOSFET-C filter, its linearity has to be studied with the rest parts of the system. Studying the linearity of a pseudo-resistor in a system yields a more meaningful result. Consider Fig. 3.8 where a single ended high-pass filter is compared to a fully differential high-pass filter. The advantage of a fully differential configuration over its single ended counterpart lies in the ability to cancel out the even order distortion terms.

![Figure 3.7, Equivalent resistance vs. drain voltage plots for different bias currents $I_B$.](image)

In Fig. 3.8, the bias current $I_B$ is not actually mirrored by $M_R$ because there is no DC current flowing into $M_R$. Small signal resistance is derived from the large signal bias current, hence, an imaginary current $I_D$ is assumed to be flowing into $M_R$. For a large time constant, $M_B$ and $M_R$ are both biased in the weak inversion region where $M_B$ is in forward saturation mode ($V_G - V_T > 5U_T$) and $M_R$ is in conductance mode (subthreshold triode...
region, \( V_A = V_B \). According to the model introduced in (3.1), their drain currents can be expressed as [13]:

\[
I_D = I_{D0R} e^{(V_{th} + V_{GB})/nU_T} (e^{-V_B/U_T} - e^{-V_A/U_T}),
\]

(3.41)

\[
I_B = I_{D0B} e^{(V_{th} + V_{GB})/nU_T} e^{-V_B/U_T},
\]

(3.42)

where \( n \) is the subthreshold slope parameter and \( U_T \) is the thermal voltage.

![Figure 3.8 High-pass filters with current biased pseudo-resistors: (a) Single ended configuration, (b) Fully differential configuration.](image)

Pseudo-resistor nonlinearities can be analyzed in a Taylor series which shows all harmonic distortion terms [24]. In order to watch all the higher order terms of \( I_D \), (3.41) is expanded in a Taylor series with respect to \( V_B \) at the point \( V_A \):

\[
I_D = f(V_B)|_{V_B = V_A}
\]

\[
= f(V_A) + \frac{f'(V_A)}{1!} (V_B - V_A) + \frac{f''(V_A)}{2!} (V_B - V_A)^2 + \frac{f'''(V_A)}{3!} (V_B - V_A)^3 + ...
\]

(3.43)
where

\[ f^{(m)}(V_A) = \frac{I_{D0B}}{n^m U_T^m} \cdot [(1 - n)^m - 1] \cdot e^{-\left(V_{gb} + V_A (1 - n) \right) / n U_T}. \]

(3.44)

Rearrange (3.42) and let \( V_B = V_A \), the exponential term in (3.44) can be expressed as:

\[ e^{-\left(V_{gb} + V_A (1 - n) \right) / n U_T} = \frac{I_B}{I_{D0B}}. \]

(3.45)

With the aid of (3.43), (3.44) and (3.45), \( I_D \) can be expressed as the sum of different order terms:

\[ I_D = \sum_{m=0}^{\infty} \frac{I_{D0B} I_B}{m! I_{D0B} n^m U_T^m} \cdot (1 - n)^m \cdot (V_B - V_A)^m. \]

(3.46)

Observing that \( I_{D0B} \equiv 2n\mu C_{ox} S_B U_T^2 e^{-\left(V_{gb} \right) / n U_T} \) and \( I_{D0R} \equiv 2n\mu C_{ox} S_R U_T^2 e^{-\left(V_{ta} \right) / n U_T} \), (3.46) can be further simplified to be:

\[ I_D = \sum_{m=0}^{\infty} \frac{S_R I_B}{m! S_R n^m U_T^m} \cdot (1 - n)^m \cdot (V_B - V_A)^m, \]

(3.47)

where \( S_B \) and \( S_R \) are the aspect ratios (W/L) of \( M_B \) and \( M_R \), respectively. The effective resistance is then given by the first order term as:

\[ R_{eq} = \frac{S_B U_T}{S_R I_B}, \]

(3.48)

and higher order terms represents harmonic distortions. Substitute (3.48) into (3.47), \( I_D \) can be expressed as:

\[ I_D = \sum_{m=0}^{\infty} \frac{1}{m! R_{eq} n^m U_T^{m-1}} \cdot (1 - n)^m \cdot (V_B - V_A)^m. \]

(3.49)
Equation (3.48) removes the exponential dependence between the effective resistance and the gate bias voltage; instead, it introduces an inverse proportional dependence between the effective resistance and the bias current $I_B$. Inverse proportion is a weaker function and the effective resistance is less affected by the external bias conditions. Now its accuracy is set by the matching between $M_R$ and $M_B$, the accuracy of $I_B$, and the accuracy of the temperature. The bias current $I_B$, which is usually generated by a beta-multiplier reference circuit, normally has a 20% error across process corners. This error comes from the deviation of the resistor used in the beta-multiplier circuit. As the chips will be implanted in animal bodies, temperature deviations can be ignored. The generation of the bias current $I_B$ will be discussed later.

In order to find the effect of the pseudo-resistor nonlinearity on the output voltage of the filter, an ideal OTA is introduced to simplify the problem, as illustrated in Fig. 3.9. An ideal OTA has such characteristics as infinite open loop gain, infinite input impedance and virtual ground inputs. With an ideal OTA, the following current and voltage relationships can be found in Fig. 3.9:

\[ I_{c1} = I_{c2} + I_{\text{Req}}, \]  
\[ I_{c1} = V_{in} sC_1, \]  
\[ I_{c2} = -V_{out} sC_2. \]

Substituting (3.51) and (3.52) into (3.50) yields the relationship between $V_{out}$ and $I_{\text{Req}}$:

\[ V_{out} = -\frac{V_{in} C_1}{C_2} + \frac{I_{\text{Req}}}{sC_2}. \]
Noticing \( I_D = I_{Req} \), \( V_B = V_{out} \), and \( V_A = 0 \) for an ideal OTA, substitute (3.49) into (3.53) gives:

\[
V_{out} = -\frac{V_{in} C_1}{C_2} + \frac{1}{s C_2} \sum_{m=0}^{\infty} \frac{1}{m! R_{eq} n^m U_T^{m+1}} [(1 - n)^m - 1] V_{out}^m. \tag{3.54}
\]

Rearranging (3.54) yields:

\[
V_{out} = V_{out,ideal} + \frac{1}{1 + s C_2 R_{eq}} \sum_{m=2}^{\infty} \frac{1}{m! R_{eq} n^m U_T^{m+1}} [(1 - n)^m - 1] V_{out}^m, \tag{3.55}
\]

where \( V_{out,ideal} = -\frac{C_1}{C_2} s C_2 R_{eq} V_{in} \) is the ideal output voltage when a linear feedback resistor is used. The ratios between a higher order term and the ideal output voltage give the percentage distortion of the output signal. When the distortion is small, \( V_o \approx V_{o,ideal} \) and the second order distortion is:

\[
\left| \frac{1}{1 + s C_2 R_{eq}} \frac{(1 - n)^2 - 1}{2 n^2 U_T} V_{out}^2 \right| \approx \left| \frac{1}{\sqrt{1 + \left(\frac{f}{f_L}\right)^2}} \frac{(1 - n)^2 - 1}{2 n^2 U_T} V_{out} \right|. \tag{3.56}
\]

The third order distortion is:

\[
\left| \frac{1}{1 + s C_2 R_{eq}} \frac{(1 - n)^3 - 1}{6 n^3 U_T^2} V_{out}^3 \right| \approx \left| \frac{1}{\sqrt{1 + \left(\frac{f}{f_L}\right)^2}} \frac{(1 - n)^3 - 1}{6 n^3 U_T^2} V_{out}^2 \right|. \tag{3.57}
\]
Equations (3.56) and (3.57) show that the distortion is a function of both the frequency and the output signal swing. The distortions are lower at higher frequencies and lower output swings.

The fact that higher frequencies yield lower distortion is mainly due to the current path formed by the parallel combination of the capacitor and the pseudo-resistor. At higher frequencies, more current passes through the linear capacitor instead of the nonlinear pseudo-resistor and the contribution of the pseudo-resistor nonlinearity to the output voltage is lower.

The fact that higher output signal swings yield higher distortions indicate that the filter can only handle small signal filtering for a reasonable distortion.
For the neural signal filter with a passband from 500 Hz to 5 KHz (which is the frequency range for extracellular action potential spikes), the distortion can be estimated at the center frequency of 1.58 KHz. Assuming $n = 1.4$ for a typical case and $U_T = 26$ mV at room temperature, if the distortion is required to be 1%, the output voltage $V_{out}$ is found to be 4 mV for the second order distortion and 18.62 mV for the third order distortion. Observing that the output swing is $V_{out_{\text{max}}} - V_{out_{\text{min}}}$, it is found to be 8 mV for the single ended configuration in Fig. 3.8(a); also noticing that a fully-differential circuit is able to cancel out all the even order distortion terms, the output swing is found to be 74.5 mV for the fully differential configuration in Fig. 3.8(b). Since the amplitude of extracellular neural action potential spikes is usually less than 500 µV before amplification [2], and 50 mV after a 40 dB amplification, a 74.5 mV output swing is usually enough for most applications.

### 3.5.2 Tunability

The equivalent resistance showing in (3.48) is a function of $I_B$, which indicates that it is tunable by varying the bias current $I_B$. Observing that the lower cutoff frequency $f_L = 1/ (2 \pi R_{eq} C_2)$, and using the expression in (3.48), the lower cutoff frequencies of the filters in Fig. 3.8 can be expressed as:

$$f_L = \frac{S_R I_B}{2 \pi C S_B U_T}.$$  \hspace{1cm} (3.58)

Equation (3.58) shows that the lower cutoff frequency is linearly proportional to the bias current. It is also a function of $S_R / S_B$ and $U_T$, indicating that for an accurate cutoff frequency, $M_R$ and $M_B$ have to be well matched and the temperature has to be stable. Matching $M_R$ and $M_B$ can be achieved by a proper layout scheme such as breaking each
of them into several fingers and using common-centroid layout pattern. A stable temperature is easily available since for an implantable application, the chips are implanted in the body of an animal whose temperature is usually stable.

3.5.3 Frequency Response of the Pseudo-Resistor Bias

The above analysis assumes that $V_{GB}$ is a constant, however in practical cases it is a function of the frequency. In the case when $V_{GB}$ changes with frequency, additional distortions will be introduced into the circuit. In order for $V_{GB}$ to be a constant, the gain from $V_B$ to $V_G$ has to be unity with zero degree phase shift. Fig. 3.10 is a small signal model of the configuration in Fig. 3.8 (a).

The transfer function from $V_b$ to $V_g$ is found to be:

$$v_g = \frac{v_b \left[ s \left( C_{gs,MR} + C_{gs,MB} \right) + g_{m,MB} + \frac{1}{r_{o,MB}} \right]}{s \left( C_{gs,MR} + C_{gs,MB} + C_{db,MB} + C_{db,MC} + C_{gd,MR} \right) + g_{m,MB} + \frac{1}{r_{o,MB}} \frac{1}{r_{o,MC}}} \quad (3.59)$$

Figure 3.10, The small signal model for Fig. 3.8 (a).
The transfer function contains a left half-plane (LHP) pole and a LHP zero:

The zero is found from the numerator:

\[ s = \frac{g_{m,MB} + \frac{1}{r_{o,MB}}}{C_{gs,MR} + C_{gs,MB}}, \]  

Equation (3.60)

And from the denominator, the pole is found to be:

\[ s = \frac{g_{m,MB} + \frac{1}{r_{o,MB}} + \frac{1}{r_{o,MC}}}{C_{gs,MR} + C_{gs,MB} + C_{db,MB} + C_{db,MC} + C_{gd,MR}}. \]  

Equation (3.61)

The simulation result for the frequency response of the pseudo-resistor bias is given in Fig. 3.11.

Figure 3.11, Frequency response of the pseudo-resistor bias – gain and phase margin.
Fig. 3.11 verifies the fact that the transfer function consists of a LHP pole and a LHP zero. Equations (3.60) and (3.61) indicate that the bias current $I_B$ (and hence $g_{m,MB}$) has to be large enough to push the pole and the zero out of the application band. Equation (3.59) indicates that the gain is a little lower that unity but the simulation result shows that the gain is slightly higher that unity at low frequencies, this inconsistency may be due to the approximation made in deriving (3.59).

3.5.4 Noise Analysis of the Pseudo-Resistor

Observing that the feedback network in Fig. 3.8 is a parallel combination of the feedback capacitor $C_2$ and the feedback pseudo-resistor $M_R$, the noise contributed by the pseudo-resistor could be predicted to be “kT over C” noise.

The noise equivalent bandwidth of the feedback network consisting of $C_2$ and $M_R$ is [31]:

$$NEB = \frac{1}{4R_{eq}C_2}.$$  \hfill (3.62)

The thermal noise conductance of the pseudo-resistor is given by [13]:

$$G_{Nth} = g_{ms} = \left. \frac{\partial I_D}{\partial V_s} \right|_{V_G,V_S} = \frac{I_F}{U_T},$$ \hfill (3.63)

where $g_{ms}$ is the source conductance. Observing that $M_B$ and $M_R$ have the same $V_G$ and $V_S$, the relationship between their forward currents $I_{FB}$ and $I_{FR}$ is:

$$I_{FB} = \frac{S_B}{S_R} I_{FR}.$$ \hfill (3.64)

And for $M_B$ which is in forward saturation mode:
\[ I_B = I_{FB}. \quad (3.65) \]

With the aid of (3.63), (3.64) and (3.65), the thermal noise conductance is found to be:

\[ G_{\text{Nth}} = \frac{S_R I_B}{S_B U_T}. \quad (3.66) \]

So the output rms noise contributed by the pseudo-resistor is:

\[ V_{\text{rms}} = \sqrt{\frac{1}{4R_{eq}C_2} \cdot 4kT \cdot \frac{S_B U_T}{S_R I_B}} = \sqrt{\frac{kT}{C_2}}. \quad (3.67) \]

Equation (3.67) verifies that the thermal noise of the feedback network consisting of \( C_2 \) and the pseudo-resistor is \( kT/C \) noise. As the noise spectrum density starts to roll off at the lower cutoff frequency \( \frac{1}{2\pi R_{eq}C_2} \), most of its power is outside the application band.

The flicker noise (1/f noise) is still not well understood, and it is modeled as a voltage source in series with the gate of the pseudo-resistor [32]:

\[ V_g^2(f) = \frac{K}{WLC_{ox}f}, \quad (3.68) \]

where \( K \) is a constant varying from process to process, \( W, L \) and \( C_{ox} \) are the transistor width, length and gate capacitance per unit area. 1/f noise only occurs when a DC current is flowing in the device. The pseudo-resistors in Fig. 3.8 don’t carry any DC current and hence their 1/f noise is zero.

3.6 Noise Analysis of the Neural Filter

The input-referred noise of the neural filter is contributed both by the pseudo-resistor and the OTA. Since the noise from the pseudo-resistor can be divided by the gain of the
filter, the total noise of the filter is mainly due to the OTA. The OTA used in this project is a folded cascode OTA as shown in Fig. 3.12.

The noise of the OTA mainly comes from M2, M3, M4, M5, M10 and M11. Assuming the OTA is actually biased in the weak inversion region, the spectral density of the thermal noise current can be expressed as [13]:

\[ I_d^2(f) = 2nkT g_m, \]  

(3.69)

and the spectral density of the input referred thermal noise voltage is:

\[ V_{n,\text{thermal}}^2(f) = \frac{4nkT}{g_{m2}} \left( 1 + \frac{g_{m4}}{g_{m2}} + \frac{g_{m10}}{g_{m2}} \right), \]  

(3.70)

Figure 3.12, The folded cascode OTA used in this project.
Since $g_m = \frac{I_D}{nU_T}$ for transistors operating in the subthreshold region, the relationships between $g_{m2}$, $g_{m4}$ and $g_{m10}$ can be found to be $2g_{m2} = 2g_{m4} = g_{m10}$ and (3.70) can be further simplified to be:

$$V_{ni,thermal}^2(f) = \frac{16nkT}{g_{m2}}. \tag{3.71}$$

The 1/f noise is modeled by (3.68) and it can also be written as the noise current spectral density:

$$I_{d,flicker}^2(f) = \frac{g_m^2K}{WLC_{ox}f}. \tag{3.72}$$

And the spectral density of the input referred 1/f noise voltage is:

$$V_{ni,flicker}^2(f) = \frac{2}{C_{ox}f} \left[ \frac{K_2}{W_2L_2} + \frac{K_4}{W_4L_4} \left( \frac{g_{m4}}{g_{m2}} \right)^2 + \frac{K_{10}}{W_{10}L_{10}} \left( \frac{g_{m10}}{g_{m2}} \right)^2 \right]. \tag{3.73}$$

Given $2g_{m2} = 2g_{m4} = g_{m10}$, (3.73) becomes:

$$V_{ni,flicker}^2(f) = \frac{2}{C_{ox}f} \left[ \frac{K_2}{W_2L_2} + \frac{K_4}{W_4L_4} + \frac{4K_{10}}{W_{10}L_{10}} \right]. \tag{3.74}$$

So the input referred OTA noise voltage spectral density is:

$$V_{ni,OTA}^2 = V_{ni,thermal}^2 + V_{ni,flicker}^2. \tag{3.75}$$

$$V_{ni,OTA}^2 = \frac{16nkT}{g_{m2}} + \frac{2}{C_{ox}f} \left[ \frac{K_2}{W_2L_2} + \frac{K_4}{W_4L_4} + \frac{4K_{10}}{W_{10}L_{10}} \right]. \tag{3.76}$$

Referring to Fig. 3.13, the input referred noise spectral density for the filter is [33]:
Substituting (3.76) into (3.77):

\[ V_{n_{i,\text{filter}}}^2 = \left( \frac{C_1 + C_2 + C_{gs}}{C_1} \right)^2 V_{n_{\text{OTA}}}^2. \] (3.77)

Equation (3.78) reveals the fact that: 1) Increasing the ratio \( C_1/C_2 \) reduces the input referred noise; 2) Increasing the transconductance (here it is \( g_{m2} \)) of the input pairs reduces the input referred noise; 3) 1/f noise dominates in the low frequency range and increasing the device sizes helps reducing 1/f noise; 4) Increasing the size of the input pair also increases the gate capacitance \( C_{gs} \), which leads to an increase in the input referred noise of the filter, offsetting some of the 1/f noise reduction.

![Figure 3.13, Noise model for a capacitive coupling filter.](image-url)
3.7 Bias Current Generation

Beta-multiplier reference circuits are widely used for reference current generation in CMOS process due to their good immunity to power supply deviations [34]. As shown in Fig. 3.14, $\beta_2$ are usually $K$ times larger than $\beta_1$ so that it is called a beta-multiplier circuit. Beta-multiplier circuits can work in both strong inversion region and weak inversion region.

Fig. 3.14 is a standard beta-multiplier reference circuit which will be used to explain its operational principle. This circuit can be modified to satisfy more advanced applications.

In strong inversion region, the following relationship can be found [34]:

$$V_{GS1} = V_{GS2} + I_{ref} R.$$  \hfill (3.79)

Figure 3.14, Beta-multiplier reference circuit.
Observing that \( V_{GS} = \sqrt{\frac{2I_{ref}}{\beta}} + V_m \) and setting \( \beta_2 = K\beta_1 \), the reference current \( I_{ref} \) can be expressed as:

\[
I_{ref} = \frac{2}{R^2 \beta_1} \left( 1 - \frac{1}{\sqrt{K}} \right)^2 ,
\]

which is independent of the supply voltage \( V_{dd} \). It can be shown that the transconductance \( g_m \) is also proportional to \( 1/R \) which is not a function of the MOSFET process shifts (constant \( g_m \)).

For small currents the circuit can be operated in weak inversion region and the EKV model can be used to analyze it [35].

In weak inversion region, the EKV model models the drain current (the reference current \( I_{ref} \)) in forward saturation as:

\[
I_{ref} = 2n\beta U_T^2 \exp \left( \frac{V_{G1} - V_{T0} - nV_S}{nU_T} \right) .
\]

Rearrange (3.81) for both \( M_1 \) and \( M_2 \):

\[
V_{G1} - V_{T0} - nV_{S1} = nU_T \ln \frac{I_{ref}}{2n\beta U_T^2} ,
\]

\[
V_{G2} - V_{T0} - nV_{S2} = nU_T \ln \frac{I_{ref}}{2n\beta_2 U_T^2} ,
\]

where \( \beta_2 = K\beta_1 \). Noticing \( V_{G1} = V_{G2} \), \( V_{S2} = V_{S1} + I_{ref} R \), (3.82) – (3.83) yields:

\[
I_{ref} R = U_T \left( \ln \frac{I_{ref}}{2n\beta_1 U_T^2} - \ln \frac{I_{ref}}{2n\beta_2 U_T^2} \right) .
\]

After simplification, (3.84) becomes:

\[
I_{ref} = \frac{U_T}{R} \ln K .
\]
Interestingly, in weak inversion region the reference current is independent of the supply voltage, the MOSFET process parameter \( (\mu C_{ox}) \) and the geometry of the devices. The reference current is set by the beta ratio between \( M_1 \) and \( M_2 \), the resistance in the source and the temperature. Knowing that in weak inversion region, \( g_m = I_D / nU_T \), \( g_m \) is also proportional \( 1/R \) and we still get a constant \( g_m \) which is independent of the MOSFET process shifts.

Beta-multiplier circuits biased in weak inversion region are used in this project for low power consumptions and large time constant applications. Remember that in (3.58) the cutoff frequency is proportional to the bias current for the pseudo-resistor and small current is required for low frequencies. In this work the bias current for the pseudo-resistor is obtained by mirroring down the reference current from the beta-multiplier circuit through several current mirror stages.

3.8 Layout Issues

The layout of an integrated circuit is critical to its performance. However, one has to recognize the fact that it is more challenging to match the drain currents of a pair of MOSFETs inside a current mirror in weak inversion region.

In strong inversion region where the drain current is proportional to the square of the overdrive voltage, the drain current mismatch can be expressed as [36]:

\[
\frac{\Delta I_D}{I_D} = -\frac{2\Delta V_t}{V_{eff}},
\]

(3.86)

where \( V_{eff} = V_{GS} - V_t \) is the overdrive voltage.

However, in weak inversion region, the drain current depends on the gate-to-source voltage exponentially, and the mismatch is expressed as[37]:

\[ \text{Mismatch} = \text{constant} \times \frac{1}{R} \]
\[
\frac{\Delta I_D}{I_D} = -\frac{\Delta V_t}{nU_T}.
\] (3.87)

In most applications, \( V_{eff} > 2nU_T \) (100 mV), which suggests that for the same threshold voltage mismatch \( \Delta V_t \), current mirrors in weak inversion region are expecting higher drain current mismatches than those in strong inversion regions.

In order to achieve better matching, one has to use larger devices and arrange the layout of the current mirror properly. Lakshmikumar et al models the mismatch of the threshold voltage in [38] as:

\[
S_{Vt} = \frac{C_{Vt}}{\sqrt{W_{eff}L_{eff}}},
\] (3.88)

where \( S_{Vt} \) is the standard deviation of the threshold voltage mismatch and \( C_{Vt} \) is a constant. Equation (3.88) suggests larger devices for better matching.

Factors that induce mismatches include oxide thickness gradients, polysilicon etch rate variations, tilted implants, etc [39]. Common-centroid layout is an effective way to reduce gradient-induced mismatches, given that the gradient is linear over distance. This is achieved by aligning the centroids of two devices. In order to do common-centroid layouts, the transistors have to be firstly divided into identical smaller elements, called fingers, and then placed in such an array that the centroids of these two devices coincide. Fig. 3.15 shows the one-dimensional and two-dimensional common-centroid used in this project.

Fig. 3.15 (a) is a one dimensional common-centroid scheme used in this project to lay out current sources, while Fig. 3.15 (b) is a two-dimensional common-centroid scheme used in this project to lay out current sources with a large number of fingers and
the input pairs of the OTAs. Dummy transistors (shorted transistors) are also placed beside the arrays to compensate for polysilicon etch rate variations.

Besides the common-centroid layout technique, matched transistors are also oriented in the same direction (both horizontally or vertically oriented) to reduce orientation-dependent factors that will affect matching. They are also arranged in such a way that the amount of right-oriented fingers equal the amount of left-oriented fingers.

The matching techniques for MOS transistors also apply for capacitor matching and resistor matching.

![Common-centroid layouts](image)

Figure 3.15, Common-centroid layouts used in this project.

3.9 Single-Ended Configuration in a Half-Micron CMOS Bulk Process

The neural filters with the proposed current biased pseudo-resistor are designed and fabricated in two processes: a 0.5 micron CMOS process and a 0.5 micron silicon-on-sapphire CMOS process. This section discusses the filters designed in the former process and those designed in the latter will be discussed in the next section.

The 0.5 micron CMOS process has a power supply of 5 volts and this will be the power supply for our neural amplifier. A single stage single-ended folded-cascode OTA as shown in Fig. 3.12 will be used in this work. One favorable advantage of folded-
cascode is that its input common mode range is independent of its output voltage [40]. When the OTA is biased in weak inversion region, it is able to give a 4.48 V peak-to-peak output swing.

Three different neural filters were fabricated in this work. They are: 1) filter A, the neural filter with a closed-loop gain of 10 V/V and $S_B/S_R = 1$; 2) filter B, the neural filter with a closed-loop gain of 10 V/V and $S_B/S_R = 40$; 3) filter C, the neural filter with a closed-loop gain of 35 V/V and $S_B/S_R = 1$. Following the design procedures discussed in section 3.3, the OTA for the 10 V/V closed-loop gain filter is biased at 60 nA ($I_{SS}$) and the OTA for the 35 V/V closed-loop gain filter is biased at 187 nA. P-type input pairs are used for the OTAs because they have lower 1/f noise. As the OTAs are biased in weak inversion region, relatively large transistors are used in the OTAs for better matching, and the length for the devices is 6 μm.

Fig. 3.16 shows the current biased pseudo-resistors for $S_B/S_R = 1$ and $S_B/S_R = 40$, the layout of the latter is also shown in the figure.

The layout showing in Fig. 3.16 includes 14 fingers in two rows. The grey ones represent the diode connected transistor (4 W/L) and the white ones represent the pseudo resistor. The grey ones are connected in parallel and the white ones in series to achieve a 40:1 ratio.

The bias current $I_B$ is generated by the same beta multiplier reference circuit as the OTA and mirrored down by several stages of current mirrors. Fig. 3.17 shows the current mirror with a 60 nA input current and a 23 pA output current.

The capacitors used in this process are poly-poly capacitors which are able to achieve relatively high capacitance in a reasonable area. Since the bottom plate has a
higher parasitic capacitance, it is connected to the output of the OTA for a more accurate closed-loop gain. The capacitors $C_1$ and $C_2$ are matched to each other in a common-centroid array.

Fig. 3.18 shows the simulation result of the frequency responses of the three different filters mentioned above. It is shown that the lower cutoff frequency around 500 Hz is achieved with the proposed pseudo resistor.

![Figure 3.16](image1.png)  
Figure 3.16, Current biased pseudo-resistors for $S_B/S_R = 1$ (left) and $S_B/S_R = 40$ (right).

![Figure 3.17](image2.png)  
Figure 3.17, Current mirror stages to mirror down the current.
Fig. 3.19 is the frequency responses for different bias currents $I_B$. The bias current $I_B$ is obtained by injecting a reference current which is 64 times larger than the bias circuit and the current is mirrored down by a 64:1 current mirror. The simulation result in Fig. 3.19 verifies the tunability of the pseudo-resistor.

The simulation result in Fig. 3.19 verifies the way of tuning the lower cutoff frequency by varying the current in the beta-multiplier reference circuit. Varying the current in the beta-multiplier circuit can be done by varying the value of the resistor. Assuming the beta-multiplier reference circuit is biased in weak inversion region, it can be found from (3.58) and (3.85) that:

$$f_L = \frac{1}{2\pi RC} \frac{S_K \ln K}{S_B m},$$

where $m$ is the mirroring ratio of the on-chip mirror, which is 64 in this case. Table 3.1 summarizes the cutoff frequencies at different $I_B$.

![Figure 3.18, Frequency responses of the three filters.](image-url)
Figure 3.19, frequency responses for the filter with $S_B/S_R = 40$ for different $I_B$.

Fig. 3.20 is the simulation result for different common mode voltage $V_{CM}$, as showing in Fig. 3.8 (a). It shows that the cutoff frequency is rarely affected by the output common mode voltage.

The input referred noise is very important for a neural filter due to the weak amplitude of the neural signal. Fig. 3.21 is the simulation result for the output noise and the input referred noise of filter A (10 V/V gain) and filter C (35 V/V gain). By integrating under the curve from 175 Hz to 10 KHz, it is found that the input referred noise is 47.52 $\mu$Vrms for filter A and 21.58$\mu$Vrms for filter C. Filter C has lower noise because it is biased at a higher current and the thermal noise of the OTA is lower. Since neural spikes usually have an amplitude range from 50 to 500 $\mu$V, filter A is probably too noisy and filter C are more practical at the expense of higher power consumption. Indeed, noise performance is improved at the expense of higher power consumption and one has
to realize the trade-offs between noise and power consumption.

Figure 3.20, Frequency responses with different common mode voltages for the filter with $S_B/S_R = 40$.

Figure 3.21, Input and output noise spectrum densities for filter A and filter C.
### Table 3.1, Cutoff Frequencies at Different $I_B$

<table>
<thead>
<tr>
<th>$I_B$ (pA)</th>
<th>15.6</th>
<th>31.3</th>
<th>46.9</th>
<th>117</th>
<th>234</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_L$ (Hz)</td>
<td>65m</td>
<td>7</td>
<td>15</td>
<td>55</td>
<td>117</td>
</tr>
</tbody>
</table>

3.10 Fully Differential Configuration in a Half-Micron SOS CMOS Process

The neural filters designed in a half-micron SOS CMOS process use folded-cascode fully differential OTAs. The advantages of using a fully differential configuration are that: 1) it gets rid of the mirror pole and gives a better phase margin; 2) it is able to cancel out even order distortions.

The filter designed in the half-micron SOS CMOS process consists of two stages, both of which share the same bias circuit. The first stage realizes a closed-loop gain of 25 V/V and the second stage realizes a closed-loop gain of 4 V/V. By cascading these two stages, a second order system with an overall gain of 100 V/V is built. The filters are both biased in weak inversion region and a power supply of 1.2 V is used. The filters, including the bias circuit, consume a total power of 1.88 µW. In order to obtain a reasonable input common mode range, low threshold MOSFETs are used as the input pairs in both the OTA and the common mode feedback circuit. The schematic of this OTA is shown in Fig. 3.22. Assuming $V_{DS} = 5U_T \approx 125$ mV, the input common mode voltage is:

\[ V_{in}^{(\text{max})} = V_{dd} - 250 \text{mV} - |V_{TP}| - \Delta V , \]  
\[ V_{in}^{(\text{min})} = V_{ss} + 250 \text{mV} - |V_{TP}| - \Delta V . \]  

In order to improve $V_{in}^{(\text{max})}$, $|V_{TP}|$ has to be reduced.

Table 3.2 summarizes the cutoff frequencies and the gains for both stages and their...
combination.

The disadvantage of cascading two first order systems to realize a second order one is that the latter has a very drooped response, in which case the old -3dB point becomes a -6dB one [41]. Figure 3.23 is the frequency responses of the second order filter and its two stages.

**Table 3.2, Fully Differential Filter Characterization**

<table>
<thead>
<tr>
<th></th>
<th>$f_L$ (-3 dB)</th>
<th>$F_U$ (-3 dB)</th>
<th>$f_L$ (-6 dB)</th>
<th>$F_U$ (-6 dB)</th>
<th>Midband Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage 1</strong></td>
<td>369 Hz</td>
<td>7.7 KHz</td>
<td>--------------</td>
<td>--------------</td>
<td>27.18 dB</td>
</tr>
<tr>
<td><strong>Stage 2</strong></td>
<td>362 Hz</td>
<td>4.77 KHz</td>
<td>--------------</td>
<td>--------------</td>
<td>11.48 dB</td>
</tr>
<tr>
<td><strong>Two Stages</strong></td>
<td>527 Hz</td>
<td>3.94 KHz</td>
<td>362 Hz</td>
<td>5.79 KHz</td>
<td>38.66 dB</td>
</tr>
</tbody>
</table>

![Figure 3.22, Fully-differential folded-cascode OTA.](image)

The fully differential filters use the same technique as the single ended ones to
achieve the lower cutoff frequency. The bias current $I_B$ is also generated by mirroring down the reference current in the beta multiplier reference circuit. The second order behavior of this filter is favorable for the following sigma-delta ADC because the ADC doesn’t have a very high oversampling rate and the neural filters act like an anti-aliasing filter.

![Frequency response graph](image)

Figure 3.23, Frequency responses of the fully differential filters.

3.11 Conclusions

This chapter discusses the issues associated with subthreshold circuit design and characterizes the current biased pseudo-resistor.

EKV model is used in the analysis of transistors biased in weak inversion in this chapter. Transistors biased in weak inversion region have a very high $g_{m}/I_D$ value and hence are power economic. Transistors in weak inversion region also have very high
output resistance and act more like an ideal current source. A single stage OTA biased in weak inversion region is able to provide sufficient gain due to the high output resistance of the transistors. However, OTAs working in weak inversion region have a very limited bandwidth (but enough for neural signal amplification applications) and relatively high input referred thermal noise (bad for neural signal amplification applications). It is better to increase the bias current for a reasonable input referred noise.

Filters with proposed pseudo-resistor are simulated in two CMOS processes in this chapter. The simulation results show that the desired cutoff frequency is achieved with such pseudo-resistors. The cutoff frequency is also tunable over a wide range and not sensitive to the output common mode voltage. The proposed structure is compact in size and simple in structure, and a distortion less than 1% can be easily achieved if the fully differential topology is utilized. This structure is suitable for neural signal front end amplification.
CHAPTER 4

LOW POWER SIGMA-DELTA ADC DESIGN

Today’s state-of-the-art digital circuits are able to perform sophisticated data computation and signal processing tasks. However, signals in our real world are analog, necessitating analog to digital converters (ADCs) to interface with the digital circuits. Neural signals, as a type of real world signal, also need to be digitized before they can be processed by any digital circuits. Digitized neural signals can also be modulated by a modulation technique called “keying”, such as phase-shift keying (PSK), frequency-shift keying (FSK) and amplitude-shift keying (ASK). This chapter discusses a low power analog to digital conversion solution for implantable neural signal applications.

4.1 Review of Nyquist Rate ADC’s and Oversampling ADC’s

ADCs can be categorized into two groups: Nyquist rate ADCs and oversampling ADCs. Nyquist rate ADCs sample the input signals at (in most cases higher than) Nyquist rate [42]. Examples of Nyquist rate ADCs are flash ADCs, subranging ADCs, Successive approximation ADCs, etc. Oversampling ADCs sample the input signal at a frequency much higher than Nyquist frequency. Normally Nyquist rate ADCs achieve higher conversion speed at the expense of resolution. On the contrary, oversampling ADCs achieve higher resolutions at the expense of conversion speeds. One desirable advantage of oversampling ADCs is that they don’t require a very high order anti-aliasing filter. In
such cases when high order anti-aliasing filters are not available, oversampling ADCs are the choice. One of the oversampling ADCs widely used today is sigma-delta ADCs. Besides the advantage of oversampling technique, the additional advantage of sigma-delta ADCs is that they are tolerant of the imperfections of analog parts. As will be discussed in the following sections, with proper topologies and circuit techniques, the effects from the imperfections of the analog parts in a sigma-delta ADC are minimized.

For implantable neural signal applications, due to the limitation of the chip area and power budget, it is hard to include an anti-aliasing filter higher than second order. In such a case a sigma-delta ADC is desired, especially when the supply voltage is low and high performance OTAs are hard to design.

4.2 Modulator Order Selection

Sigma-delta ADCs achieve signal to noise ratio (SNR) improvement by averaging and noise shaping [43]. Due to the area limitation of this project, up to second order modulators can be built. The selection of modulator order and oversampling rate can be done by comparing the signal to noise ratio of a general ADC and that of a sigma delta ADC:

\[
\text{SNR}_{\text{general}} = 6.02N_1 + 1.76, \quad (4.1)
\]

\[
\text{SNR}_{\text{1st\_order}} = 6.02N_2 + 1.76 - 5.17 + 30\log K_1, \quad (4.2)
\]

\[
\text{SNR}_{\text{2nd\_order}} = 6.02N_3 + 1.76 - 12.9 + 50\log K_2, \quad (4.3)
\]

where \(\text{SNR}_{\text{general}}\) is the ideal signal to noise ratio of a general ADC (without oversampling and noise shaping), \(N_I\) is the projected number of bits, \(\text{SNR}_{\text{1st\_order}}\) is the signal to noise ratio of a first order sigma delta ADC, \(N_2\) is the number of bits of the quantizer in the first
order sigma delta ADC, $\text{SNR}_{\text{2nd\_order}}$ is the signal to noise ratio of a second order sigma delta ADC, $N_3$ is the number of bits of the quantizer in the second order sigma delta ADC, and $K$ is the oversampling rate. Requiring a 10-bit accuracy, i.e. $N_1 = 10$, and assuming single bit quantizers are used for both types of modulators, i.e. $N_2 = N_3 = 1$, it is found that $K_1 = 95$ and $K_2 = 22$. As the bandwidth requirement and hence the power consumption of the OTA is determined by the base bandwidth and the oversampling rate, a second order sigma-delta ADC consumes less power than a first order one. An oversampling rate of 32 is finally used for the second order sigma-delta ADC because 32 is a multiple of 2. The architectural resolution of the ADC is hence 11 bits.

4.3 Modulator Design

With a supply voltage of 1.2 V, and limited power and area budget, it’s difficult to use gain enhancement techniques such as gain boosting and two stage OTA to improve the gain. What’s more, when the OTA is biased in subthreshold region, matching becomes challenging and higher offset voltages are expected. Under such circumstances, circuit techniques for reducing the effects of OTA imperfections are crucial to the ADC performance.

4.3.1 Reduced Harmonic Distortion CIFB

Consider the two different topologies in Fig. 4.1, which is a cascade of integrators in the feedback form. Fig. 4.1 (a) is an ordinary second order topology while Fig. 4.1 (b) is a low harmonic distortion topology which is discussed by Steensgaard in [44]. In Fig. 4.1 (a), the output of the first integrator contains a component that represents $\beta_2x[n]$ because this component is to be subtracted by $\beta_2y[n]$ to reduce the power in the signal band before.
the second integrator; however, the component $\beta_2x[n]$ is distorted by the first integrator and finally introduces harmonic distortion in the modulator. In order to compensate the component $\beta_2x[n]$, an additional signal path $\alpha_2x[n]$ is introduced. $\alpha_2x[n]$ doesn’t pass the first integrator and hence is not distorted. If $\alpha_2 = \beta_2$, $\alpha_2x[n]$ balances $\beta_2y[n]$ and gets rid off the distorted component $\beta_2x[n]$ at the output of the first integrator. As a result, the total harmonic distortion of the modulator is reduced. In order to avoid integrator saturation and instability, the following coefficient values are used: $\alpha_1 = \beta_1 = 0.5$, $\alpha_2 = \beta_2 = c_1 = 0.4$ and $c_2 = 1$ [45].

![Figure 4.1](image-url)

**Figure 4.1**, Ordinary second order topology (a) and reduced harmonic distortion topology (b).
4.3.2 Correlated Double Sampling

Correlated double sampling is a technique to reduce the effects of OTA imperfections, such as offset, 1/f noise and finite gain, on the integrators. This technique is used in this project due to the difficulty of designing a well performed OTA at low supply voltages. Consider the three different integrator topologies in Fig. 4.2.

![Integrator Comparison Diagram]

Figure 4.2, Integrator comparison: (a) ordinary integrator, (b) integrator with CDS technique, (c) integrator with CDS technique and slew-prevention capacitor.

Fig. 4.2 (a) is an ordinary integrator without correlated double sampling technique.

If the OTA has an infinite gain, the transfer function of the integrator is [46]:

\[
V_{\text{out}}(z) = \frac{C_1}{C_2} \frac{V_{\text{in}}(z) \cdot z^{-1}}{1 - z^{-1}}. \tag{4.4}
\]
However, a practical integrator has a finite gain and the transfer function in a general case is [43]:

\[
V_{\text{out}}(z) = \frac{C_1}{C_2} \frac{V_{\text{in}}(z) \cdot z^{-1}}{1 + \frac{C_1}{C_2} \frac{1}{A_{\text{OL}}(f)} - z^{-1}}.
\]  

(4.5)

The term \( \varepsilon_{\text{gain}} = \frac{C_1}{C_2} \frac{1}{A_{\text{OL}}(f)} \) is called gain error which causes integrator leakage when some of the charge in \( C_1 \) is not transferred to \( C_2 \) due to the finite gain of the OTA. However, if an auxiliary capacitor \( C_{\text{DS}} \) is introduced in the circuit, the finite gain error and the offset voltage of the OTA can be compensated [47]. During phase 1, the error voltage due to OTA offset, OTA finite gain and 1/f noise is stored in capacitor \( C_{\text{DS}} \); during phase 2 when the integration is performed, this offset voltage is in series with the inverting input of the OTA and hence is cancelled. The value of \( C_{\text{DS}} \) is not important in this case and can be made very small.

During the intervals between phase 1 and phase 2, both switches (switch 1 and switch 2) are opened and the integrator is in an open loop configuration. Slewing problem will occur because the feedback loop is broken at this time. The output of the OTA may swing towards the supply rails and generates spikes. In order to prevent this problem, Matsumoto et al proposed a slew-prevention capacitor \( C_M \) is introduced as in Fig. 4.2 (c) [48]. \( C_M \) provides a feedback loop for the OTA during the intervals between phase 1 and phase 2 and prevents slewing. The size of \( C_M \) is also not important and can be made very small.
4.3.3 Building Block Design

An important task for designing the OTA is to calculate its unity gain \( f_u \) frequency and transconductance \( g_m \). Assuming a settling time of \( 5\tau \) and a duty cycle of 1/3 of the clock period, the following relationships can be found [43]:

\[
\frac{1}{3f_s} = 5\tau, \tag{4.6}
\]

\[
\tau = \frac{1}{2\pi f_u \beta}, \tag{4.7}
\]

\[
\beta = \frac{C_2}{C_2 + C_1}. \tag{4.8}
\]

And the unity gain frequency is found to be:

\[
f_u = \frac{15f_s(C_1 + C_2)}{2\pi C_2}. \tag{4.9}
\]

The transconductance of the OTA can be calculated as:

\[
g_m = 2\pi f_u C_{\text{eff}}, \tag{4.6}
\]

\[
C_{\text{eff}} = C_L + C_1 + \frac{C_L C_1}{C_2}. \tag{4.7}
\]

In order to avoid slewing problem, the tail current has to satisfy the full power bandwidth \( f_m \):

\[
I_{\text{tail}} = \pi f_m V_{FS} C_{\text{eff}}. \tag{4.8}
\]
The clocked comparator used in this project is shown in Fig. 4.3 (a), an RS flip-flop consisting of two cross couple NOR gates is introduced to make sure that the comparator is a falling edge comparator.

Fig. 4.3 (b) is the implementation of the 1-bit DAC, which consists of two inverters in a parallel combination. When $V_{ip}$ is logic one and $V_{im}$ is logic zero, $V_{o+} = V_{ref+}$ and $V_{o-} = V_{ref-}$; however when $V_{ip}$ is logic zero and $V_{im}$ is logic one, $V_{o-} = V_{ref+}$ and $V_{o+} = V_{ref-}$.

![Diagram of the clocked comparator and DAC](image)

Figure 4.3, (a) The clocked comparator (the 1-bit ADC) and (b) The 1-bit DAC used in this project.

4.4 Simulation Results

The modulator in this project is shown in Fig. 4.4, where $C_{11} = 0.5 C_2$ and $C_{12} = C_{13} = 0.4 C_2$. For a signal bandwidth of 4.88 KHz and an oversampling rate of 32, the sampling
frequency is 312.5 KHz. The modulator was simulated in a 0.5 micron SOS process with Cadence Analog Environment, and the simulation results are given in Fig. 4.5. The plots on the left are the output waveforms of the 2-stage modulator, showing that the modulator is not saturated. The plots on the right are the input (top) and the output (bottom) of the modulator. The OTAs and the bias circuit consume a total static power of 15.3 µW.

4.5 Conclusions

A digitization system for the neural recording frontend, which is chosen to be a second order sigma-delta modulator, is designed and simulated in this chapter in a 0.5 micron SOS CMOS process. This modulator consumes a static power of 15.3 µW, which is among the lowest in published literatures, as showing in Table 4.1. The table shows both continuous time (CT) and switched capacitor (SC) modulators.

Figure 4.4, Implementation of the second order sigma-delta modulator.
Figure 4.5, Transient simulation results of the second order modulator.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Power and Supply Voltage</th>
<th>$f_s$ and $f_B$ (Hz)</th>
<th>OSR and Resolution</th>
<th>Type</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. H. Nielsen et al. [49]</td>
<td>108 µW 1.8 V 1.4 M 4 K</td>
<td>175; 10 bits</td>
<td>3rd order CT</td>
<td>0.35 µm CMOS</td>
<td></td>
</tr>
<tr>
<td>S. Kim et al. [50]</td>
<td>26.8 µW 0.9 V 2.048 M 8 K</td>
<td>128; 13 bits</td>
<td>2nd order SC</td>
<td>0.25 µm CMOS</td>
<td></td>
</tr>
<tr>
<td>E. Farshidi et al. [51]</td>
<td>80 µW 1 V 100 K 390</td>
<td>128; 8 bits</td>
<td>2nd order CT</td>
<td>0.35 µm CMOS</td>
<td></td>
</tr>
<tr>
<td>A. Gerosa et al. [52]</td>
<td>1.8 µW 1.8 V 8.192 K 256</td>
<td>16; 8 bits</td>
<td>3rd order SC</td>
<td>0.8 µm CMOS</td>
<td></td>
</tr>
<tr>
<td>S. S. Bazarjani et al. [53]</td>
<td>100 µW 1 V 1 M 4K</td>
<td>128; 8 bits</td>
<td>1st order SC</td>
<td>0.5 µm CMOS</td>
<td></td>
</tr>
<tr>
<td>This Work</td>
<td>15.3 µW 1.2 V 312.5 K 4.88K</td>
<td>32; 10 bits</td>
<td>2nd order SC</td>
<td>0.5 µm SOS</td>
<td></td>
</tr>
</tbody>
</table>
MEASUREMENT RESULTS

The filters discussed in Chapter III were fabricated in two CMOS processes: a half-micron process for the single ended configuration and a half-micron SOS process for the fully differential configuration. This section discusses the test results for these chips.

5.1 Single-Ended Filter Test Results

5.1.1 Dies for Testing

OTAs biased in weak inversion region usually have an output resistance of up to several hundred mega ohms; however the instruments used for testing usually have an input resistance of one mega ohms. The neural filters in this work are designed to drive a capacitance less than 10 pF but the parasitic capacitance of the connection cable and the input capacitance of the test instruments add up to more than 100 pF. As a result, on-chip buffers have to be introduced between the output of the filters and the input to the instruments to increase the driving capability of the chips. Fig. 5.1 is the structure of the die for testing filter A mentioned in section 3.9. Dies for testing filters B and C are similar to this one. The on-chip buffer is a voltage follower consisting of a unity gain configuration OTA biased in strong inversion region. Three transmission gates are used to switch the channels for testing the buffer, the open loop OTA and the neural
bandpass filter.

Figure 5.1, Die structure for the single ended filter in the half-micron CMOS process.

Fig. 5.2 (a) is the die photo for the chip and Fig. 5.2 (b) is the die photo for filter A, which shows an area of 680 μm × 305 μm. An independent bias circuit for the pseudo-resistor is also included for the purpose of testing. In a practical application, the pseudo-resistor can share the same bias circuit with the OTA, and the area can be reduced.

5.1.2 Frequency Response of the OTA

The frequency response of the OTA fabricated in the bulk process was tested with an oscilloscope and a function generator, the gain was calculated by dividing the output signal amplitude by the input signal amplitude. Sine wave was used in this test and the result is shown in Fig. 5.3. The frequency was varied from 1 Hz to 10 KHz.
Fig. 5.3 shows the measured open loop gain, the simulated open loop gain and the trend line that fits the measurement result. The measured gain is about 5 dB below the simulated one, which is due to the error of the simulator. The test result shows that at 500 Hz where the lower cutoff frequency (and the first pole of the filter) locates, the open
loop gain is about 40 dB, which is 23 dB higher than the desired circuit gain (17 dB); hence the finite OTA gain doesn’t affect the lower cutoff frequency much.

5.1.3 Frequency Response of the Filter

The tested frequency responses for filters A, B and C are shown in Fig. 5.5, Fig. 5.6 and Fig. 5.7, with the test setup shown in Fig. 5.4. The simulation results are also given in the same figure for comparison. The measurements were done with an Agilent 89441A vector signal analyzer, and the power supply was ±2.5 V. The midband gains for filters A, B and C are 19.5 dB, 19.12 dB and 29.4 dB respectively. It is found that the measured frequency responses for filter A and B fit the simulation results very well, however, the measured frequency response for filter C shifts to higher frequencies. This can be explained by looking into the bias current for the OTAs. Higher cutoff frequencies are due to higher bias currents. By measuring the resistance in the beta-multiplier circuit, it is found that the resistance is 20% lower than the laid out value, which causes an increase in the bias current, and leads to the increase in the cutoff frequencies. The simulation results also verify the fact that the current biased pseudo-resistor is able to set an accurate cutoff frequency given that the bias current is accurate. The accuracy of the bias current is most affected by the resistance in the beta-multiplier reference circuit.

5.1.4 Tunability Test

The change of frequency responses with bias current $I_B$ for filter B is shown in Fig. 5.9. The bias currents for the pseudo-resistor were generated by a HP 4155A semiconductor parameter analyzer and then they were scaled down by a 64:1 on chip current mirror to give $I_B$, as in Fig. 5.8. The test result verifies the tenability of the filter.
As the bias current decreases, the cutoff frequency shifts to lower end, this is consistent with (3.58) and the simulation results in Table 3.1.

![Frequency response of the OTA.](image)

**Figure 5.3, Frequency response of the OTA.**

5.1.5 Input Common Mode Voltage Sensitivity

Fig. 5.11 is the measured frequency responses for filter B at different common mode voltages $V_{CM}$. The common mode voltages for this measurement were generated by a HP 4155A semiconductor parameter analyzer, as shown in Fig. 5.10. $\pm800$ mV were used in the testing, and this should be much larger than the offset voltage due to transistor mismatches. The input common mode voltages actually set the output common mode voltages. The test results show that the lower cutoff frequency is not affected by the input common mode voltage, which in turn verify that the value of the pseudo-resistor is not affected by the output common mode voltage as a voltage biased pseudo-resistor does. This characteristic is very useful for fully implantable applications.
5.1.6 Noise Test

The output noise and input-referred noise for filter A and filter C are shown in Fig. 5.13 and 5.14. The input-referred noise spectrum density is obtained by dividing the output noise by the gain of the filter at the corresponding frequencies. The shapes of the measured noise spectrum densities are similar to the simulated ones. By integrating the input-referred noise spectrum densities from 175 Hz to 10 KHz, the input-referred noise for filter A and filter C is found to be 67.7 µV rms and 24.1 µV rms respectively. Filter A may not be very practical for extracellular action potential spike detection because these spikes usually have an amplitude range from 100 µV to 500 µV. Filter C is a more practical choice due to its lower input-referred noise, but it consumes higher power and larger chip area. The noise contribution by the on-chip buffer is ignorable due to the gain of the pre-amplifier filter stage.

Neural filter A has been used to amplified the neural signal from the auditory cortex of an awake rat, as described in [54]. The biomedical test result verifies the performance of the filter which suppresses the low frequency components and emphasizes the EAP spikes.

![Figure 5.4, Frequency response test setup.](image)
Figure 5.5, Frequency response of filter A.

Figure 5.6, Frequency response of filter B.
Figure 5.7, Frequency response of filter C.

Figure 5.8, Tunability test setup.
Figure 5.9, Measured frequency responses for filter B for different $I_B$.

Figure 5.10, Frequency response test setup for different input common mode voltages.
Figure 5.11, Measured frequency responses at different common mode voltage for filter B.

Figure 5.12, Noise test setup.
Figure 5.13, Measured noise spectrum densities for filter A.

Figure 5.14, Measured noise spectrum densities for filter C.
5.2 Fully-Differential Filter Test Results

5.2.1 Dies for Testing

The fully-differential filters were fabricated in a half micron SOS process. Due to the limited area assigned for the fully differential filters, it is not possible to lay out metal-metal capacitors and MOS capacitors are used instead. Fig. 5.15 shows the structure of the filter and the experimental setup for die testing. MOS capacitors used in the filters are also explicitly shown. Each MOS capacitor consists of three MOSFETs with their gates connected together. Two of them are zero threshold NMOSFETs (IN) placed back to back to form the capacitor. The regular PMOS (RP), whose gate is also tied to the gates of the zero threshold NMOSFETs, is small and acts as a switch for charge injection. The injected charge establishes a bias voltage at the gates of the NMOS in order to form gate oxide capacitance.

Fig. 5.16 is the die photos of the filters, which includes the bias circuits, two filters with different gains (25 V/V for the first stage and 4 V/V for the second stage) and the corresponding open-loop OTAs for either stage.

5.2.2 OTA Test Results

Fig. 5.17 is the experimental test setup for testing the OTA. Since the OTA biased in weak inversion region is not able to drive the input resistance and input capacitance of the test instruments, two low threshold PMOS source followers are included as the on-chip buffers. The bias currents for the buffers were generated by a HP4155A semiconductor parameter analyzer. One of the inputs of the OTA was connected to a 600 mV common mode voltage and the other input was swept from 300 mV to 900 mV. The
change of the output voltages for both $V_{out+}$ and $V_{out-}$ were recorded in Fig. 5.18 for the first stage OTA which was biased at 186 nA. The supply voltage for the OTA was 1.2 V. And the on-chip MOS load capacitance is about 1.5 pF. The test results show that the offset voltage is about 4 mV.

5.2.3 Filter Test Results

The frequency responses for the first stage filter are shown in Fig. 5.19 before charge injection. The simulation result is also shown in the same figure for comparison. Charge injection was performed but no injection current was observed indicating a failure of injection. The lower cutoff frequency shifts to a high frequency because the capacitance of the MOS capacitors is smaller without a bias voltage. The pads for charge injection were left floating in the test and the frequency responses were recorded with an Agilent 89441A vector signal analyzer.

Figure 5.15, Filter structure and test setup.
Figure 5.16, Die photo for the fully differential filters fabricated in the half micron SOS process.

Figure 5.17, OTA test setup.
Figure 5.18, OTA test results.

Figure 5.19, Frequency responses of stage 1 filter.
CHAPTER 6

CONCLUSION

A tunable low power fully integrated filter for neural signal recording applications and a low power sigma-delta ADC modulator were designed in this work. The filter with proposed current biased pseudo-resistors was fabricated in two difference CMOS processes, and the sigma-delta ADC modulator was simulated in a half micron SOS CMOS process.

The neural signal filter uses a novel current biased pseudo-resistor to set the cutoff frequency. With the proposed pseudo-resistor, the cutoff frequency is more predictable, tunable over a wide range and not sensitive to the output common mode voltage. Compared to gm-C method, filters with the proposed pseudo-resistor are more compact in size and suitable for low power implantable applications.

The sigma-delta ADC modulator works at an over sampling rate of 32. It utilizes techniques such as reduced distortion topology, correlated double sampling to reduce the drawback coming from OTA imperfection. The OTA’s are biased in moderate inversion region and the power consumption of the modulator is approximately 15.3 µW, which is among the lowest reported by published literatures. The sigma-delta ADC discussed in this work is suitable for low power implantable RFID application due to its low power characteristics.

The fully differential filters described in section 3.10 can be integrated with the
sigma-delta ADC modulator described in Chapter IV to make an amplification and digitization front end for implantable neural signal recording systems, and shown in Fig. 6.1. The cascaded filters can work as both the amplification front end of the system and the anti-alias filter for the ADC. However the integration work has not been done in this dissertation yet.

![System diagram of the amplification and digitization front end.](image)

Figure 6.1, System diagram of the amplification and digitization front end.

One of the problems associated with the filters biased in subthreshold region is the input referred noise. According to the simulation results in Chapter III and the measurement results in Chapter V, the filter with a bias current of 60 nA and a closed loop gain of 10 V/V has a simulated input referred noise of 47.52 µVrms and a measured input referred noise of 67.7 µVrms; however, the filter with a bias current of 187 nA and a closed loop gain of 35 V/V has a simulated input referred noise of 21.58 µVrms and a measured input referred noise of 24.1 µVrms. The former one is not very practical for EAP neural signal spike recording applications because these spikes have amplitudes ranging from 100 µV to 500 µV. In order to lower the input referred noise, the bias current and hence the power and area have to be increased.

Future work would include the integration of the filter and the sigma-delta ADC on a proper process which yields a reasonable area.


[52] A. Gerosa and A. Neviani, "A 1.8-μW sigma-delta modulator for 8-bit digitization of cardiac signals in implantable pacemakers operating down to 1.8 V," in *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 52, no. 2,


VITA

Sitong Yuan

Candidate for the Degree of

Doctor of Philosophy

Dissertation:  AN ULTRA LOW POWER AMPLIFICATION AND DIGITIZATION SYSTEM FOR NEURAL SIGNAL RECORDING APPLICATIONS

Major Field:  Electrical Engineering

Biographical:

Education:
Completed the requirements for the Doctor of Philosophy in Electrical Engineering at Oklahoma State University, Stillwater, Oklahoma in May, 2010.

Experience:  Sitong Yuan received his Master’s degree in Natural and Applied Sciences in May, 2005. From Aug, 2005 he has been working as a research assistant and a teaching assistant in the School of Electrical and Computer Engineering at Oklahoma State University. He earned his Bachelor’s degree in Applied Physics in Jul, 1999, from South China University of Technology.

Professional Memberships:  IEEE Student Member
Name: Sitong Yuan Date of Degree: May, 2010

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: AN ULTRA LOW POWER AMPLIFICATION AND DIGITIZATION SYSTEM FOR NEURAL SIGNAL RECORDING APPLICATIONS

Pages in Study: 100 Candidate for the Degree of Doctor of Philosophy

Major Field: Electrical Engineering

Scope and Method of Study: The scope is to develop a tunable low power fully integrated bandpass filter and a low power second order sigma-delta ADC modulator for implantable neural signal amplification and digitization applications, with subthreshold circuit design techniques in different CMOS processes. Since biopotentials usually contain low frequency components, the neural filters in this project have to be able to achieve large and predictable time constant for implantable applications. Voltage biased pseudo-resistors are vulnerable to process variations and circuit imperfections, and hence not suitable for implantable applications. A current biased pseudo-resistor is implemented in the neural filters in this work to set the cutoff frequency, and a Taylor series is used to study its linearity.

Findings and Conclusions: The filters with proposed current biased pseudo-resistors were fabricated in two different CMOS processes and tested. The test results verify that the filters with current biased pseudo-resistors are tunable, and not vulnerable to process variations and circuit imperfections. The filters with current biased pseudo-resistors meet the design requirements of fully integrated, implantable applications. The sigma-delta ADC modulator was designed and simulated in a half micron SOS CMOS process. The simulation results of the ADC confirm the possibility of an ultra low power ADC for neural signal recording applications.

ADVISER’S APPROVAL: Dr. Louis G. Johnson