

**An Ultra-Low-Power Neural Recording Amplifier
And Its Use in Adaptively-Biased Multi-Amplifier
Arrays**

by

Woradorn Wattanapanitch

Submitted to the Department of Electrical Engineering and Computer
Science

in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

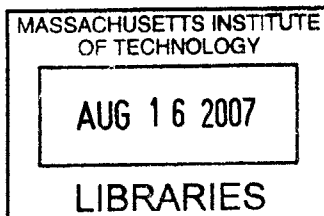
May 2007

© Massachusetts Institute of Technology 2007. All rights reserved.

Author
Department of Electrical Engineering and Computer Science
May 18, 2007

Certified by
Rahul Sarpeshkar
Associate Professor
Supervisor

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



BARKER

An Ultra-Low-Power Neural Recording Amplifier And Its Use in Adaptively-Biased Multi-Amplifier Arrays

by

Woradorn Wattanapanitch

Submitted to the Department of Electrical Engineering and Computer Science
on May 18, 2007, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering and Computer Science

Abstract

The design of a micropower energy-efficient neural recording amplifier is presented. The amplifier appears to be the lowest power and most energy-efficient neural recording amplifier reported to date. I describe low-noise design techniques that help the neural amplifier achieve an input-referred noise that is near the theoretical limit of any amplifier using a differential pair as an input stage. The bandwidth of the amplifier can be adjusted for recording either neural spikes or local field potentials (LFP). When configured for recording neural spikes, the amplifier yielded a midband gain of 40.8 dB and -3 dB bandwidth from 45 Hz to 5.32 kHz; the amplifier's input-referred noise was measured to be $3.06 \mu V_{rms}$ while consuming $7.56 \mu W$ of power from a 2.8 V supply corresponding to a Noise Efficiency Factor (NEF) of 2.67 with the theoretical limit being 2.02. When configured for recording LFPs, the amplifier achieved a midband gain of 40.9 dB and a -3 dB bandwidth from 392 mHz to 295 Hz; the input-referred noise was $1.66 \mu V_{rms}$ while consuming $2.08 \mu W$ from a 2.8 V supply corresponding to an NEF of 3.21. The amplifier was fabricated in AMI's $0.5 \mu m$ CMOS process and occupies 0.16 mm^2 of chip area.

The designs of two previous amplifiers that have been attempted are also presented. Even though they do not achieve optimal performances, the design insights obtained have led to a successful implementation of the energy-efficient neural amplifier discussed above. Finally, the adaptive biasing technique is discussed. The design and the detailed analysis of a feedback calibration loop for adjusting the input-referred noise of the amplifier based on the information extracted from the recording site's background noise is also presented. With such an adaptive biasing scheme, significant power savings in a multi-electrode array may be achieved since each amplifier operates with just enough power such that its input-referred noise is significantly but not overly below the neural noise.

Thesis Supervisor: Rahul Sarpeshkar
Title: Associate Professor

Acknowledgments

I would like to thank many people who have supported me during the time I was working on this thesis project.

First of all, I would like to thank my thesis supervisor, Professor Rahul Sarpeshkar who always provides valuable insights and suggestions for my research. The knowledge I have learned from him since I became his student is enormous. I would also like to thank Professor Michale Fee who spent his precious time helping me obtain the neural recording waveforms that were used in this thesis.

Next, I would like to thank Soumyajit Mandal whom I always turned to whenever I had any technical or writing problem. I also would like to thank all the previous and current members of the Analog VLSI & Biological Systems Group who supported me in many ways. They are Christopher Salthouse, Ji-Jon Sit, Keng-Hoong Wee, Lorenzo Turicchia, Scott Arfin, Michael Baker, Micah O'Halloran, Serhii Zhak, Benjamin Rapoport, Heemin Yang and Maziar Tavakoli-Dastjerdi. I also would like to thank Gretchen Jones, the past administrative assistant of AVBS who helped me make a smooth transition into this group. I also owe thanks to my academic advisor, Professor Steven Leeb.

Next, I would like to thank my parents, Somsak and Naruemol, and my sister, Warinsinee, who supported me through all these years. Their loves and cares for me are unsurpassed. Last of all, I would like to thank my girlfriend, Methichit Chayosumrit, who is always there for me for all these years.

Contents

1	Introduction	15
1.1	Power-Noise Tradeoff in the design of Neural Recording Amplifier . .	16
1.2	Existing Works on Neural Amplifier design	18
1.3	Adaptive Biasing Technique	19
2	A Feedforward Distributed-Gain Amplifier	21
2.1	A feedforward amplifier gain stage	22
2.2	Design of a feedforward distributed-gain amplifier using MOS-bipolar pseudoresistor elements for setting the amplifier's DC operating point	26
2.3	Robustness problems due to the uses of MOS-bipolar pseudoresistor elements	30
2.4	A feedforward distributed-gain amplifier using linear-region MOS transistors for setting the amplifier's DC operating point	32
2.5	Conclusion	36
3	A Feedback Neural Amplifier Using a Folded-Cascode OTA as the Gain Stage	37
3.1	Overall Amplifier Design	38
3.2	Small-Signal Analysis	40
3.3	Noise Analysis	43
3.4	Measurement Results	47
3.5	Problems in The Implementation of This Design	52
3.6	Conclusion	53

4	An Energy-Efficient Micropower Neural Recording Amplifier	55
4.1	Overall System Design	55
4.2	Low-Power Low-Noise OTA Design for Gain Stage	58
4.2.1	Device Sizing for Maximizing G_m	61
4.2.2	OTA Noise Analysis	64
4.2.3	Current Mirror Mismatch Analysis	67
4.2.4	Noise Efficiency Factor and Its Theoretical Limit for Any OTA	69
4.3	Measurement Results	70
4.4	Measurements of Local Field Potentials	75
4.5	Conclusion	78
5	Variable Input-Referred Noise Amplifier and Adaptive-Biasing Technique	79
5.1	A Variable Input-Referred Noise Amplifier	80
5.2	Feedback Calibration Scheme for Adjusting the Amplifier's Input-Referred Noise	80
5.3	Input-Referred Noise Calibration Loop's Building Blocks	83
5.3.1	Neural Amplifier's Noise Detection Circuitry	84
5.3.2	Current Subtractor	89
5.3.3	Lead Compensation Network	90
5.3.4	Bias Decision Circuitry	90
5.4	Feedback Analysis of the Input-Referred Noise Calibration Loop	94

List of Figures

1-1	A common-source amplifier with its noise sources	16
1-2	A noise distribution function of an array of 64 electrodes collected from a sampling statistic.	20
2-1	A schematic of a two-stage amplifier with input-referred noise sources	22
2-2	A schematic of a feedforward gain stage	23
2-3	A comparison of thermal noise in a MOS transistor and a resistor . .	25
2-4	A MOS-bipolar pseudoresistor high-resistance element	27
2-5	A feedforward distributed gain amplifier using psuedoresistor elements for rejecting DC offsets at electrode-tissues interfaces	28
2-6	The implementation of each amplifier stage: A_1 (Top); A_2 (Bottom) .	29
2-7	Schematic of the first stage where a step change in DC offset of ΔV_{ref} is applied.	31
2-8	Schematic of the feedforward distributed-gain amplifier using MOS transistors in linear region to set the amplifier's DC operating point. .	32
2-9	The transfer function of the feedforward distributed-gain amplifier using MOS transistors in linear region to set the amplifier's DC operating point.	33
2-10	An input-referred noise spectral density of the amplifier.	34
2-11	A schematic for calculating the input-referred noise from noise of M_{b1} and M_{b2}	35
3-1	A high-level schematic of the feedback neural amplifier.	39

3-2	A circuit schematic for analyzing the operation of the folded-cascode gain stage.	41
3-3	A preliminary block-diagram describing the operation of the feedback amplifier.	41
3-4	A unity-gain feedback block diagram describing the operation of the feedback amplifier.	42
3-5	A folded-cascode OTA schematic used in this design.	44
3-6	A small-signal schematic for describing the operation of folded-cascode OTA.	45
3-7	A small-signal block diagram describing the operation of folded-cascode OTA.	45
3-8	A micrograph of the feedback amplifier using a current-scaled folded cascode OTA as a gain stage.	48
3-9	A measured transfer function of the feedback amplifier.	49
3-10	A measured input-referred noise spectral density of the feedback amplifier.	50
3-11	Neural recording from the RA region of a zebra-finch bird brain.	51
4-1	Overall System Schematic of the Neural Amplifier.	56
4-2	Block diagram of our neural amplifier including the input noise source of the OTA.	57
4-3	Schematic of the low-noise OTA used in this design.	58
4-4	Circuit schematic for analyzing current scaling in the source-degenerated current mirrors of Fig. 4-3.	59
4-5	Schematic of a standard folded-cascode OTA.	61
4-6	Circuit schematics for obtaining admittance formula.	62
4-7	Circuit schematics for analyzing V_T and R mismatches in source-degenerated current mirrors.	67
4-8	A die micrograph of our neural amplifier.	70

4-9	Measured transfer function of the neural amplifier configured for recording neural spikes.	71
4-10	Measured and simulated (smooth curve) input-referred noise spectra of the neural amplifier configured for recording neural spikes.	72
4-11	NEF of published neural amplifiers as a function of the total supply current.	73
4-12	CMRR and PSRR measurements of the neural amplifier configured for recording action potentials.	74
4-13	Neural recording from a zebra finch's brain: (a) A zebra finch (b) Long time trace (c) Short time trace.	76
4-14	Transfer function of the amplifier configured for recording LFP.	77
4-15	Measured and simulated (smooth curve) input-referred noise spectra for the amplifier configured for recording LFP.	77
5-1	Noise power spectrum for three different supply currents.	81
5-2	Overall architecture of adaptive biasing loop.	81
5-3	Schematic of noise calibration feedback loop implemented in this design.	83
5-4	High-level schematic of the neural amplifier's noise detection circuitry.	85
5-5	A schematic of the envelope detector used in this design.	86
5-6	A schematic of the current-mode peak detector.	88
5-7	A schematic of current subtracter and lead compensation network.	90
5-8	A schematic of the Biasing Decision Circuitry.	91
5-9	A schematic of the G_{mB} OTA.	91
5-10	A schematic of current inversion circuit.	92
5-11	A schematic of the current limiting circuit.	94
5-12	A small-signal feedback block diagram of the input-referred noise calibration loop.	97
5-13	Root locus plots (a) Neural's amplifier's noise envelope is increasing (b) Neural's amplifier's noise envelope is decreasing.	98

List of Tables

3.1	Measured Performance Characteristics for the feedback amplifier . . .	52
4.1	Operating Points for Transistors in the OTA with $I_{tot} = 2.7 \mu\text{A}$. . .	64
4.2	Measured Performance Characteristics	73
4.3	Measured Performance Characteristics of LFP Amplifier	78
5.1	Parameters of the peak detector	89

Chapter 1

Introduction

Large-scale chronic multi-electrode neural-recording systems are being built to enable us to understand how the brain works [4,7]. With the help of such systems, a number of experiments have shown that it is possible to predict intended limb movements by simultaneously recording from many neurons, and interpreting their cortical activities [1,15]. For example, brain-machine interfaces are being built to help a paralyzed patient move a computer cursor by thoughts alone. Portable, chronic use of such interfaces may eventually play an important role in treatment of paralyzed patients, and enable large-scale monitoring of the brain in experimental neuroscience.

One of the most important parts in the development of brain-machine interfaces is the neural signal amplifier. Neural signals from extracellular recording are very weak (typically between $10\ \mu\text{V}$ and $500\ \mu\text{V}$). As a result, amplification is needed before they can be processed further. Next generation multi-electrode recording systems will be entirely implanted within the skull and incorporate a large number of neural amplifiers (on the order of 100-1000, one for every electrode). For such applications, ultra-low-power operation is very important. To get clean neural signal recording, it is important that the input-referred noise of the amplifier is kept low. Practically, the input-referred noise of the amplifier should be kept below the background noise of the recording site ($5\ \mu\text{V}$ - $10\ \mu\text{V}$) [4].

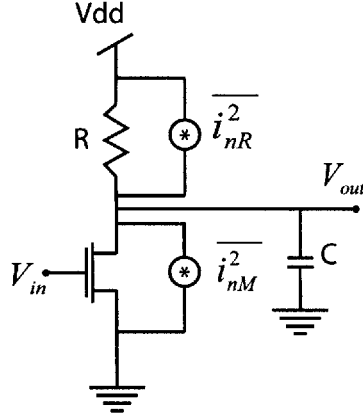


Figure 1-1: A common-source amplifier with its noise sources

1.1 Power-Noise Tradeoff in the design of Neural Recording Amplifier

To achieve the low-noise performance, designers must address the power-noise tradeoff in the design of an amplifier. As an example, let's consider a common-source amplifier shown in Fig. 1-1. Let's assume that output impedance of each transistor is much higher than its resistive load. We shall consider the input-referred thermal noise of this amplifier and see how it trades off with the power consumption. The thermal current noise source in an MOS transistor can be modeled as

$$\overline{i_{nM}^2} = 4kT\gamma g_m \quad (1.1)$$

where k is the Boltzmann constant and is equal to 1.38×10^{-23} , T is the absolute temperature, and γ is equal to $2/3$ in strong inversion and $1/(2\kappa)$ in weak inversion. The parameter κ is the subthreshold gate coupling coefficient of a MOSFET which has a typical value of 0.6-0.7 and g_m is the MOSFET's transconductance. The thermal current noise in a resistor can be expressed as

$$\overline{i_{nR}^2} = \frac{4kT}{R}. \quad (1.2)$$

The input-referred thermal noise of this amplifier can be calculated as the output noise divided by the gain of the amplifier to be

$$\overline{v_{ni}^2} = \frac{1}{g_m^2} \left(4kT\gamma g_m + \frac{4kT}{R} \right) = \frac{4kT}{g_m} \left(\gamma + \frac{1}{g_m R} \right) \approx \frac{4kT\gamma}{g_m} \quad (1.3)$$

assuming that $g_m R$, which is the gain of the amplifier, is much greater than $1/\gamma$, thus $1/(g_m R)$ is negligible compared to γ if the amplifier has a high-enough gain. The total input-referred thermal noise of the amplifier can be calculated by integrating the noise over the entire frequency range to be

$$V_{ni,thermal} = \sqrt{\frac{\pi}{2} \frac{4kT\gamma}{g_m} \frac{1}{2\pi RC}} = \sqrt{\frac{kT\gamma}{g_m RC}}. \quad (1.4)$$

In weak inversion where an MOS transistor achieves a maximum g_m/I_D where I_D is the drain current of the transistor, we have $\gamma = 1/(2\kappa)$ and $g_m = \kappa I_{tot}/U_T$, where I_{tot} is the total current of the common-source amplifier in Fig. 1-1. Therefore, we can express the total input-referred thermal noise of the common-source amplifier with the transistor operating in weak inversion as

$$V_{ni,thermal} = \frac{1}{\kappa} \sqrt{\frac{1}{I_{total}} \frac{U_T \cdot kT}{2RC}}. \quad (1.5)$$

Since the total power consumption is $P = I_{total} V_{dd}$, we can express the total power consumption of the amplifier as a function of input-referred thermal noise as

$$P = \frac{1}{V_{ni,thermal}^2} \frac{U_T \cdot kT \cdot V_{dd}}{2RC\kappa^2}. \quad (1.6)$$

Equation (1.6) clearly shows the tradeoff between the power consumption and the total input-referred thermal noise of a subthreshold amplifier for a given supply voltage and bandwidth (denoted by RC product in this case). To reduce the input-referred thermal noise by a factor of 2, the total power consumption must be increased by a factor of 4. This relationship shows a steep power cost of achieving low-noise performance in a thermal-noise limited amplifier, even without taking a flicker noise into

account. The power-noise tradeoff in the amplifier is aggravated if the transistor is operating in strong inversion. In strong inversion, the transconductance g_m is proportional to $\sqrt{I_{tot}}$. As a result, the total power consumption scales as $1/V_{ni,thermal}^4$ instead of $1/V_{ni,thermal}^2$ as in the subthreshold case.

1.2 Existing Works on Neural Amplifier design

Many designs of neural amplifiers have been reported in the literature [5, 6, 8–10]. Most amplifiers consume power near $100 \mu\text{W}$ to achieve less than $10 \mu\text{V}_{rms}$ input-referred noise for bandwidths of 5-10 kHz. The designs in [8, 10] consume power near $100 \mu\text{W}$ to achieve about 8-9 μV_{rms} input-referred noise with approximately 10 kHz of bandwidth. The design in [6] achieves an input-referred noise of $2.2 \mu\text{V}_{rms}$ with 7.2 kHz of bandwidth while consuming $80 \mu\text{W}$ of power. If such amplifiers are to be used in a multi-electrode array, with a power near $100 \mu\text{W}$ per amplifier for most designs, the power required for the neural amplifiers can become the limiting factor for the whole multi-electrode system.

The design in [6] presents many useful techniques for designing a neural amplifier. The use of MOS-bipolar pseudoresistor element as a high-resistance element and on-chip AC-coupling capacitors enable the amplifier to reject large DC offsets at electrode-tissue interfaces while being able to pass the neural signals of interest. Since high-resistance elements can be implemented in a small area on chip, large off-chip components are not needed. The amplifier in [6] uses a standard wide-output swing operational transconductance amplifier (OTA) with capacitive feedback to realize a gain of approximately 40 dB. The work presents design techniques that minimize the input-referred noise of the amplifier by operating some devices of the OTA in strong inversion to minimize their noise contributions. Even though the design achieves the power-noise tradeoff near theoretical limit of that particular OTA topology, the OTA used is actually not power-efficient since a large portion of total current are wasted in the current mirrors. The power efficiency of the amplifier can be greatly improved if a new OTA topology that makes use of the supply current more efficiently is used.

Therefore, as a part of this thesis project, I have developed a new neural amplifier's topology that appears to be the most power-efficient neural amplifier reported to date. With such design, the power consumption per amplifier is low enough such that the total power consumption of a multi-electrode array may no longer be the bottleneck for the design of brain-machine interfaces. The detail operation of the amplifier will be discussed in Chapter 4.

1.3 Adaptive Biasing Technique

One problem of the existing neural amplifiers is that they are designed to handle the worst-case signal-to-noise ratio expected at any recording situation. As a result, amplifiers in a multi-amplifier array are usually designed to have the input-referred noise below the recording site background noise at any location. However, in practice, the background noise strength can vary significantly from one recording site to the other. An example of noise distribution function for the recording site's background noise is shown in Fig. 1-2. The plot was obtained from the lab of Professor Richard Andersen at Caltech. Notice a long tail skewed toward large background noise. The lowest noise is approximately at $4 \mu V_{rms}$ while the noise distribution function peaks at $15 \mu V_{rms}$. The probability that the background noise is lower than $14 \mu V_{rms}$ is approximately only 21%. Thus, if every amplifier in the array is biased such that its input-referred noise is $4 \mu V_{rms}$ to handle the lowest background noise situation, a large fraction of the total power in the array is wasted in those amplifiers whose recording sites have much higher background noise than their input-referred noise. In other words, those amplifiers' power consumptions are higher than they are necessary. A significant power saving can be achieved if each amplifier in the array has some freedom in adjusting its input-referred noise. For instance, when the background noise at one recording site is very high, say, $30 \mu V_{rms}$, instead of biasing the amplifier at that recording site to have an input-referred noise of $4 \mu V_{rms}$, the total current of the amplifier can be lowered significantly such that their input-referred noise is about $30 \mu V_{rms}$. In this thesis project, I also investigated the idea of how to adaptively

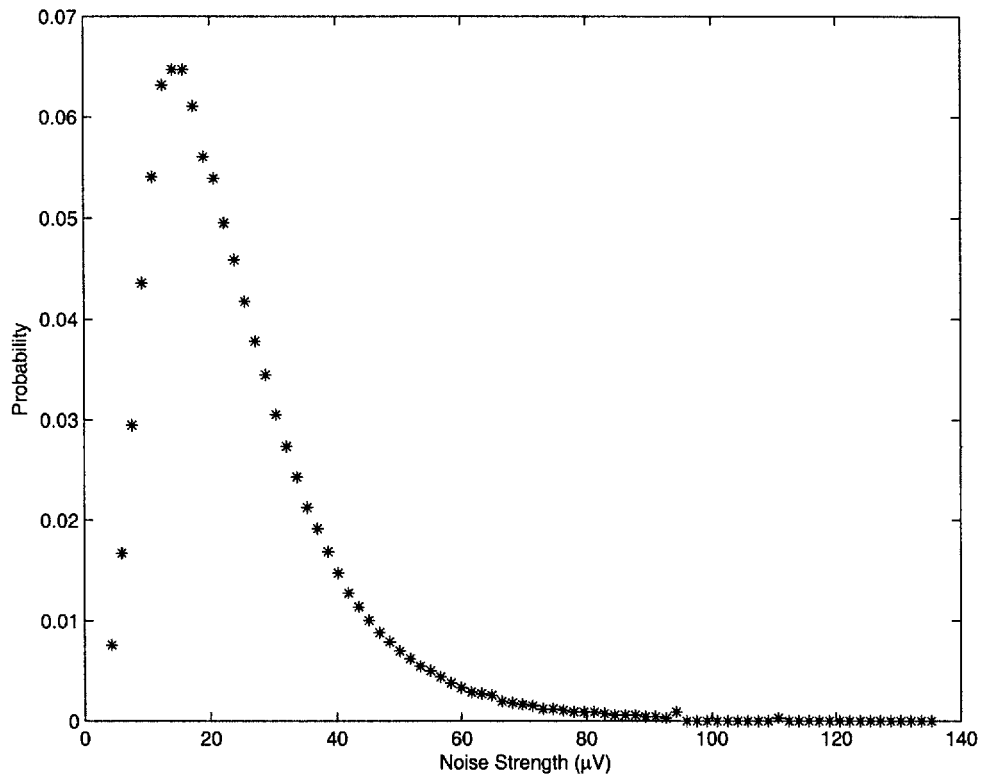


Figure 1-2: A noise distribution function of an array of 64 electrodes collected from a sampling statistic.

bias each neural amplifier such that its input-referred noise can be adjusted to suite the background noise at the recording site. With such adaptive-biasing scheme, the total power consumption of a multi-amplifier array is determined by the average case, rather than the worst case as in other fixed-power neural amplifier arrays.

Chapter 2

A Feedforward Distributed-Gain Amplifier

A feedforward amplifier topology instead of a feedback topology appeared to be a strong candidate for realizing low-power low-noise neural amplifier at first. During the first part of this thesis project, I investigated the idea of using a feedforward distributed-gain amplifier topology to realize a low-power low-noise neural amplifier. Unfortunately, the topology posed some challenges that remained unsolved. However, the design insights obtained from the feedforward distributed-gain amplifier design led to a successful design of an energy-efficient micropower neural amplifier which will be discussed later in Chapter 4. In this chapter, I will present the basic ideas behind the feedforward distributed-gain amplifier and technical problems that I encountered during the design and verification phases that prevented this feedforward distributed-gain amplifier to be used in real neural recording situations.

To achieve the desired overall gain, the gain of the amplifier can be distributed among many stages. If the gain of the first stage is high, the total input-referred noise of the overall amplifier is dominated by the input-referred noise of the first stage. This idea can be illustrated with a two-stage amplifier shown in Fig. 2-1. The gain and the input-referred noise per unit bandwidth of the i^{th} stage are modeled as A_i and $\overline{V_{ni}^2}$ respectively. The overall gain of the amplifier is $A = A_1 \cdot A_2$. We can then calculate

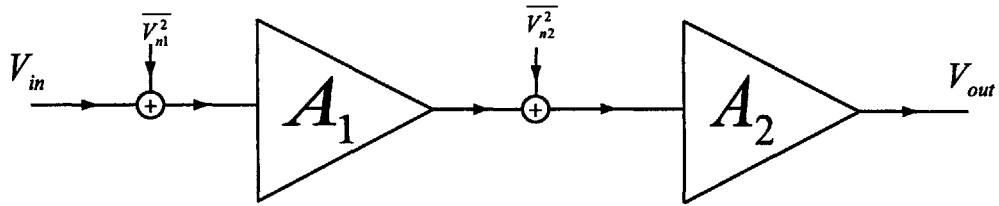


Figure 2-1: A schematic of a two-stage amplifier with input-referred noise sources

the input-referred noise per unit bandwidth of the overall amplifier to be

$$\overline{V_{n,in}^2} = \overline{V_{n1}^2} + \frac{\overline{V_{n2}^2}}{A_1^2}. \quad (2.1)$$

From 2.1, the input-referred noise power of the second-stage amplifier is attenuated by a factor of A_1^2 . Therefore, if the first-stage amplifier's gain A_1 is high, the input-referred noise requirement of the second-stage amplifier can be significantly relaxed. To achieve low-noise performance and desired overall gain, the first-stage amplifier should be designed to have low input-referred noise with enough gain while the subsequent stages just need to provide sufficient gains to meet the gain requirement for the overall amplifier while their input-referred noise requirements need not be as low as that of the first-stage amplifier. As discussed in Chapter 1, the input-referred thermal noise of the amplifier is proportional to $1/V_n^2$ where V_n is the total input-referred noise of the amplifier. Therefore, subsequent amplifier stages' power consumptions can be significantly lowered without severely degrading their input-referred noise per unit bandwidth. Thus, for a distributed-gain amplifier, most of the overall power consumption should be consumed in the first-stage amplifier since its input-referred noise is the most critical and its gain should be sufficiently high such that the noise contributions from subsequent amplifier stages become insignificant.

2.1 A feedforward amplifier gain stage

Since weak-inversion MOS transistors achieve the highest transconductance at a given bias current, they are thus the most suitable choice for use as input differential-

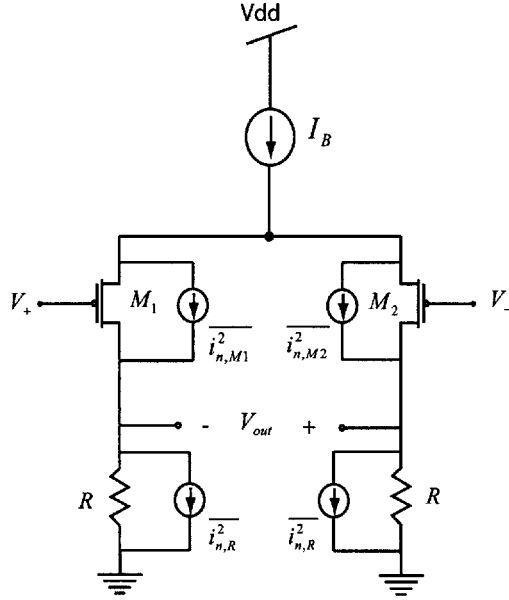


Figure 2-2: A schematic of a feedforward gain stage

pair transistors. I have shown in Chapter 1 that a resistively-loaded common-source amplifier can achieve low input-referred noise since the noise contribution from the resistive load can be made negligible compared to the noise from the transistor if the amplifier has a sufficiently high gain. Therefore, a resistively-loaded differential amplifier was chosen as the topology of a gain stage. The amplifier along with its current noise sources are shown in Fig. 2-2. Define the differential input voltage as $V_{in} = V_+ - V_-$. The low-frequency gain of the amplifier can be expressed as

$$A_{diff} = \frac{V_{out}(s)}{V_{in}(s)} = g_m (R \parallel 1/r_o) \approx g_m \cdot R \quad (2.2)$$

assuming that r_o which is the output resistance of M_1 and M_2 is much greater than R . Let's calculate the input-referred noise of this amplifier. To get the lowest input-referred noise per unit bandwidth, M_1 and M_2 operate in subthreshold. Thus their thermal noise current per unit bandwidth is given by

$$\overline{i_{n,M1}^2} = \overline{i_{n,M2}^2} = 2qI_{M1} = qI_B \quad (2.3)$$

where I_{M_1} is the channel current of M_1 and M_2 . The thermal noise current per unit bandwidth of each resistor can be expressed as

$$\overline{i_{n,R}^2} = \frac{4kT}{R}. \quad (2.4)$$

The thermal output voltage noise per unit bandwidth of the amplifier is calculated to be

$$\overline{V_{n,out}^2} = \overline{i_{n,M_1}^2} R^2 + \overline{i_{n,M_2}^2} \cdot R^2 + 8kTR \quad (2.5)$$

$$= 2qI_B \cdot R^2 + 8kTR \quad (2.6)$$

The input-referred thermal noise per unit bandwidth of the amplifier can be calculated from the output thermal noise per unit bandwidth divided by the gain of the amplifier. Using the fact that g_m of one of the input differential-pair transistors operating in subthreshold is $\kappa I_{M_1}/U_T = \kappa I_B/(2U_T)$, we can write the total input-referred thermal noise per unit bandwidth of this amplifier as

$$\overline{V_{n,in}^2} = \frac{2qI_B R^2 + 8kTR}{\left(\frac{\kappa I_B}{2U_T}\right)^2 \cdot R^2} \quad (2.7)$$

$$= \left(\frac{2U_T}{\kappa}\right)^2 \cdot \frac{2q}{I_B} \left(1 + \frac{2kT}{q} \cdot \frac{1}{I_B R/2}\right) \quad (2.8)$$

$$= S \cdot (1 + \alpha). \quad (2.9)$$

The quantity $S = \left(\frac{2U_T}{\kappa}\right)^2 \cdot \frac{2q}{I_B}$ is the input-referred noise per unit bandwidth of the amplifier assuming that the input differential-pair transistors are the only noise sources in the amplifier. The quantity $\alpha = \frac{2kT}{q} \cdot \frac{1}{I_B R/2}$ is the noise excess factor that captures the remaining thermal noise sources in the amplifier. Equation (2.7) suggests that for a given bias current, we can achieve low-noise performance by making $I_B R/2$, which is the voltage drop across each resistor, several times of $2\frac{kT}{q} = 2U_T$. For instance, if the voltage drop across each resistor is 300 mV, the parameter α equals to 0.17, thus the input-referred thermal noise per unit bandwidth of this amplifier is $\overline{V_{n,in}^2} = S \times 1.17$ which is close to the ideal value S .

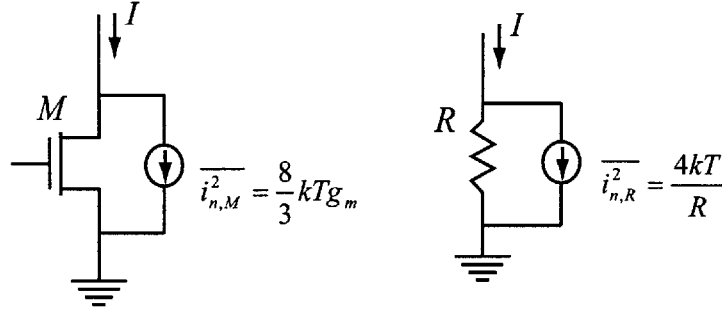


Figure 2-3: A comparison of thermal noise in a MOS transistor and a resistor

One might consider using linear-region MOS transistors as the loads instead of resistors to save the chip area. However, this might not be the case since in order to achieve the same low-noise performance, the transistor needs to be strongly inverted thus may need to have large area. To illustrate this point, let's compare the thermal noise in a resistor and in an MOS transistor operating in strong inversion and in linear region while they are running at the same current level as shown in Fig. 2-3. In strong inversion, the thermal noise of an MOS transistor in linear region can be expressed as

$$\overline{i_{n,M}^2} = \frac{8}{3}kTg_m \left(\frac{1 + \eta + \eta^2}{1 + \eta} \right), \quad (2.10)$$

$$\eta = 1 - \frac{V_{DS}}{V_{DSAT}} \quad (2.11)$$

where V_{DS} and V_{DSAT} are the drain-source voltage and saturation voltage of the transistor and g_m is the transconductance of the same transistor if it is operating in saturation. Let assume that the transistor is operating at the verge of saturation, thus $\eta = 0$ (for lowest noise at the same inversion level). Let the voltage drop across the resistor be $V = IR = 300$ mV. The thermal noise current in the resistor can then be expressed as

$$\overline{i_{n,R}^2} = \frac{4kT}{R} = \frac{4kT}{(V/I)} = \frac{4}{0.3}kT \cdot I \quad (2.12)$$

In order for the current noise of the transistor to be equal to that of the resistor, we need

$$\overline{i_{n,M}^2} = \frac{8}{3}kTg_m = \frac{4}{0.3}kT \cdot I \quad (2.13)$$

This would result in $g_m/I = 5 V^{-1}$. In subthreshold, an MOS transistor exhibits $g_m/I = \kappa/U_T \approx 29 V^{-1}$ and g_m/I progressively decreases as the transistor enters farther into strong inversion. This means that in order for the transistor to exhibit the same amount of thermal current noise as that of the resistor with a voltage drop of 300 mV across it, the transistor needs to operate far into strong inversion. This means that the transistor needs to be very long in order to achieve such level of inversion. Therefore, using MOS transistors as the loads may not offer any area advantage over the use of resistors. Furthermore, an MOS transistor also exhibits large 1/f noise unless its area is made large. This 1/f noise turns out to be a very important consideration for the design of a neural amplifier in which low-frequency operation is required.

2.2 Design of a feedforward distributed-gain amplifier using MOS-bipolar pseudo-resistor elements for setting the amplifier's DC operating point

The resistively-loaded differential amplifier in Fig. 2-2 can be used to realize a low-noise distributed-gain amplifier. Since the input-referred noise of the first stage amplifier dominates, most of the power consumption should be in the first stage. The power in the second stage amplifier can be much lower since it can tolerate much higher input-referred noise.

In real recording situations, neural amplifiers must be able to reject the DC offset voltage due to electro-chemical effects developed at the electrode-tissue interfaces. This DC offset voltage can vary by a few hundreds of millivolts thus it may saturate a high gain amplifier. The most widely used method to reject this DC offset voltage is to use an AC-coupling capacitor together with a MOS-bipolar pseudo-resistor element as a high-resistance element to create a high-pass cutoff at very low frequency such that

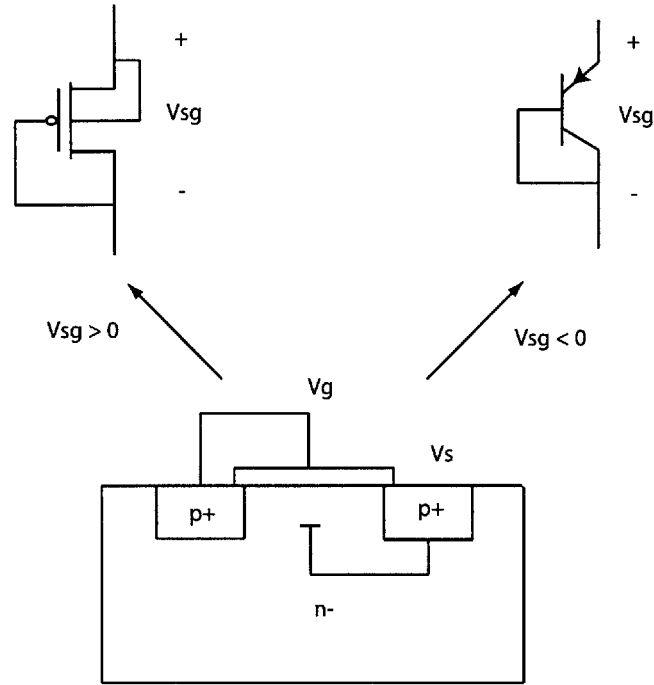


Figure 2-4: A MOS-bipolar pseudoresistor high-resistance element

the amplifier only rejects DC offset voltage while still passes the signal of interest [2]. The schematic of a MOS-bipolar pseudoresistor element is shown in Fig. 2-4. When $V_{sg} > 0$, the MOS-bipolar pseudoresistor element acts just like a diode-connected PMOS transistor. When $V_{sg} < 0$, the element acts like a diode-connected PNP transistor. However, when $V_{sg} \approx 0$ the incremental resistance of this element is very high ($\gg 10^{12} \Omega$). Since this high-resistance element can be realized in a small area thus eliminating the needs for large off-chip components, it is widely used in the designs of neural amplifiers to realize a low-frequency highpass cutoff for rejecting DC offset voltage.

The first design of a feedforward distributed-gain amplifier is shown in Fig. 2-5. The MOS-bipolar pseudoresistor elements M_{b1} - M_{b4} are used to set the DC input voltage of each amplifier. The amplifiers A_1 and A_2 are implemented as resistively-loaded differential amplifiers as shown in Fig. 2-6 Let V_{in1} and V_{out1} be the differential input voltage and the differential output voltage of A_1 respectively. Similarly V_{in2} and V_{out2} are the differential input voltage and the differential output voltage of A_2

respectively. The transfer function of each amplifier can be approximated as

$$\frac{V_{out1}}{V_{in1}}(s) = A_1(s) = \frac{g_{m1}R_1}{1 + sR_1C_1} \quad (2.14)$$

and

$$\frac{V_{out2}}{V_{in2}}(s) = A_2(s) = \frac{g_{m3}R_2}{1 + sR_2C_2}. \quad (2.15)$$

Let r_a denote the incremental resistance of a MOS-bipolar pseudoresistor element when its gate-source voltage is approximately zero. We obtain the overall transfer function of the feedforward distributed-gain amplifier in Fig. 2-5 to be

$$\frac{V_{out}}{V_{in}}(s) = \left(\frac{sr_a C_{in1}}{1 + sr_a C_{in1}} \cdot \frac{g_{m1}R_1}{1 + sR_1C_1} \right) \cdot \left(\frac{sr_a C_{in2}}{1 + sr_a C_{in2}} \cdot \frac{g_{m3}R_2}{1 + sR_2C_2} \right). \quad (2.16)$$

Due to AC coupling at the input of each amplifier stage, the DC gain of the overall amplifier is zero thus the amplifier should be able to reject a DC offset voltage at the electrode-tissue interfaces in real recording situations.

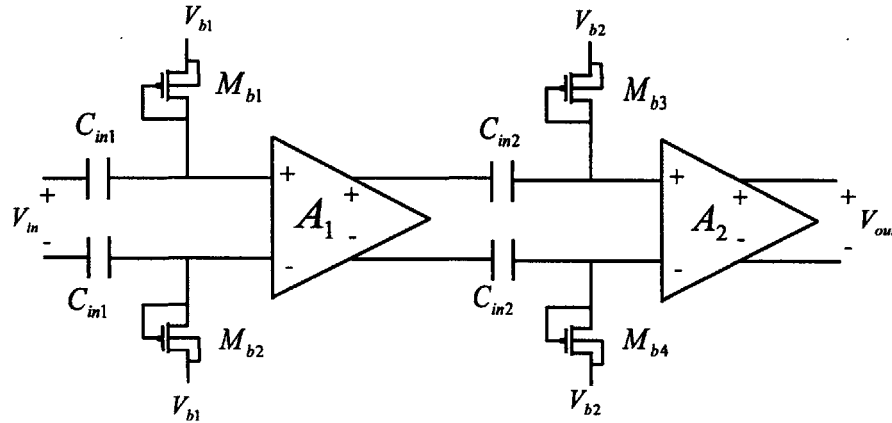


Figure 2-5: A feedforward distributed gain amplifier using pseudoresistor elements for rejecting DC offsets at electrode-tissues interfaces

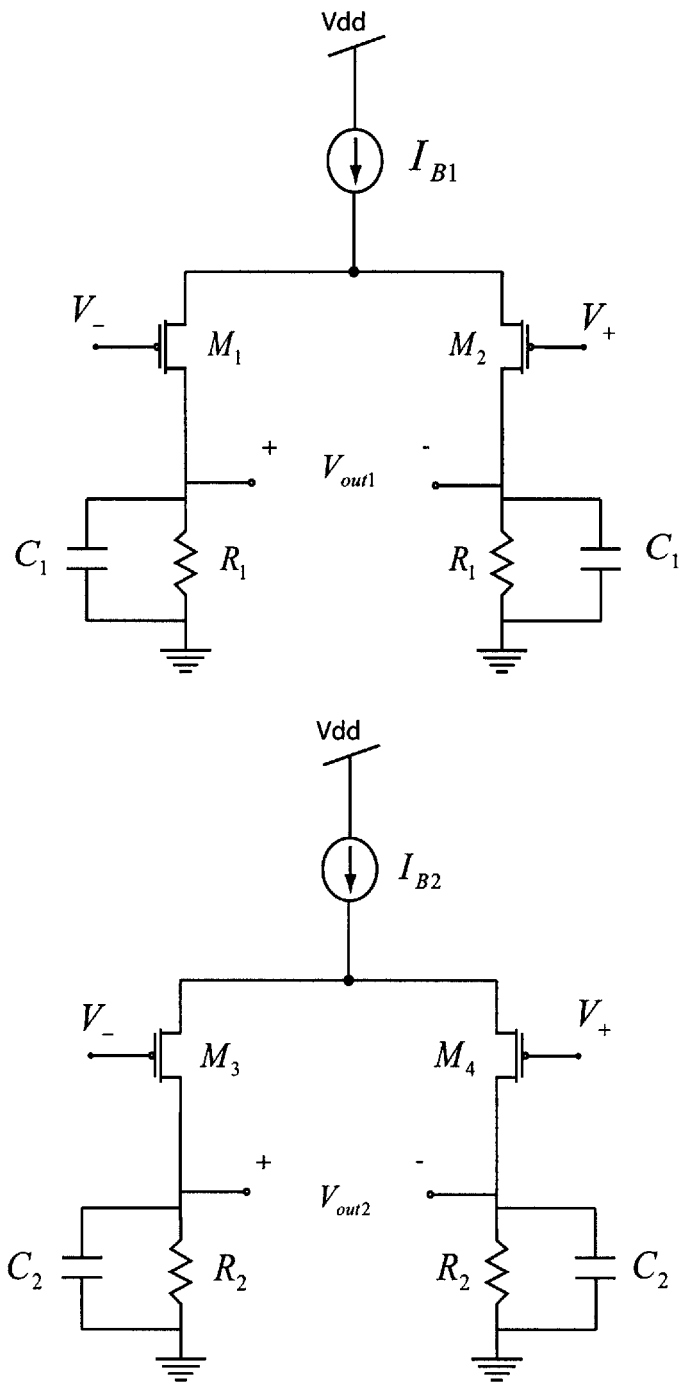


Figure 2-6: The implementation of each amplifier stage: A_1 (Top); A_2 (Bottom)

2.3 Robustness problems due to the uses of MOS-bipolar pseudoresistor elements

It was found during experiments that the feedforward distributed-gain amplifier discussed earlier was not robust to large fluctuations in the input voltages. When the BNC input cables were disconnected and reconnected from one of the amplifier's input terminals, the output signal disappeared for many minutes before the amplifier resumed the normal operation. Furthermore, when a step change in an input offset voltage of a few hundreds of millivolts was intentionally applied, the amplifier exhibited the same behavior. This behavior is a severe problem for the neural recording system that needs to operate continuously once it is turned on.

The problem arises because the feedforward distributed-gain amplifier does not have the mechanism to control its DC input operating point. Let's consider the situation depicted in Fig. 2-7 in which a large step change in the DC input voltage is applied at one of the input terminals. Suppose the amplifier is in a steady state for $t < 0$, that is, $V_{g1} = V_{g2} = V_B$, thus M_1 and M_2 both carry the same current of $I_{B1}/2$. At $t=0$, the input terminal on the V_{g1} side experiences a step change in voltage of ΔV_{ref} . For simplicity, let's ignore any parasitic capacitance at the gate of M_1 . The voltage across a capacitor cannot change instantaneously, thus, at $t = 0$, $V_{g1}(t = 0) = V_B + \Delta V_{ref}$. For $t > 0$, V_{g1} is discharged through r_a , thus we can set up a differential equation for describing V_{g1} after $t = 0$ to be

$$\frac{V_{g1} - V_B}{r_a} = C_{in1} \cdot \frac{d(V_{in} - V_{g1})}{dt} \quad (2.17)$$

with an initial condition

$$V_{g1}(t = 0) = V_B + \Delta V_{ref}. \quad (2.18)$$

We then solve for V_{g1} as a function of time as

$$V_{g1}(t) = V_B + \Delta V_{ref} \cdot e^{-t/(r_a C_{in1})}. \quad (2.19)$$

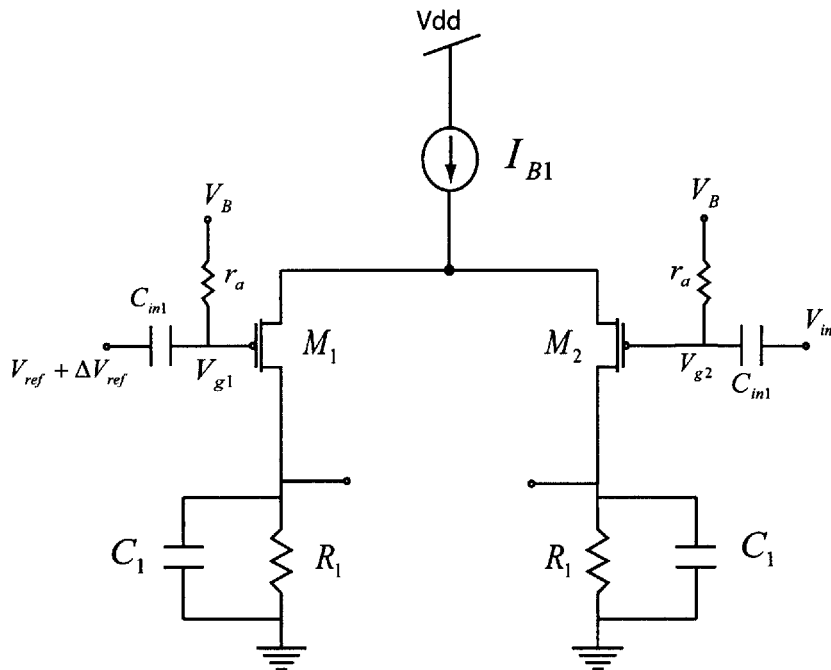


Figure 2-7: Schematic of the first stage where a step change in DC offset of ΔV_{ref} is applied.

Since $V_{g2}(t) = V_B$ for $t > 0$, the differential input voltage of the first stage amplifier is described by

$$V_{diff} = V_{g1}(t) - V_{g2}(t) = \Delta V_{ref} \cdot e^{-t/(r_a C_{in1})}. \quad (2.20)$$

During normal operation, an input differential pair must have its input differential voltage V_{diff} to be within its input linear range, which is approximately 150 mV for a subthreshold differential pair. If V_{diff} exceeds this linear range, all the bias current I_{B1} will flow in only one of the differential-pair transistors and the amplifier will lose all its incremental gain. Let's consider when the step change ΔV_{ref} exceeds 150 mV. At $t = 0$, $V_{diff} = \Delta V_{ref}$ thus the amplifier has no gain at this instant of time. From (2.20), V_{diff} slowly decays toward zero with a time constant of $r_a C_{in1}$ as t increases. Since r_a is very high due to a small gate-source voltage of a MOS-bipolar pseudoresistor element, the time constant $r_a C_{in1}$ is very large. Therefore it takes a long time before V_{diff} decays to within the linear range of the input differential pair, thus taking a long time before the amplifier resumes its normal amplification.

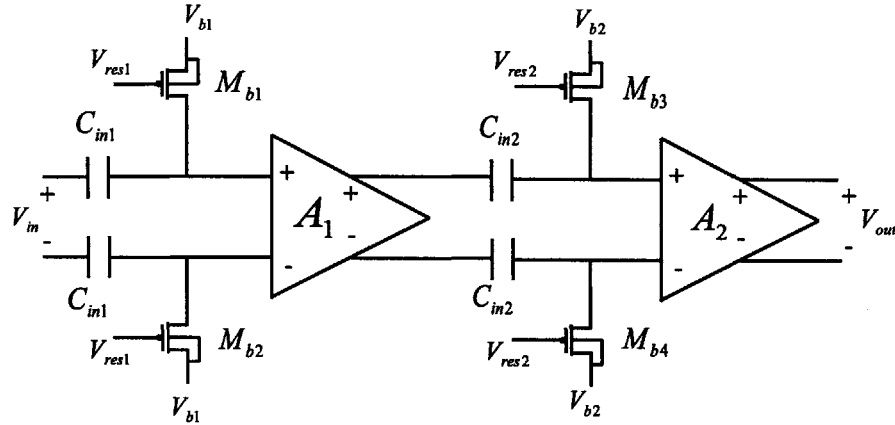


Figure 2-8: Schematic of the feedforward distributed-gain amplifier using MOS transistors in linear region to set the amplifier's DC operating point.

Even after V_{diff} already decays to be within the linear range of the differential pair, it would take much longer before it decays to almost zero. This is because the incremental resistance r_a increases significantly since the pseudo-resistor element's gate-source voltage becomes smaller. It was observed during the experiments that the amplifier always exhibited a large input-referred offset voltage even when it was amplifying the input signal normally.

2.4 A feedforward distributed-gain amplifier using linear-region MOS transistors for setting the amplifier's DC operating point

To solve the robustness problem mentioned in 2.3, MOS-bipolar pseudo-resistor elements were replaced by MOS transistors operating in linear region to increase the incremental resistance of these biasing elements. The schematic of an improved feedforward distributed-gain amplifier is shown in Fig. 2-8. The gate-source voltages $V_{b1} - V_{res1}$ and $V_{b2} - V_{res2}$ are set to be about 500 mV. As a result M_{b1} - M_{b4} operate in linear region and their drain-source incremental resistances are not as high as those of the MOS-bipolar pseudo-resistor elements. As a result, the gate voltages of the

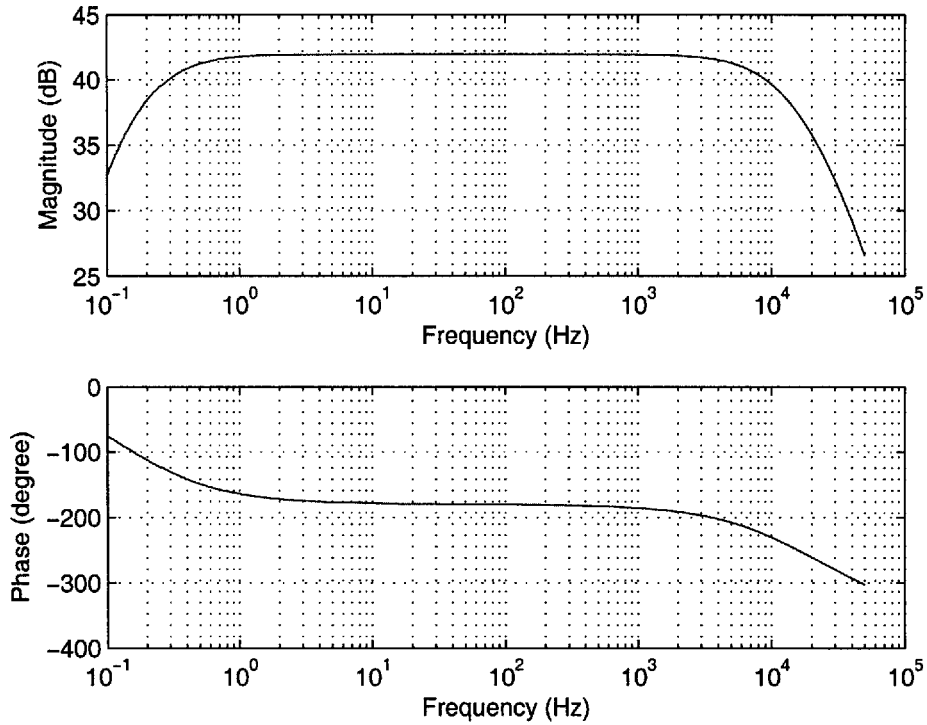


Figure 2-9: The transfer function of the feedforward distributed-gain amplifier using MOS transistors in linear region to set the amplifier's DC operating point.

input differential-pair transistors can decay to V_{b1} much faster in the event of a step change in input DC offset voltages because the time constants created by M_{b1} - M_{b4} in Fig. 2-8 are much smaller. In this topology, the high-pass cutoff frequency that is used to reject the DC offset voltages appears at much higher frequency than in the previous topology.

From the experiments, the amplifier was able to recover quickly when a step change in input voltage was applied. The transfer function of the amplifier is shown in Fig. 2-9. Unfortunately, the use of linear-region MOS transistors for setting the DC operating points of the input differential-pair transistors poses another severe problem. It appears that the amplifier exhibits much larger low-frequency noise than it was expected. The noise spectral density at low frequency rolls off as $1/f^2$ in power unit instead of $1/f$ if it is a flicker noise. An input-referred noise spectral

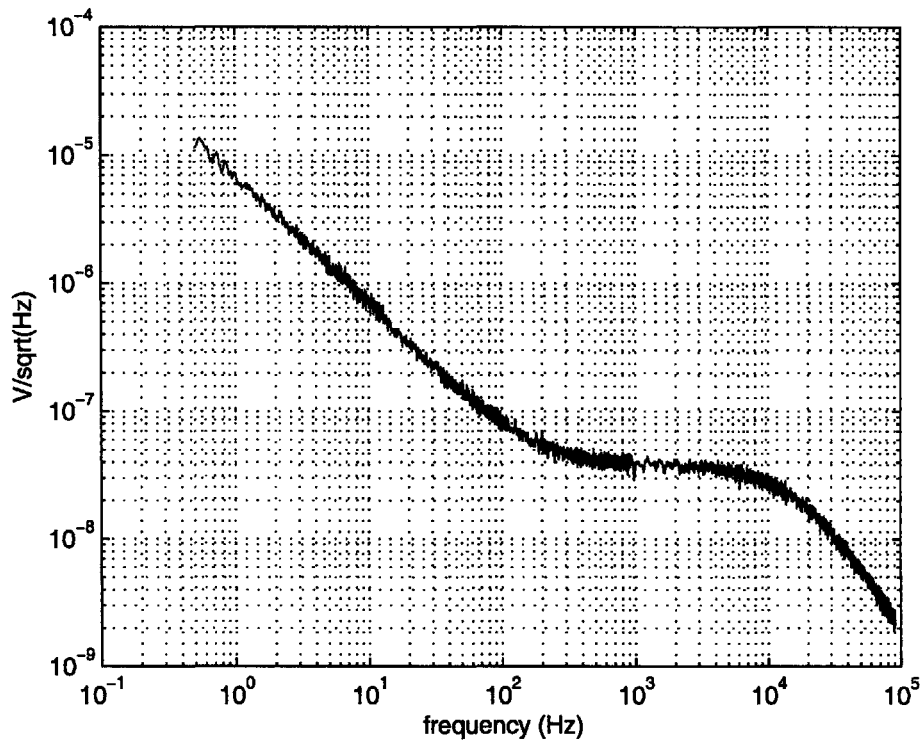


Figure 2-10: An input-referred noise spectral density of the amplifier.

density of this amplifier showing a significant amount of low-frequency noise is shown in Fig. 2-10.

With some analysis, it appears that the low-frequency noise that rolls off as $1/f^2$ in power unit is due mainly to the filtering of the thermal noise in M_{b1} and M_{b2} . These thermal noise sources are at the very frontend of the amplifier which is the most critical stage. To understand why such $1/f^2$ noise is present, let's calculate a part of the amplifier's input-referred noise that is contributed by M_{b1} and M_{b2} . A circuit schematic illustrating this situation is shown in Fig. 2-11. For simplicity, let's assume that M_{b1} and M_{b2} have the same noise current and the same incremental resistance (denoted r_{a1}) since they are biased at the same operating point. Similarly, let's assume that M_{b3} and M_{b4} have the same incremental resistance denoted r_{a2} . We can safely ignore the noise contributions from M_{b3} and M_{b4} since the gain of the first stage makes their noise contributions insignificant. The input-referred noise

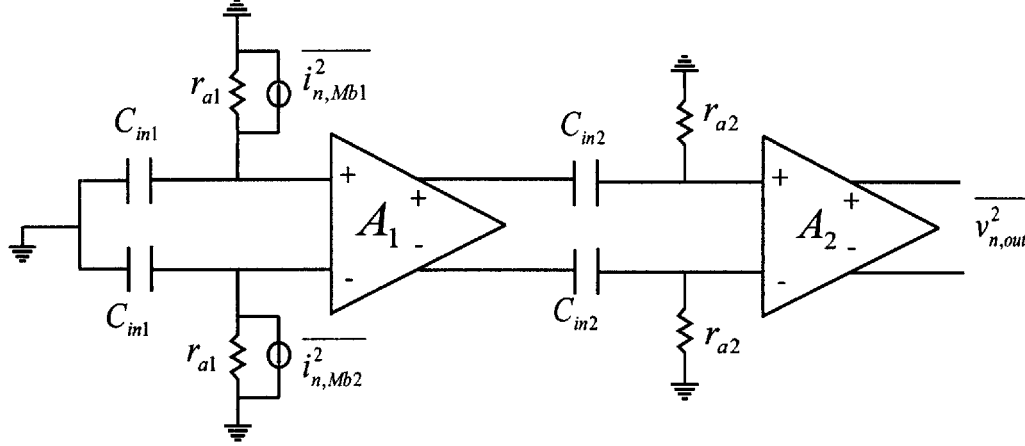


Figure 2-11: A schematic for calculating the input-referred noise from noise of M_{b1} and M_{b2} .

contributed by M_{b1} and M_{b2} is calculated to be

$$\overline{v_{n,in}^2} = (i_{n,Mb1}^2 + i_{n,Mb2}^2) \cdot \left(\frac{r_{a1}}{1 + sr_{a1}C_{in1}} \cdot \frac{sr_{a2}C_{in2}}{1 + sr_{a2}C_{in2}} \right)^2. \quad (2.21)$$

The transistors M_{b1} and M_{b2} are biased in subthreshold such that the highpass cutoff frequency of the amplifier is below 1 Hz. Therefore, the thermal noise in M_{b1} and M_{b2} can be approximated by

$$\overline{i_{n,Mb1}^2} = \overline{i_{n,Mb2}^2} = 2qI_1 (1 + e^{-V_{ds1}/U_T}) \approx 4qI_1 \quad (2.22)$$

since their drain source voltage V_{ds1} is approximately zero and I_1 is the channel current of M_{b1} if it is in saturation. For the frequency range of $f \gg 1/(2\pi r_{a1}C_{in1}), 1/(2\pi r_{a2}C_{in2})$, the input-referred noise in (2.21) can be approximated by

$$\overline{v_{n,in}^2} = \frac{8qI_1}{(2\pi)^2 \cdot C_{in1}^2} \cdot \frac{1}{f^2} \quad (2.23)$$

which agrees well with the $1/f^2$ rolloff at low frequency of the noise spectral density in Fig. 2-11. Therefore, in order to reduce this low-frequency noise in the passband, the saturation current of M_{b1} and M_{b2} should be made very small, thus M_{b1} and M_{b2} should have as small gate-source voltages as possible. However, this would lead to

the robustness problem due to a very long time constant as described in Section 2.3.

2.5 Conclusion

Despite the fact that the feedforward distributed-gain amplifier promises a very low-noise operation since most of the input-referred noise is mainly from the two input differential-pair transistors, the AC-coupling and DC input biasing of such amplifier pose many problems that prevent it to be used in the real recording environments. Nevertheless, the design of the feedforward distributed-gain amplifier provides many useful insights. First, if the neural amplifier is to be AC-coupled, the pseudoresistor elements should be made with the incremental resistance as high as possible such that the thermal noise in these pseudoresistor elements is filtered out well before the pass-band. Thus, the MOS-bipolar pseudoresistor elements should be used. Even though the use of such elements leads to robustness problem in a feedforward distributed-gain amplifier, this problem can easily be solved if the feedback topology with a high loop gain is used.

Chapter 3

A Feedback Neural Amplifier

Using a Folded-Cascode OTA as the Gain Stage

Using the linear-region MOS transistors to set the DC operating points of the feedforward distributed-gain amplifier poses a severe problem since the thermal noise in the linear-region MOS transistors appears at the frontend which is the most critical stage of any low-noise amplifier. Instead of achieving a low-noise performance, the feedforward distributed-gain amplifier in Chapter 2 appears to have a much higher total integrated input-referred noise than it was originally desired due to these biasing elements. By setting the gate-source voltages of the linear-region MOS transistors such that the highpass cutoff frequency of the amplifier happens at a very low frequency, the thermal noise in these linear-region MOS transistors can be filtered out well before the frequency band of interest. However, the robustness of the amplifier is compromised due to a very slow time constant caused by the high incremental resistance of these biasing elements. If there is any large fluctuation at the input of the amplifier during recording such as the movements of the electrode that cause the DC offset voltage at the electrode-tissue interface to change abruptly, the amplifier may stop amplifying for a period of several minutes before it resumes normal operation. This behavior is intolerable for a recording system which needs to operate

continuously once it is turned on. Therefore, a new amplifier that exhibits a lower input-referred noise and is also robust to changes in the recording environment is needed.

This chapter describes a design of a feedback amplifier that uses a folded-cascode operational transconductance amplifier (OTA) as the gain stage. The folded-cascode OTA offers many advantages over other OTA topologies for low-frequency applications if it is used in a feedback topology with a high closed-loop gain. The first advantage is that the frequency compensation of the feedback amplifier can be achieved with a simple dominant-pole compensation at the output since the internal nodes of the OTA have low impedances. Thus the non-dominant poles always appear at much higher frequencies than the dominant pole. Furthermore, the output impedance of the folded-cascode OTA is very high due to cascoding of the output stage, thus only one gain stage is needed to achieve a desired open-loop gain. The most important advantage is that for low-frequency applications such as in neural recordings, the current in the folded branch of the OTA can be made much lower than the current in the input differential-pair transistors without affecting the stability of the overall feedback amplifier. Lowering the current in the folded branch has two main benefits. First, the total power consumption of the OTA decreases. Second, the noise contributions from the transistors in the folded branch decrease due to a lower current level if the overall transconductance of the OTA can be maintained. The design presented in this chapter makes use of this technique to try to simultaneously reduce the power consumption and the input-referred noise of the amplifier. However, the fabricated amplifier exhibited poor performance since many design issues were overlooked. These problems will be addressed at the end of this chapter and are very important for a successful design of an energy-efficient amplifier that will be presented in Chapter 4.

3.1 Overall Amplifier Design

The high-level schematic of the amplifier is shown in Fig. 3-1. The MOS-bipolar pseudoresistor elements M_{b1} and M_{b2} are used to set the DC operating point of the

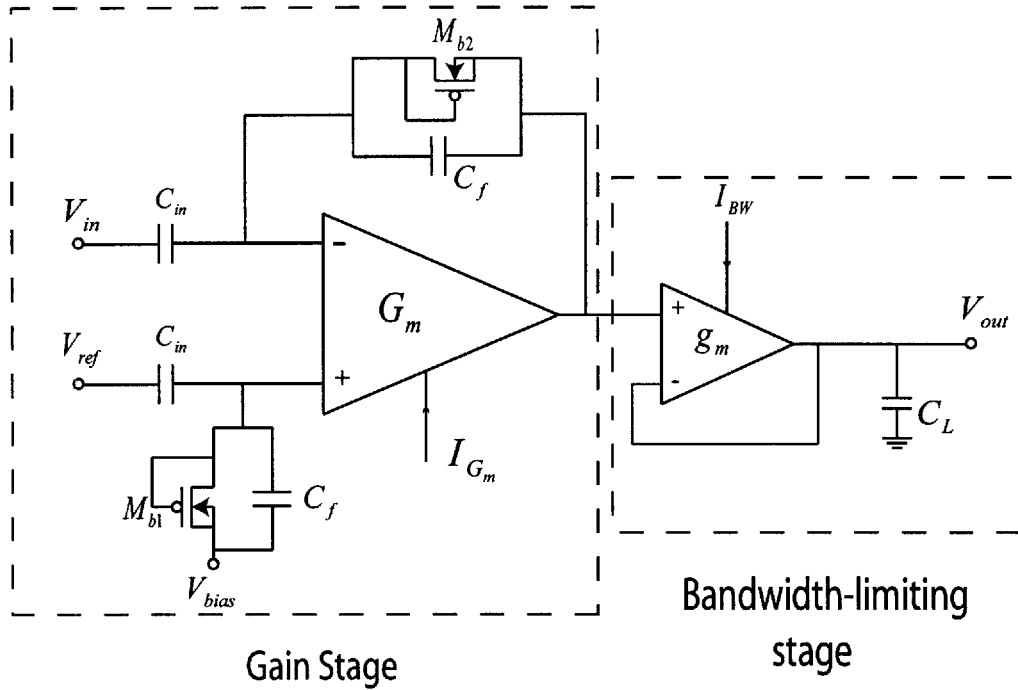


Figure 3-1: A high-level schematic of the feedback neural amplifier.

amplifier. To understand why this feedback topology does not suffer from the robustness problem described in Chapter 2, let's consider the situation when there is a large fluctuation in the DC offset voltage at the recording site. Suppose that V_{ref} experiences a voltage excursion of ΔV_{ref} . At the moment the voltage excursion occurs, the positive terminal's voltage of the G_m OTA will be at $V_+ = V_{bias} + \frac{C_{in}}{C_{in} + C_f} \cdot \Delta V_{ref}$. If the feedback path formed by M_{b2} and C_f is not present and ΔV_{ref} is larger than the input linear range of G_m OTA, one of the transistors in the input differential pair of G_m OTA will carry all the bias current, making the amplifier to lose all its gain. Now let's consider when the feedback path is present. At the moment the input voltage excursion occurs, the G_m OTA has a large differential input voltage. Therefore, the output of the G_m OTA quickly moves toward and stays at one of the supply rails since the OTA has a very high gain. As a result, M_{b2} will have a large gate-source voltage. During this phase, M_{b2} no longer acts as a high-resistance element but becomes either a diode-connected MOS transistor or a diode-connected BJT depending on the output voltage polarities. The turned-on M_{b2} then quickly charges the voltage at the

negative terminal V_- of the G_m OTA such that it becomes close to V_+ once again. As a result, the feedback topology can adjust to the fluctuations at the recording site much faster than the feedforward amplifier that uses the MOS-bipolar pseudoresistor elements to set the DC operating points. It was verified during the experiments that a large step change in DC input voltage does not cause the feedback amplifier to stop amplifying. Thus, this feedback amplifier is suitable for use in a real recording situation due to its robustness to the recording site's fluctuations.

3.2 Small-Signal Analysis

Let's analyze the operation of the amplifier in the Laplace's domain with the feedback block diagram approach. First, let us consider the operation of the gain stage. Let assume that the transfer function of the G_m OTA can be approximated by

$$A(s) = \frac{G_{m,eff}R_o}{1 + sR_oC_{L,p}} \quad (3.1)$$

where $G_{m,eff}$ and R_o are the effective total transconductance and the output resistance of the G_m OTA respectively. The loading effect at the output node of the gain stage is modeled as a $C_{L,p}$ parasitic capacitance connecting between the output node of the gain stage to an incremental ground. Let $C_{in,p}$ denote the parasitic capacitance connecting between the negative terminal of the G_m OTA to an incremental ground. Let v_- denote the small-signal voltage at the negative terminal of G_m OTA. Furthermore, let r_a denote the incremental resistance of M_{b2} when its gate-source voltage is close to zero. The circuit diagram for analyzing the operation of the gain stage is shown in Fig. 3-2. We can write v_- as a superposition of v_{in} and $v_{o,1}$ as

$$v_- = \frac{\frac{1}{sC_{in,p}} \parallel \frac{r_a}{1+s r_a C_f}}{\frac{1}{sC_{in}} + \left(\frac{1}{sC_{in,p}} \parallel \frac{r_a}{1+s r_a C_f} \right)} \cdot v_{in} + \frac{\frac{1}{s(C_{in,p}+C_{in})}}{\frac{1}{s(C_{in,p}+C_{in})} + \frac{r_a}{1+s r_a C_f}} \cdot v_{o,1} \quad (3.2)$$

$$= \frac{s r_a C_{in}}{1 + s r_a (C_{in} + C_f + C_{in,p})} \cdot v_{in} + \frac{1 + s r_a C_f}{1 + s r_a (C_{in} + C_f + C_{in,p})} \cdot v_{o,1}. \quad (3.3)$$

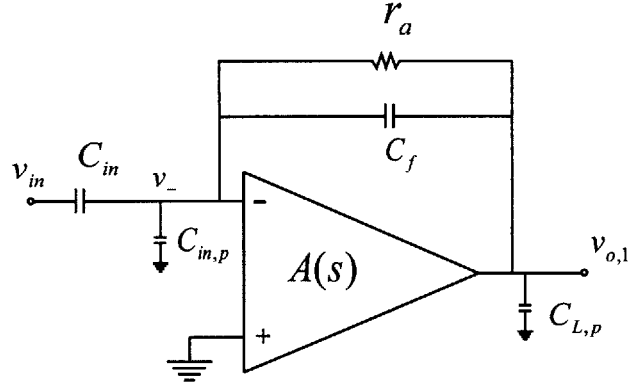


Figure 3-2: A circuit schematic for analyzing the operation of the folded-cascode gain stage.

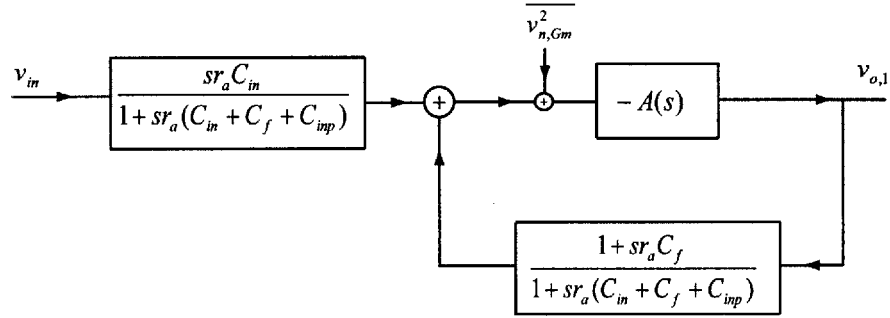


Figure 3-3: A preliminary block-diagram describing the operation of the feedback amplifier.

We can also write $v_{o,1}$ as a function of v_- as

$$v_{o,1} = -A(s) \cdot v_- \quad (3.4)$$

Equations (3.2)-(3.4) can be captured in a feedback block diagram shown in Fig. 3-3. The input-referred noise of the G_m OTA is included in the block diagram with the $\overline{v_{n,Gm}^2}$ term being added to the input of the G_m OTA, where $\overline{v_{n,Gm}^2}$ represents the input-referred noise per unit bandwidth of the G_m OTA. However, the noise analysis of this amplifier is deferred until Section 3.3. The block diagram in Fig. 3-3 can be simplified into a unity-gain feedback form as shown in Fig. 3-4. In practice, the pole denoted by $1/(r_a C_f)$ is at a very low frequency (on the order of a few millihertz). We can thus consider the operation of the amplifier when the frequency of operation

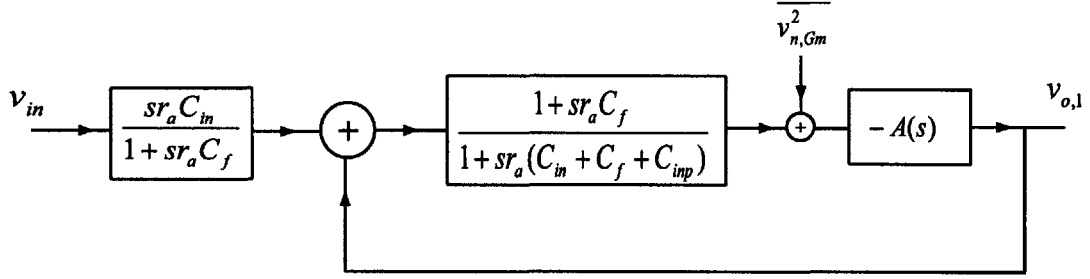


Figure 3-4: A unity-gain feedback block diagram describing the operation of the feedback amplifier.

$\omega \gg 1/(r_a(C_{in} + C_f + C_{inp}))$. Then the term $\frac{1+sr_a C_f}{1+sr_a(C_{in}+C_f+C_{inp})}$ can be approximated by $C_f/(C_{in} + C_f + C_{inp})$. Using (3.1) and the Black's formula, we can estimate the transfer function of the gain stage to be

$$\frac{v_{o,1}}{v_{in}}(s) \approx -\frac{sr_a C_{in}}{1 + sr_a C_f} \cdot \frac{1}{1 + sC_{L,p}A_{CL}/G_{m,eff}} \quad (3.5)$$

where $A_{CL} = \frac{C_{in}+C_f+C_{inp}}{C_f} \approx \frac{C_{in}}{C_f}$ is the closed-loop gain of the amplifier, assuming that $C_{in} \gg C_f, C_{inp}$. Equation (3.5) suggests that the highpass cutoff frequency due to AC coupling is at $f_h = 1/(r_a C_f)$ and the lowpass cutoff frequency due to the loading effect at the output of the G_m OTA is at $f_L = G_{m,eff}/(2\pi A_{CL}C_{L,p})$. In our design, we want f_L to be as high as possible so that it does not limit the bandwidth of the amplifier. Instead, the bandwidth-limiting stage is designed to provide a fixed lowpass cutoff frequency. In this way, we can vary the bias current of the gain stage without affecting the overall bandwidth as long as the pole due to the loading effect at the output of the gain stage is at a much higher frequency than the pole provided by the bandwidth-limiting stage. The transfer function of the bandwidth-limiting stage is provided by

$$\frac{V_{out}}{v_{o,1}}(s) = \frac{1}{1 + sC_L/g_m}. \quad (3.6)$$

If $g_m/C_L \ll G_{m,eff}/(2\pi A_{CL}C_{L,p})$ the bandwidth of the amplifier is controlled by the bandwidth-limiting stage. Therefore, the overall transfer function of the amplifier

can be expressed as

$$\frac{V_{out}}{v_{in}}(s) = -\frac{sr_a C_{in}}{1 + sr_a C_f} \cdot \frac{1}{1 + sC_L/g_m}. \quad (3.7)$$

At a midband frequency in which $1/(r_a C_f) < \omega < g_m/C_L$, the gain of the amplifier can be approximated by

$$A_M = -\frac{C_{in}}{C_f}. \quad (3.8)$$

As a result, the mid-band gain of the amplifier is controlled by the ratio of two capacitors and can be well controlled.

3.3 Noise Analysis

The amplifier can be thought of as a cascade of two amplifiers. The first stage is the gain stage which provides a midband gain of approximately 40 dB. The second stage is the bandwidth-limiting stage which is designed to provide a fixed lowpass cutoff frequency at 5 kHz. Since the input signal has been gained up by approximately 100× by the gain stage, the input-referred noise contribution from the bandwidth-limiting stage is insignificant. Therefore, we can ignore the input-referred noise of the bandwidth-limiting stage in the following noise calculation. From the feedback block diagram of Fig. 3-4, we can estimate the input-referred noise of the overall amplifier as

$$\overline{v_{n,amp}^2} = \left(\frac{C_{in} + C_f + C_{in,p}}{C_{in}} \right)^2 \cdot \overline{v_{n,Gm}^2}. \quad (3.9)$$

Equation (3.9) emphasizes the importance of the parasitic capacitance $C_{in,p}$ at the negative input terminal of the OTA. While making the input differential-pair transistors large may reduce $1/f$ noise in the amplifier, the parasitic capacitances of large input devices can degrade the input-referred noise of the overall amplifier according to (3.9).

To achieve a low-noise performance, the input-referred noise $\overline{v_{n,Gm}^2}$ of the gain stage OTA must be minimized. This section discusses the low-noise techniques that

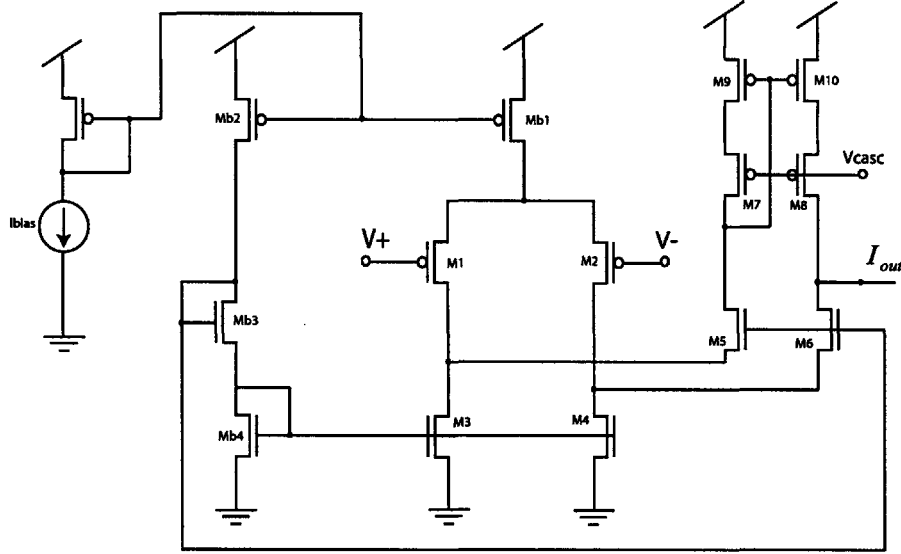


Figure 3-5: A folded-cascode OTA schematic used in this design.

are used in this design and also the implementation problems that prevent this design from achieving an optimal performance. The schematic of the folded-cascode OTA used in the gain stage is shown in Fig. 3-5. The OTA itself can be thought of as a two-stage amplifier. The first stage is the transconductance stage that has a voltage input and a current output. The second stage is a common-gate amplifier stage that takes in an input current and converts this current into a voltage at the output. The transconductance stage composes of M_{b1} and M_1 - M_4 while the common-gate amplifier stage composes of M_5 - M_{10} . We can express the folded-cascode OTA by their equivalent small-signal diagram as shown in Fig. 3-6. In Fig. 3-6, R_{o1} and R_{o2} are the output resistance of the transconductance stage and output resistance of the common-gate amplifier stage respectively and they can be approximated by

$$R_{o1} = r_{o2} || r_{o4} \quad (3.10)$$

and

$$R_{o2} \approx ((g_{s8}r_{o8})r_{o10}) || ((g_{s6}r_{o6})(r_{o2} || r_{o4})) \quad (3.11)$$

where r_{oi} and g_{si} are the Early-Effect resistors and the incremental source admittance

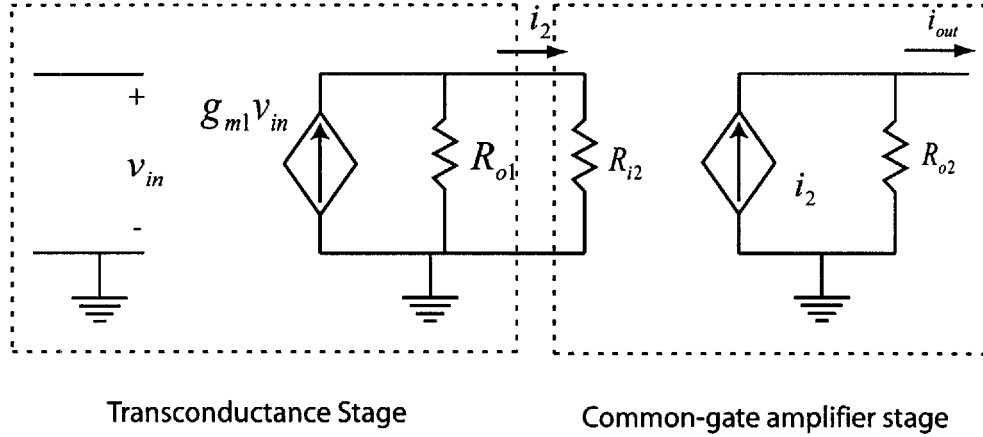


Figure 3-6: A small-signal schematic for describing the operation of folded-cascode OTA.

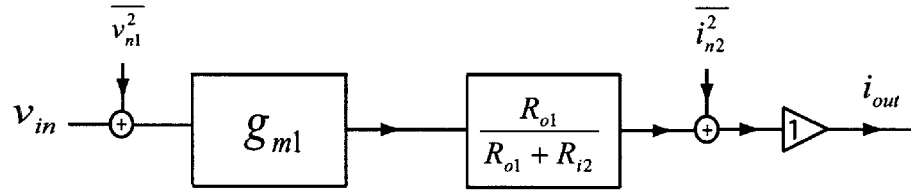


Figure 3-7: A small-signal block diagram describing the operation of folded-cascode OTA.

of M_i respectively. The resistance R_{i2} is the input resistance of the common-gate amplifier stage which can be approximated by

$$R_{i2} = \frac{1}{g_{s5}} \quad (3.12)$$

where g_{s5} is the incremental source admittance of M_5 and M_6 .

The noise analysis of the OTA can be best understood by the small-signal block diagram shown in Fig. 3-7. The amount of the transconductance stage's output current that flows into the source of M_5 and M_6 is determined by the current divider formed by R_{o1} and R_{i2} . The current that flows into R_{i2} appears directly at the output of the common-gate stage. This is described by a unity-gain buffer shown in Fig. 3-7. The input-referred noise of the transconductance stage is represented by $\overline{v_{n1}^2}$ while the input-referred noise of the common-gate stage which has a current input is represented with a current noise source $\overline{i_{n,2}^2}$.

The input-referred noise of the transconductance stage can be calculated to be

$$\overline{v_{n1}^2} = \frac{1}{g_{m1}^2} \cdot \left(\overline{i_{n,M1}^2} + \overline{i_{n,M2}^2} + \overline{i_{n,M3}^2} + \overline{i_{n,M4}^2} \right). \quad (3.13)$$

In order to minimize this input-referred noise, we shall maximize g_{m1} . Therefore, the input differential-pair transistors M_1 and M_2 are made with large W/L such that they operate in deep in subthreshold and achieve the maximum g_m for a given bias current. Even though M_3 and M_4 should be biased in strong inversion to reduce their g_m in order to reduce their noise contribution, in this design they operate in subthreshold so that their saturation voltages can be small. The amplifier was designed to work with a 2 V supply, thus minimizing the noise contributions from M_3 and M_4 by operating them well above threshold proved to be impractical. Thus, the input-referred noise of the transconductance stage can be expressed in terms of the transistors' small-signal parameters as

$$\overline{v_{n1}^2} = \frac{2kT}{\kappa g_{m1}} \cdot \left(2 + 2 \times \frac{g_{m3}}{g_{m1}} \right). \quad (3.14)$$

To simplify the input-referred noise calculation of the common-gate amplifier stage, we make an assumption that the noise contributions from M_5 - M_8 are negligible since they act as cascode transistors and these transistors self-shunt their own current noise sources. Thus the transistors in the common-gate amplifier stage that significantly contribute noises are M_9 and M_{10} . Due to supply voltage constraint, M_9 and M_{10} are also biased in weak-inversion such that they can operate with small saturation voltages. Thus, the input-referred current noise of the common-gate amplifier stage can be expressed as

$$\overline{i_{n2}^2} = \overline{i_{n,M9}^2} + \overline{i_{n,M10}^2} \quad (3.15)$$

$$= 2 \times \frac{2kT}{\kappa} g_{m9}. \quad (3.16)$$

Let $G_{m,eff} = i_{out}/v_{in}$ be an effective total transconductance of the folded-cascode

OTA. From the circuit diagram in Fig. 3-6, $G_{m,eff}$ can be calculated to be

$$G_{m,eff} = \frac{i_{out}}{v_{in}} = g_{m1} \cdot \frac{R_{o1}}{R_{o1} + R_{i2}}. \quad (3.17)$$

Thus the total input-referred voltage noise of the OTA can be expressed as

$$\overline{v_{n,OTA}^2} = \overline{v_{n1}^2} + \frac{1}{G_{m,eff}^2} \cdot \overline{i_{n2}^2} \quad (3.18)$$

$$= \frac{4kT}{\kappa g_{m1}} \cdot \left(1 + \frac{g_{m3}}{g_{m1}} + \left(\frac{R_{o1} + R_{i2}}{R_{o1}} \right)^2 \cdot \left(\frac{g_{m9}}{g_{m1}} \right) \right). \quad (3.19)$$

In order to minimize the input-referred noise for a given bias current, g_{m1} should be maximized by running M_1 and M_2 in subthreshold. Furthermore, g_{m9} should be minimized and the current divider ratio $\frac{R_{o1}}{R_{o1} + R_{i2}}$ should be minimized. In this design, I made the current in M_9 and M_{10} to be much smaller than the current in M_1 and M_2 . In this way, the ratio g_{m9}/g_{m1} is made small compared to other terms in (3.19). Moreover, lowering the current in the folded branch makes the term g_{m3}/g_{m1} which is usually larger than 1 becomes close to 1 since the currents in M_3 and M_4 are almost the same as the current in M_1 and M_2 . For this topology, the ideal input-referred noise that can be achieved while all the transistors are operating in subthreshold is

$$\overline{v_{n,OTA}^2} = 4 \cdot \frac{2kT}{\kappa g_{m1}} \quad (3.20)$$

assuming that $g_{m3} \approx g_{m1}$ and $g_{m9}/g_{m1} \ll 1$. The ideal input-referred noise in (3.20) is equivalent to the input-referred noise of an OTA with effectively four subthreshold devices that contribute noise.

3.4 Measurement Results

The amplifier was fabricated in a commercial $0.5 \mu m$ CMOS process. A micrograph showing the amplifier is shown in Fig. 3-8. All the measurements in this Section were obtained from a Stanford Research System SR785 Signal Analyzer. The amplifier

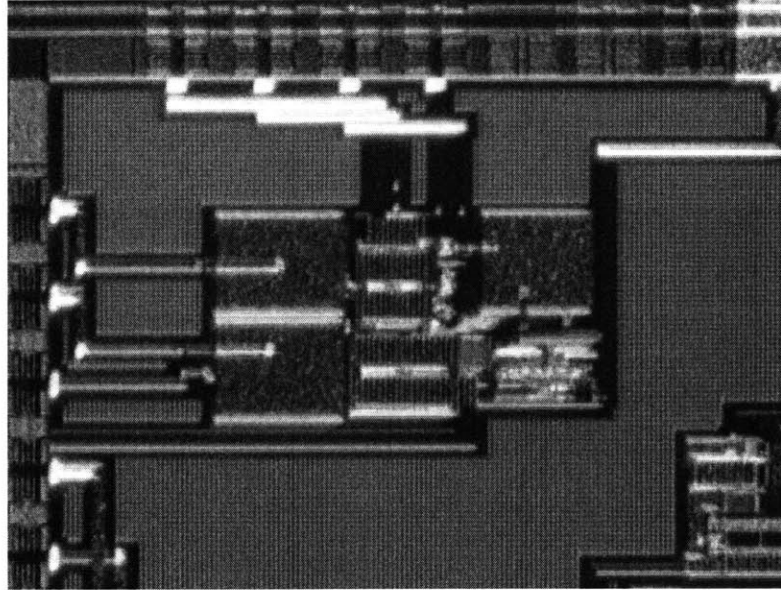


Figure 3-8: A micrograph of the feedback amplifier using a current-scaled folded cascode OTA as a gain stage.

was biased with a total supply current of $3.5 \mu\text{A}$ from a 2 V supply voltage. The bandwidth-limiting stage provided a lowpass cutoff frequency of 5 kHz. The transfer function of the amplifier is shown in Fig. 3-9 where the midband gain was measured to be 41.57 dB and its highpass cutoff was measured to be approximately at 1 mHz. The measured input-referred noise power spectral density (output noise spectrum divided by the midband gain) of the amplifier is shown in Fig. 3-10. The total input-referred noise of the amplifier was calculated to be $5.5 \mu\text{V}_{rms}$ by integrating the noise spectral density in Fig. 3-10 from 0.5 Hz to 90 kHz. The performance characteristics from the test-bench experiments are summarized in Table 3.1. This amplifier was also used to successfully record neural action potentials (neural spikes) from the RA region in a zebra finch's brain with a Carbostar electrode. The long-time trace and short-time trace of the recording normalized to the gain of the amplifier are shown in Fig. 3-11. Notice that the waveforms shown in Fig 3-11 do not center at 0 V. We speculated that it was due to a low-frequency signal which was present as a background voltage of about $200 \mu\text{V}$. In a longer time scale, we noticed the frequency of this low-frequency signal to be approximately 300 mHz. The presence of this low-frequency signal was

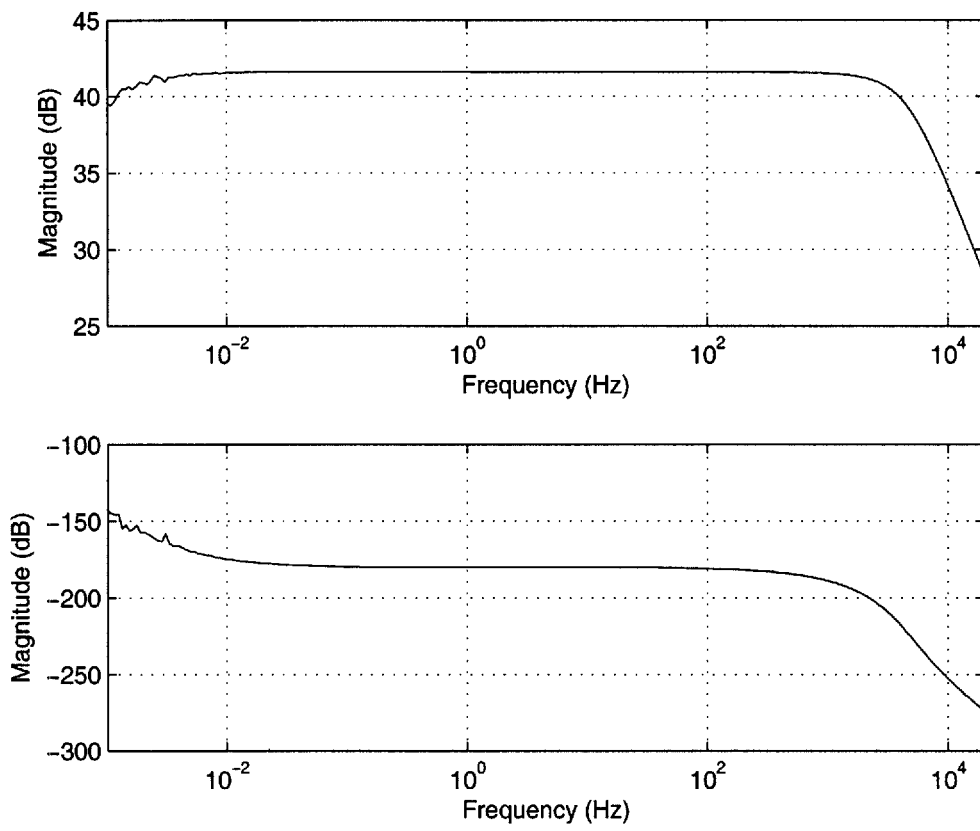


Figure 3-9: A measured transfer function of the feedback amplifier.

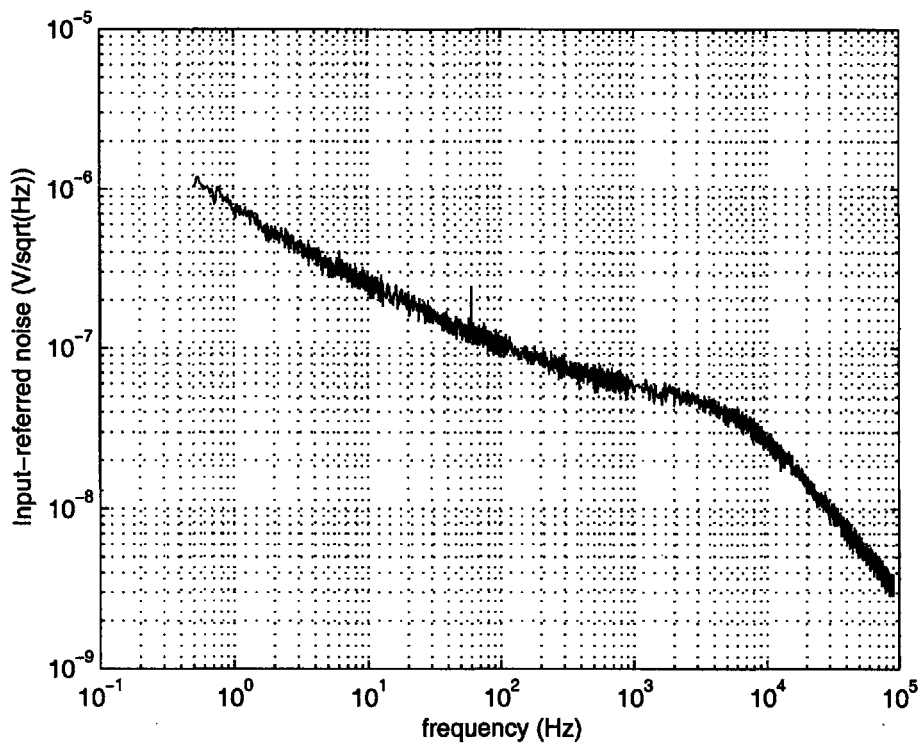


Figure 3-10: A measured input-referred noise spectral density of the feedback amplifier.

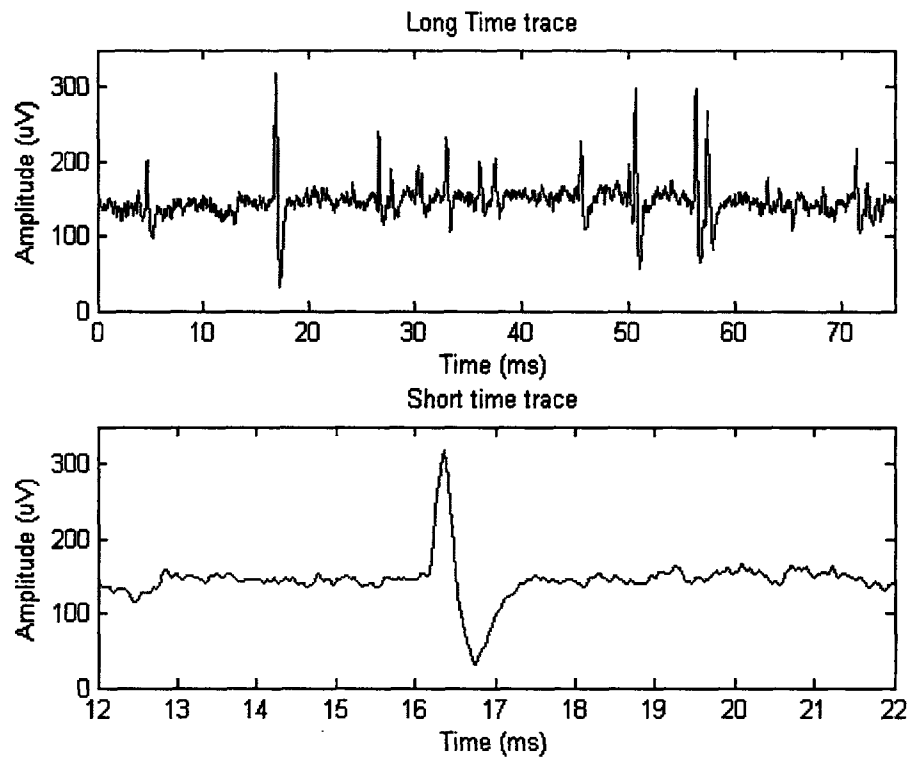


Figure 3-11: Neural recording from the RA region of a zebra-finch bird brain.

due to the fact that the highpass cutoff of the amplifier was at about 1 mHz, thus, the amplifier could not filter out this low-frequency signal.

Table 3.1: Measured Performance Characteristics for the feedback amplifier

Parameter	Measured
Supply voltage	2 V
Total current	3.5 μ A
Gain	41.7 dB
Bandwidth	1 mHz-5 kHz
Input-referred noise	5.5 μ V _{rms}
Max. signal (1% THD @ 1.024 kHz)	6 mV _{pp}
Dynamic Range (1% THD)	50 dB
CMRR (45 Hz-5.32 kHz)	55 dB
Area (in 0.5 μ m CMOS)	0.09 mm ²

3.5 Problems in The Implementation of This Design

Low Effective G_m due to Current Scaling

During the design phase of this amplifier, I did not realize the importance of the scaling term $\left(\frac{R_{o1}+R_{i2}}{R_{i2}}\right)^2$. In subthreshold, the source admittance of a MOS transistor is $g_s = I_D/U_T$, where I_D is the drain current of the transistor. Intuitively, as the currents in M_5 and M_6 become much smaller than the currents in M_3 and M_4 , the resistance looking into the sources M_5 and M_6 (R_{i2} in Fig. 3-7) becomes comparable to the resistance looking into the drains of M_3 and M_4 (R_{o1} in Fig. 3-7). The amount of the incremental current that can flow into the sources of M_5 and M_6 is determined by the current-divider ratio $R_{o1}/(R_{o1} + R_{i2})$. Therefore, if R_{i2} becomes larger, smaller incremental current can flow to the output of the folded-cascode OTA, resulting in a smaller value of $G_{m,eff}$. As a result, instead of the term $\left(\frac{R_{o1}+R_{i2}}{R_{o1}}\right)^2 \cdot \left(\frac{g_{m2}}{g_{m1}}\right)$ to be negligible compared to other terms in (3.19), it becomes significant since $\left(\frac{R_{o1}+R_{i2}}{R_{o1}}\right)^2$ is larger. Thus, the technique of making the current in the folded branch much smaller than the current in the input differential-pair transistors was not as effective

as desired.

Parasitic Capacitance at The Input Terminals of The Gain-Stage OTA

From (3.9), the parasitic capacitance at the input terminals of the OTA should be kept as small as possible in order to have low a input-referred noise. For this design, a critical mistake was made by connecting the bottom plates of the input capacitors C_{in} to the input terminals of OTA. The poly-poly capacitors used to realize the amplifier exhibit a parasitic capacitance between the poly1 bottom plate to the substrate of approximately 15-20% of the total capacitance. This mistake resulted in an increase of the input-referred noise by almost 20% from the simulated value.

3.6 Conclusion

A feedback amplifier with a high open-loop gain folded-cascode OTA was presented. Analysis shows that the feedback amplifier solved the robustness problem due to large the fluctuations in the input voltage encountered in the design of a feedforward distributed gain amplifier that uses MOS-bipolar pseudo-resistor elements to set the DC operating points. The low-power low-noise design technique that reduces the currents in the folded branch of a folded-cascode OTA was presented. Experimental results showed that in order to achieve the optimal performance with this topology, the overall transconductance of the folded-cascode OTA must be maintained even with the reduction of the current in the folded branch. Furthermore, the parasitic capacitances at the input terminals of the gain-stage OTA must be kept small to achieve low input-referred noise.

Chapter 4

An Energy-Efficient Micropower Neural Recording Amplifier

The technique of reducing the current in the folded branch of the folded-cascode OTA promises a very low-noise operation if the overall transconductance of the gain stage does not degrade due to current scaling. In this chapter, an improved version of the feedback amplifier in Chapter 3 that maximizes the overall transconductance even with a severe current scaling in the folded-branch is presented.

4.1 Overall System Design

The overall schematic of the improved neural amplifier is shown in Fig 4-1. The topology of the gain stage is similar to that in 3. This design includes a bandpass filter stage following the gain stage to shape the passband of the amplifier to filter out low-frequency signals present in the amplifier of Chapter 3. The low-frequency high-pass cutoff of the gain stage is created by the MOS-bipolar pseudoresistor element formed by M_{b1} - M_{b2} and the capacitance C_f . The capacitive feedback formed by C_f and C_{in} sets the midband gain of the amplifier to approximately 40.8 dB. The high-pass cutoff and the low-pass cutoff frequencies of the amplifier can be adjusted via V_{tune} and the bias current of the g_m -OTA in the bandpass-filter stage respectively. With the addition of the bandpass-filter stage, the amplifier can be configured to record either

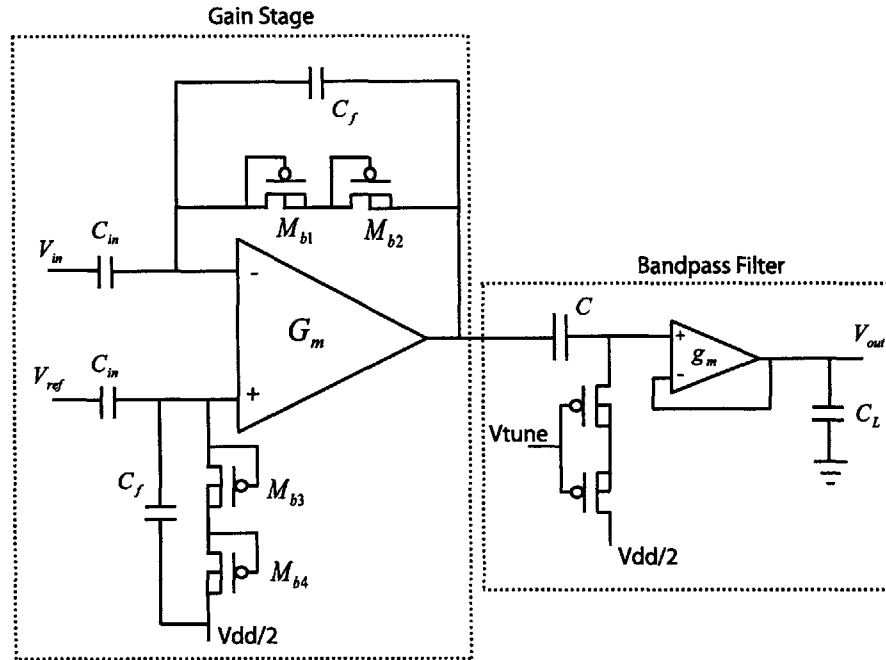


Figure 4-1: Overall System Schematic of the Neural Amplifier.

LFPs (< 1 Hz to 100 Hz) or neural spikes (100 Hz to > 1 kHz). For low-bandwidth LFP recording, the bias current of the OTA in the gain stage can be lowered to conserve power. It is worth mentioning that the high-pass cutoff frequency of the gain stage should be kept as low as possible. As reported in [9] and discussed in Chapter 2, placing a weak-inversion MOS transistor in parallel with C_f to create a high-pass filter with a cutoff frequency at a few hundred Hz introduces low-frequency noise that rolls off as $1/f^2$ in power units due to the noise from the transistor being low-pass filtered by C_f . This low-frequency noise appears at the front-end and gets amplified by the gain of the amplifier thereby degrading the minimum detectable signal. In our design as well as in [6], however, the MOS-bipolar pseudoresistor element's noise is at very low frequencies since the MOS-bipolar pseudoresistor element has a much higher impedance than a weak-inversion MOS transistor. Therefore, low-frequency noise due to this element is filtered out well before the passband and does not appear in the frequency band of interest.

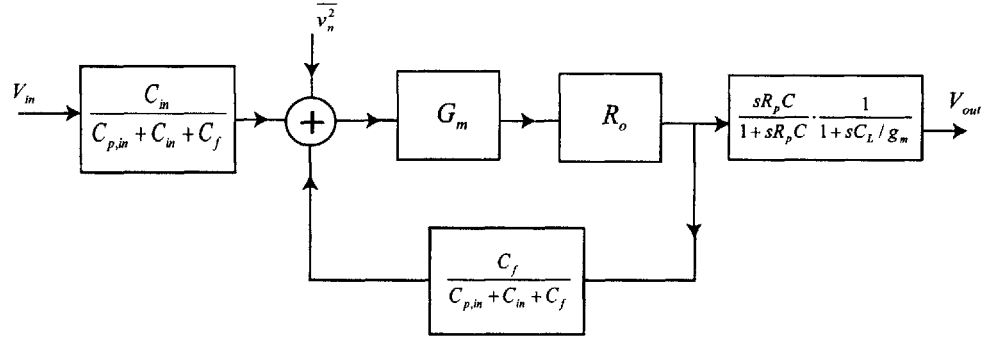


Figure 4-2: Block diagram of our neural amplifier including the input noise source of the OTA.

The operation of our amplifier can easily be understood by the block diagram of Fig 4-2. We include $C_{p,in}$ to model parasitic gate capacitances at input terminals of the gain-stage OTA. The input referred-noise of the OTA is modeled as a $\overline{v_n^2}$ term added to the system at the input of the gain-stage OTA. The gain-stage OTA is used as a high-gain amplifier and is modeled by G_m and R_o blocks where G_m and R_o represents the transconductance and the output resistance of the gain-stage OTA respectively. In the bandpass-filter stage, R_p is the resistance of the series PMOS transistors operating in the triode regime. The value of R_p is set by V_{tune} . The combination of C and R_p realizes the highpass cutoff frequency for the amplifier. From the small-signal block diagram in Fig. 4-2, assuming that $G_m R_o$ is much higher than 1, we can express the transfer function of the neural amplifier as

$$H(s) = \frac{v_{out}(s)}{v_{in}(s)} = -\frac{C_{in}}{C_f} \cdot \frac{sR_p C}{1 + sR_p C} \cdot \frac{1}{1 + sg_m C_L}. \quad (4.1)$$

The midband gain of the amplifier is $A_v = -C_{in}/C_f$. The highpass cutoff frequency is at $f_{HP} = 1/(2\pi R_p C)$ whereas the lowpass cutoff frequency is at $f_{LP} = g_m/(2\pi C_L)$. We can relate the input-referred noise $\overline{v_n^2}$ of the gain-stage OTA to the input-referred noise $\overline{v_{n,amp}^2}$ of the overall amplifier as

$$\overline{v_{n,amp}^2} = \left(\frac{C_{in} + C_f + C_{p,in}}{C_{in}} \right)^2 \cdot \overline{v_n^2}. \quad (4.2)$$

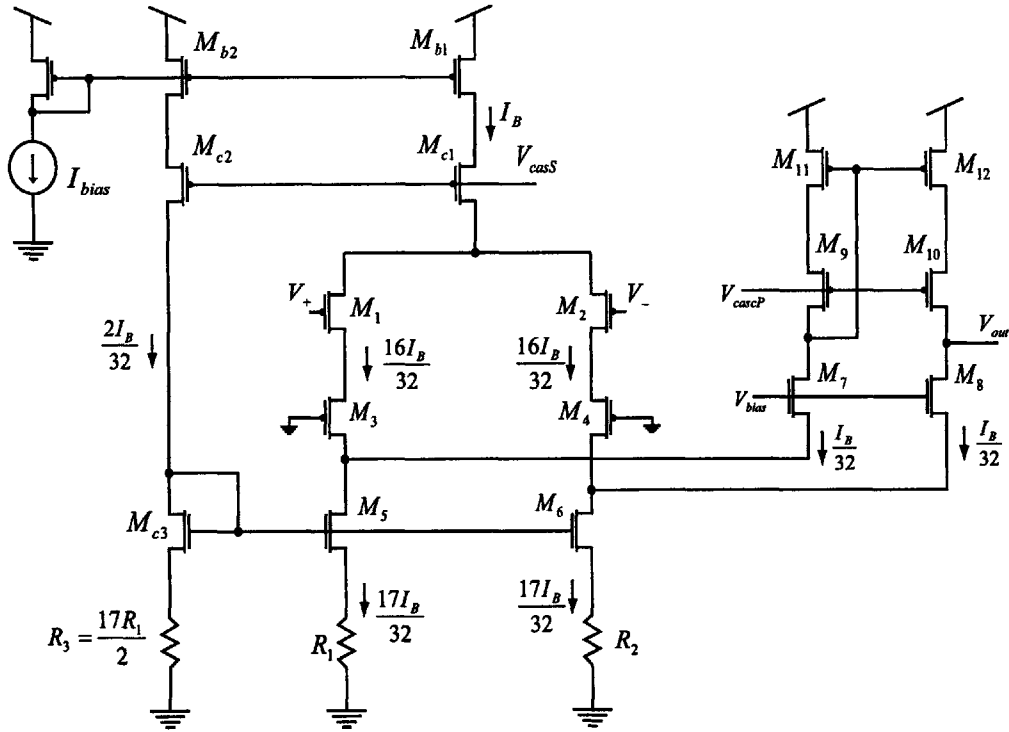


Figure 4-3: Schematic of the low-noise OTA used in this design.

The input-referred noise of the bandpass filter stage is insignificant and is not included in the block diagram since the gain of 40 dB of the gain stage alleviates the bandpass-filter stage's input-referred noise requirement. As a result, the power consumption of the bandpass filter stage is much smaller than that of the gain stage. Thus, to achieve low-noise performance, it is important to design the gain-stage OTA to have low input-referred noise. Section 4.2 describes the low-noise low-power design techniques used in this OTA.

4.2 Low-Power Low-Noise OTA Design for Gain Stage

The schematic of the low-noise OTA is shown in Fig. 4-3. It is a modified version of a standard folded-cascode topology shown in Fig. 4-5. The OTA in Fig. 4-3 is biased such that the currents of the transistors in the folded branch M_7 - M_{12} are only a small

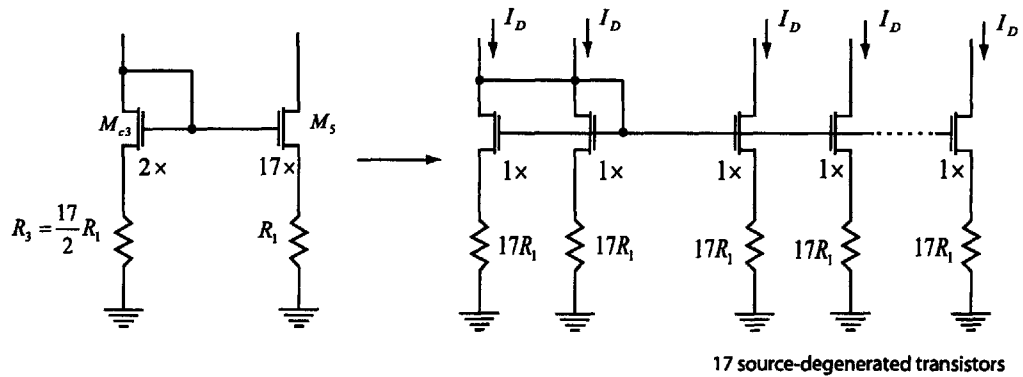


Figure 4-4: Circuit schematic for analyzing current scaling in the source-degenerated current mirrors of Fig. 4-3.

fraction of the current in the input differential pair transistors M_1 and M_2 . In our design, the channel current in M_7 - M_{12} is scaled to approximately $1/16^{th}$ of the current in M_1 and M_2 . The much lower current in M_7 - M_{12} makes the noise contributed by them negligible compared to that from M_1 and M_2 . As a result, we simultaneously lower the total current and the total input-referred noise of the OTA.

To ensure that such severe current scaling is achieved, we carefully set the bias currents of M_5 and M_6 through the use of the bias circuit formed by M_{b2} , M_{c2} and M_{c3} . The current sources M_{b1} , M_{b2} are cascoded to improve their output impedances and thereby ensure accurate current scaling. They operate in strong inversion to reduce the effect of threshold voltage variations. The source-degenerated current mirrors formed by M_{c3} , M_5 and M_6 and resistors R_1 and R_2 set the currents in M_5 and M_6 such that the currents in M_7 and M_8 (the difference between the current in M_3 and M_5 and between the current in M_4 and M_6) are a small fraction of the currents in M_1 and M_2 . An analysis of mismatches in source-degenerated current mirrors is deferred until Section 4.2.3 and is important for robust biasing performance. In order to save power in the bias circuit, the current scaling ratio between M_{b1} and M_{b2} is 16:1 ($2I_B/32$) as shown in Fig. 4-3. To set the currents in the folded-branch transistors to be $I_B/32$, which is $1/16^{th}$ of the currents in differential-pair transistors, we set the current in M_5 and M_6 to be $17I_B/32$. Such current ratioing is achieved by making R_3 to be $17R_1/2 = 17R_2/2$, and constructing M_{c3} as a parallel combination

of two unit transistors while M_5 and M_6 are each constructed from 17 unit transistors in parallel. To clarify this scaling further, the current mirror formed by M_{c3} , R_3 and M_5 , R_1 in Fig. 4-3 is transformed into an equivalent circuit comprised of many source-degenerated unit transistors as shown in Fig. 4-4. All source-degenerated unit transistors are identical and have the same gate voltage. For any gate voltage there is only one source voltage at which a unit resistor's current equals a unit transistor's current. Thus, the nominal channel currents in all unit transistors are identical and the total current in M_5 is $17/2$ times the current in M_{c3} as desired.

For the amplifier to have low input-referred noise, the transconductance G_m of the OTA needs to be maximal for a given current level. For the standard folded-cascode OTA shown in Fig. 4-5, the impedance looking into the sources of M_5 and M_6 is much smaller than the impedance looking into the drains of M_1 - M_4 . As a result, the standard folded-cascode OTA achieves an overall transconductance G_m near g_{m1} , the g_m of M_1 . However, if we lower the current in M_5 - M_{10} to be a small fraction of the current in M_1 and M_2 , the impedance looking into the sources of M_5 and M_6 can be a significant portion of the impedance looking into the drains of M_1 - M_4 such that incremental currents do not almost all go through the sources of M_5 and M_6 in the current divider formed between the sources of M_5 and the drains of M_1 and M_3 . Therefore, G_m is significantly less than g_{m1} . Section 4.2.1 explains how we achieve G_m near g_{m1} even with our extreme current scaling via the use of source-degenerated transistors M_5 and M_6 in Fig. 4-3.

In the standard folded-cascode topology shown in Fig 4-5, the current sources formed by M_3 and M_4 contribute a significant amount of noise due to their large channel currents. In this design, we replace the current-source transistors M_3 and M_4 in Fig. 4-5 with source-degenerated current sources formed by M_5 and M_6 and degeneration resistors R_1 and R_2 as shown in Fig. 4-3. With an appropriate choice of degeneration resistance, the noise contributions from the source-degenerated current sources are mainly from the resistors and can be made much smaller than the noise contributions from MOS transistors operating at the same current level. Another benefit of using source-degenerated current sources is that the noise from resistors

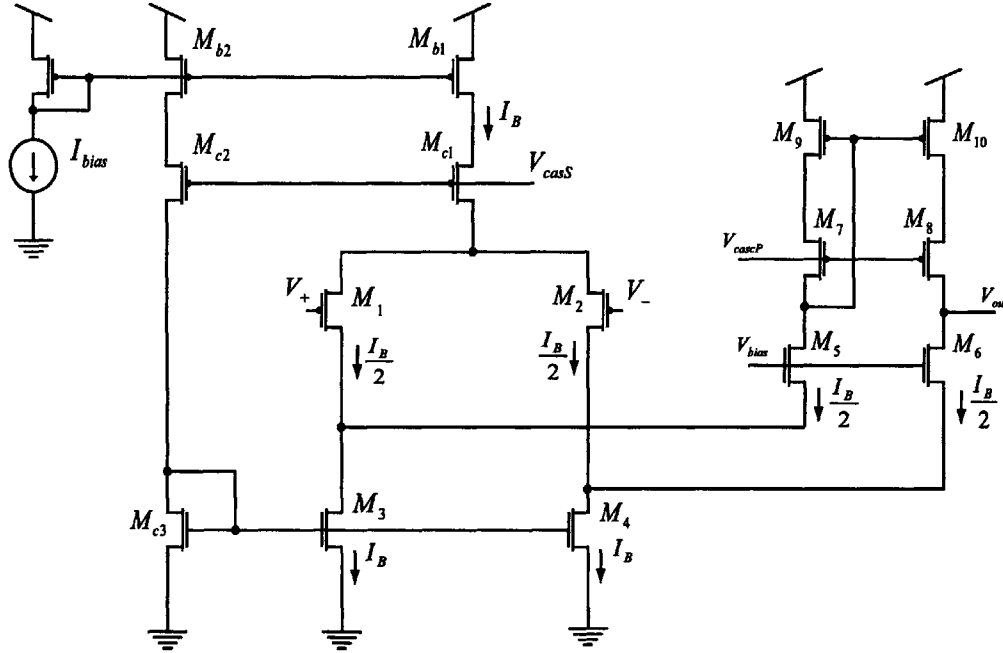


Figure 4-5: Schematic of a standard folded-cascode OTA.

is mainly thermal noise while NMOS current sources contribute a large amount of $1/f$ noise unless they are made with very large area. As a result, the $1/f$ noise in our neural amplifier is mainly from the input differential pair. Therefore, the input-differential pair is made with large-area PMOS transistors, which have lower $1/f$ noise than similarly-sized NMOS transistors in most CMOS processes.

4.2.1 Device Sizing for Maximizing G_m

To achieve low input-referred noise, it is important that the transconductance of the OTA be maximized for a given total current. The maximum transconductance of the standard folded-cascode OTA that can be achieved is the transconductance of one of the transistors in the input-differential pair, say g_{m1} . As a result, it is advantageous to operate M_1 and M_2 in the subthreshold regime where a transistor's g_m is maximized for a given current level. Therefore, M_1 and M_2 need to have large W/L ratios. The lengths of M_1 and M_2 then need to be small such that their widths stay relatively small and the input capacitance of the amplifier is not too large.

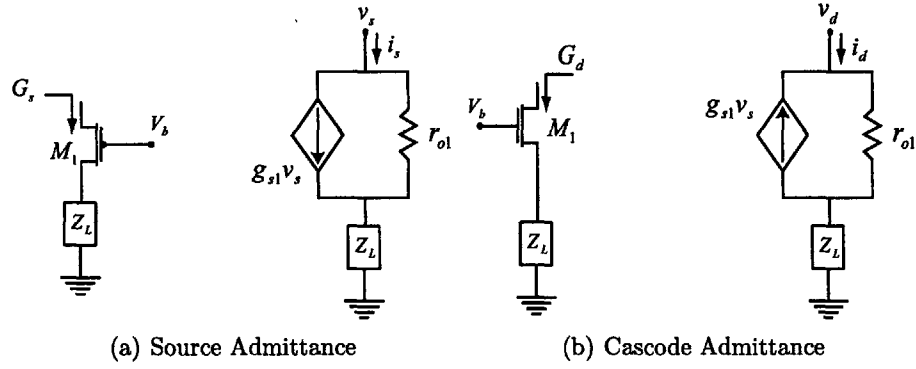


Figure 4-6: Circuit schematics for obtaining admittance formula.

In order to make sure that all the incremental current caused by the differential input goes through the sources of M_7 and M_8 , we cascode the input differential-pair transistors with M_3 and M_4 to increase their output impedances. The source-degenerated current sources formed by M_5 and R_1 and by M_6 and R_2 are designed to have large output impedances as well. The output impedances of the cascoded input-differential pair and the source-degenerated current sources need to be much larger than the impedance looking into the sources of M_7 and M_8 such that G_m is near g_{m1} .

Before we analyze the operation of the OTA in Fig. 4-3, we shall briefly review two useful admittance formulas. The first one is the formula for the admittance looking into the source of an MOS transistor when its drain is connected to an impedance to incremental ground as shown in Fig. 4-6(a). The second useful formula is the admittance looking into the drain of a cascode transistor as shown in Fig. 4-6(b). Using a nodal analysis, we obtain the two admittances to be

$$G_s = \frac{i_s}{v_s} = \frac{g_{s1} + 1/r_{o1}}{1 + Z_L/r_{o1}}, \quad (4.3)$$

$$G_d = \frac{i_d}{v_d} = \frac{1}{r_{o1}} \cdot \left(\frac{1}{1 + g_{s1}Z_L + Z_L/r_{o1}} \right) \quad (4.4)$$

Let G_{s3} be the admittance looking into the sources of M_3 and M_4 , G_{d5} be the admittance looking into the drains of M_5 and M_6 , and G_{s7} be the admittance looking into

the sources of M_7 and M_8 of the OTA in Fig. 4-3. We can express the transconductance G_m of the OTA as

$$G_m = g_{m1} \cdot \left(\frac{G_{s7}}{G_{s7} + G_{d5}} \right) \left(\frac{G_{s3}r_{o1}}{1 + G_{s3}r_{o1}} \right). \quad (4.5)$$

We can express G_{s3} , G_{s7} and G_{d5} by using (4.3) and (4.4) as

$$G_{s3} = \frac{g_{s3} + 1/r_{o3}}{1 + 1/(r_{o3}(G_{s7} + G_{d5}))} \quad (4.6)$$

$$\approx \frac{g_{s3}}{1 + 1/(r_{o3}(G_{s7} + G_{d5}))}, \quad (4.7)$$

$$G_{s7} = \frac{g_{s7} + 1/r_{o7}}{1 + (1/g_{m11})/r_{o7}} \quad (4.8)$$

$$\approx \left(\frac{g_{m11}r_{o7}}{1 + g_{m11}r_{o7}} \right) \cdot g_{s7}, \quad (4.9)$$

and

$$G_{d5} = \frac{1}{r_{o5}} \frac{1}{1 + R_1/r_{o5} + g_{s5}R_1} \quad (4.10)$$

where g_{si} and r_{oi} are the incremental source admittance of M_i with its drain at incremental ground, and the output resistance of M_i respectively. The expressions from (4.7)-(4.10) present the design constraints for sizing and biasing each device to achieve G_m close to g_{m1} . The size, the channel current and the simulated intrinsic gain ($g_s r_o$) of each transistor in the OTA are shown in Table 4.1. From (4.5), in order to make G_m close to g_{m1} , the ratios $G_{s7}/(G_{s7} + G_{d5})$ and $G_{s3}r_{o1}/(1 + G_{s3}r_{o1})$ should be made as close to 1 as possible. The ratio $G_{s7}/(G_{s7} + G_{d5})$ represents the incremental current gain from the drain of M_3 and M_4 to the output. The incremental current gain from the input differential pair transistors to the drain of the cascode transistors M_3 and M_4 is $G_{s3}r_{o1}/(1 + G_{s3}r_{o1})$.

In order to maximize the ratio $G_{s7}/(G_{s7} + G_{d5})$, we try to make $G_{d5} \ll G_{s7}$. Since M_{11} and M_7 have the same channel current, $g_{m11} \approx g_{m7}$. Therefore, $g_{m11}r_{o7} \approx g_{m7}r_{o7} \gg 1$ and we have $G_{s7} \approx g_{s7}$. In order to make $G_{d5} \ll G_{s7}$, we need to

minimize G_{d5} . From (4.10), we can minimize G_{d5} by making r_{o5} large and also making $g_{s5}R_1 \gg 1$. Therefore, we make M_5 and M_6 with large W/L ratios and with long channel lengths to achieve large g_{s5} and r_{o5} respectively. Then we choose R_1 such that $g_{s5}R_1 \gg 1$.

In order to maximize the ratio $G_{s3}r_{o1}/(1 + G_{s3}r_{o1})$, we need to make $G_{s3}r_{o1} \gg 1$. From (4.7), G_{s3} is approximately g_{s3} if $G_{s7}r_{o3}$ is much greater than 1. Since $G_{s7} \approx g_{s7}$, we have $G_{s7}r_{o3} \approx g_{s7}r_{o3}$. Since the current in M_7 is about 1/16 of the current in M_3 and both transistors are operating in subthreshold, $g_{s7} \approx g_{s3}/16$. From simulation, we achieve $g_{s3}r_{o3}$ of 119 which results in a $g_{s7}r_{o3}$ of 7.43. The expression in (4.7) is thus reduced to $G_{s3} \approx 0.88g_{s3}$. Note that M_1 and M_3 have the same currents and the same channel lengths. Thus M_1 and M_3 should have $r_{o1} = r_{o3}$. As a result, $G_{s3}r_{o1} \approx G_{s3}r_{o3} \approx (0.88g_{s3})r_{o3} = 104$. Therefore, the ratio $G_{s3}r_{o1}/(1 + G_{s3}r_{o1})$ is close to 1. As a result, we are able to achieve G_m close to g_{m1} even with sixteen-fold current scaling between the input differential-pair transistors and the folded-branch transistors.

4.2.2 OTA Noise Analysis

The noise in cascode transistors typically contributes little to the overall noise in an OTA [11] because these transistors self shunt their own current noise sources: A cascode transistor's current noise is attenuated by a factor of $1/(1 + g_s R)^2$ where g_s is its incremental source transconductance and R is the effective source-degeneration resistance respectively. Therefore, the only noise sources that are significant in Fig. 4-3 are due to non-cascode transistors, i.e., the differential-pair input transistors M_1

Table 4.1: Operating Points for Transistors in the OTA with $I_{tot} = 2.7 \mu A$

Devices	W/L (μm)	I_D	$g_s r_o$	Operating Region
M_1, M_2	399/1.2	$1.18 \mu A$	133	subthreshold
M_3, M_4	100.5/1.2	$1.18 \mu A$	119	subthreshold
M_5, M_6	204/6	$1.25 \mu A$	322	subthreshold
M_7, M_8	3.6/1.5	$68 nA$	164	subthreshold
M_9, M_{10}	6/1.2	$68 nA$	123	subthreshold
M_{11}, M_{12}	3.6/2.2	$68 nA$	458	above-threshold

and M_2 , the resistors R_1 and R_2 , and the current-mirror transistors M_{11} and M_{12} . We now perform an OTA noise analysis using a method similar to that described in [11].

The admittances looking into the sources of M_3 , M_5 , and M_7 are approximately g_{s3} , g_{s5} , and g_{s7} respectively. Then the current transfer function from each significant current noise source in the OTA to an incrementally grounded output can be calculated to be

$$\frac{\overline{i_{n,out}^2}}{\overline{i_{n,M1}^2}} = \left(\frac{G_{s3}r_{o1}}{1 + G_{s3}r_{o1}} \cdot \frac{G_{s7}}{G_{s7} + G_{d5}} \right)^2 \quad (4.11)$$

$$\approx \left(\frac{g_{s3}r_{o1}}{1 + g_{s3}r_{o1}} \cdot \frac{g_{s7}}{g_{s7} + G_{d5}} \right)^2, \quad (4.12)$$

$$\frac{\overline{i_{n,out}^2}}{\overline{i_{n,R1}^2}} = \left(\frac{G_{s5}R_1}{1 + G_{s5}R_1} \cdot \frac{G_{s7}}{G_{s7} + G_{d3}} \right)^2 \quad (4.13)$$

$$\approx \left(\frac{g_{s5}R_1}{1 + g_{s5}R_1} \cdot \frac{g_{s7}}{g_{s7} + G_{d3}} \right)^2, \quad (4.14)$$

and

$$\frac{\overline{i_{n,out}^2}}{\overline{i_{n,M11}^2}} = 1. \quad (4.15)$$

Since this circuit is biased such that $g_{s3}r_{o1} \gg 1$, $g_{s5}R_1 \gg 1$ and $g_{s7} \gg G_{d5}, G_{d3}$ as explained in Section 4.2.1, the expressions from (4.12)-(4.15) are reduced to 1. For the following discussion, we model the MOSFET's current noise as

$$\overline{i_n^2} = 4\gamma kT g_m \quad (4.16)$$

where k is Boltzmann's constant, T is the absolute temperature, g_m is the transconductance of the MOSFET, and $\gamma = 2/3$ for above-threshold operation and $\gamma = 1/(2\kappa)$ for subthreshold operation. From this noise model, we can calculate the input-referred noise of the OTA as the total output current noise divided by its transconductance g_{m1}^2 to be

$$\overline{v_n^2} = \frac{1}{g_{m1}^2} \left(\frac{4kT g_{m1}}{\kappa} + \frac{8kT}{R_1} + \frac{16}{3} kT g_{m11} \right) \quad (4.17)$$

where M_1 and M_2 operate in weak inversion and M_{11} and M_{12} operate in strong inversion. Let IC be the inversion coefficient of the transistor which is defined as the ratio of its channel current I_D to the moderate inversion characteristic current I_S where I_S is given by [14]

$$I_S = \frac{2\mu C_{ox} U_T^2}{\kappa} \cdot \frac{W}{L} \quad (4.18)$$

where U_T is the thermal voltage and is equal to kT/q , where q is the electron charge. Using the EKV model [3], we can estimate the g_m of each transistor to be

$$g_m = \frac{\kappa I_D}{U_T} \cdot \frac{2}{1 + \sqrt{1 + 4 \cdot IC}} \quad (4.19)$$

We can then rewrite (4.17) as

$$\overline{v_n^2} = \frac{1}{g_{m1}} \cdot \frac{4kT}{\kappa} \left(1 + \frac{2U_T}{I_1 R_1} + \frac{4}{3} \kappa \alpha \frac{I_{11}}{I_1} \right) \quad (4.20)$$

where $\alpha = 2 / (1 + \sqrt{1 + 4 \cdot IC_{11}})$, which is less than 1, and IC_{11} is the inversion coefficient of M_{11} and M_{12} . Equation (4.20) suggests that in order to minimize the input-referred noise of the OTA, $I_1 R_1$ should be large compared to $2U_T$. Furthermore, the current ratio I_1/I_{11} should be large compared to $\frac{4}{3}\kappa\alpha$. For our implementation, the ratio I_1/I_{11} is 16. For a total supply current of $2.7 \mu\text{A}$ and 5.3 kHz bandwidth, I_1 and I_{11} are approximately $1.18 \mu\text{A}$ and 68 nA respectively. For $R_1 = 240\text{k}\Omega$, the second and the third terms in (4.20) are 1.8×10^{-1} and 5.4×10^{-2} respectively, assuming a temperature of $T=300 \text{ K}$, $\kappa = 0.7$ and $\alpha = 1$. Equivalently, (4.17) is reduced to

$$\overline{v_n^2} = \frac{2kT}{\kappa g_{m1}} \times 2.47. \quad (4.21)$$

Equation (4.21) can be interpreted as 2.47 times the input-referred noise of a MOS transistor operating in weak inversion with a transconductance of g_{m1} . This means that our OTA effectively has only 2.47 subthreshold devices that contribute noise. This value is close to the theoretical limit of 2 noise sources in any OTA that uses two subthreshold MOS differential-pair transistors as an input stage. Effectively, our design has minimized almost all other sources of noise except for that of M_1 and M_2 .

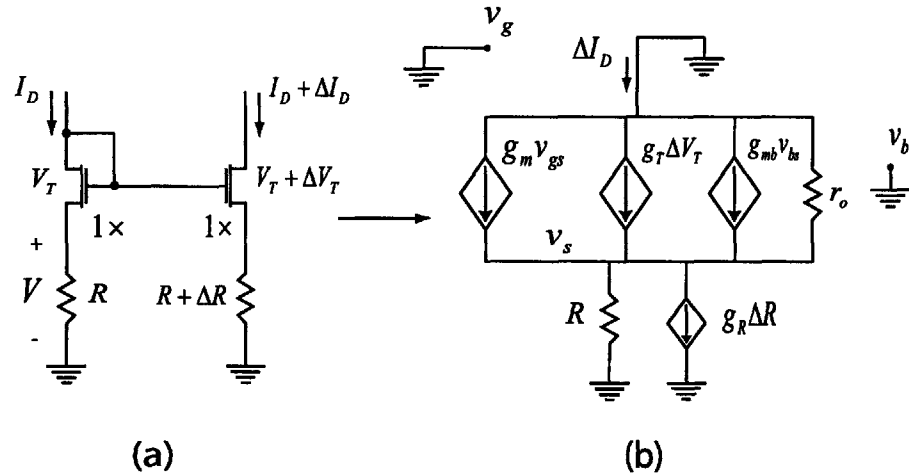


Figure 4-7: Circuit schematics for analyzing V_T and R mismatches in source-degenerated current mirrors.

4.2.3 Current Mirror Mismatch Analysis

The key techniques for achieving good power-noise tradeoff in this amplifier are the uses of source-degenerated current mirrors and the severe current scaling ratio between the input-differential pair transistors and the folded-branch transistors. The severe current scaling scheme can work only if the current errors due to mirroring are well controlled: The amplifier would not work if the error due to current scaling is too large such that none of the current flows in M_7 - M_{12} in the OTA of Fig.4-3. Thus, we address and investigate this concern to ensure the correct operation of our amplifier. Let us consider the current matching between two unit transistors in Fig. 4-4 due to variations in the threshold voltage and variations in the source-degeneration resistance. We shall model these variations as errors in the parameters of each of the unit transistors of Fig. 4-4. Let the nominal current in one of the unit transistors of M_{c3} be I_D and consider the deviation in current ΔI_D in one of the unit transistors of M_5 from its nominal value due to deviations in the threshold voltage ΔV_T and deviations in the source-degeneration resistor ΔR as shown in Fig. 4-7(a). To model the threshold-voltage mismatch, we use the body-referenced current equation in saturation for an MOS transistor operating in weak inversion [14]. Let the nominal

current in each unit transistor be described by

$$I_D = I_s e^{\kappa(V_{GS}-V_T)/U_T} \cdot e^{(1-\kappa)V_{BS}/U_T}. \quad (4.22)$$

where I_s is a constant scaling current which is the same for all unit transistors. Let V be the nominal DC voltage drop across R such that $I_D = V/R$. We define

$$g_T = \frac{\partial I_D}{\partial V_T} = -\frac{\kappa}{U_T} \cdot I_D = -g_m, \quad (4.23)$$

$$g_R = \frac{\partial I_D}{\partial R} = \frac{\partial}{\partial R} \left(\frac{V}{R} \right) = -\frac{1}{R} \cdot \frac{V}{R} = -\frac{1}{R} \cdot I_D. \quad (4.24)$$

and

$$g_{mb} = \frac{\partial I_D}{\partial V_{BS}} = \frac{1-\kappa}{U_T} \cdot I_D. \quad (4.25)$$

By assuming that ΔV_T and ΔR are small, we can use a small-signal circuit model as shown in Fig. 4-7b to calculate the variation in nominal current ΔI_D when ΔV_T and ΔR are considered as inputs to the system. With some analysis, the variation in the channel current due to variations in V_T and R is obtained to be

$$\Delta I_D = g_T \cdot \Delta V_T - (g_m + g_{mb} + 1/r_o) \cdot (\Delta I_D - g_R \cdot \Delta R) \cdot R. \quad (4.26)$$

Combining (4.26) with the results from (4.23) and (4.24) and using the relationship $g_s = g_m + g_{mb}$, we obtain the fractional change in channel current as a function of the fractional change in V_T and R to be

$$\frac{\Delta I_D}{I_D} = -\frac{1}{1 + g_s R + R/r_o} \cdot \frac{\Delta V_T}{I_D/g_m} - \frac{g_s R + R/r_o}{1 + g_s R + R/r_o} \cdot \frac{\Delta R}{R}. \quad (4.27)$$

Since M_{c3} , M_5 and M_6 are biased in weak-inversion regime, their I_D/g_m is approximately 40 mV at room temperature. As seen from (4.27), the mismatch in threshold voltage as a fraction of 40 mV is attenuated by a factor of $1 + g_s R + R/r_o$ and is negligible if $g_s R \gg 1$. In our design, we have $g_s R \approx 12$, thus, the fractional mismatch in threshold voltage is attenuated by more than a factor of 10 and does not play

a significant role in current mirror mismatch. In contrast, the fractional mismatch in channel current scales almost 1:1 to the fractional mismatch in R . However, the matching of passive components in most CMOS processes is much better controlled than the matching of transistors' threshold voltages. In our design, therefore, we try to achieve good resistor matching with careful layout.

4.2.4 Noise Efficiency Factor and Its Theoretical Limit for Any OTA

To compare the power-noise tradeoff among amplifiers, we adopt the noise efficiency factor (NEF) proposed in [13] and widely used to compare neural-amplifier designs:

$$\text{NEF} = V_{ni,rms} \sqrt{\frac{2I_{tot}}{\pi \cdot U_T \cdot 4kT \cdot BW}} \quad (4.28)$$

where $V_{ni,rms}$ is the total input-referred noise, I_{tot} is the total supply current, and BW is the -3 dB bandwidth of the amplifier respectively. The theoretical limit of the NEF of an OTA that uses a differential pair as an input stage is when the two differential-pair transistors are the only noise sources in the circuit. The input-referred noise of the OTA is then $\overline{V_{ni}^2} = 2 \times 2kT / (\kappa g_m) = 4kT / (\kappa g_m)$ where g_m is the transconductance of a single differential-pair transistor. For minimum input-referred noise, the transistors should run in subthreshold, such that we have $g_m = \kappa I_D / U_T$. Assuming a first-order roll-off of the frequency response, the input-referred noise of the ideal OTA is expressed as

$$V_{ni,rms} = \sqrt{\frac{4kT \cdot U_T}{\kappa^2 I_D} \cdot \frac{\pi}{2} \cdot BW}. \quad (4.29)$$

Combining (4.28) and (4.29) and setting $I_{tot} = 2I_D$, we obtain the theoretical limit for NEF of any OTA that uses a subthreshold MOS differential pair to be

$$\text{NEF} = \frac{\sqrt{2}}{\kappa} \approx 2.02 \quad (4.30)$$

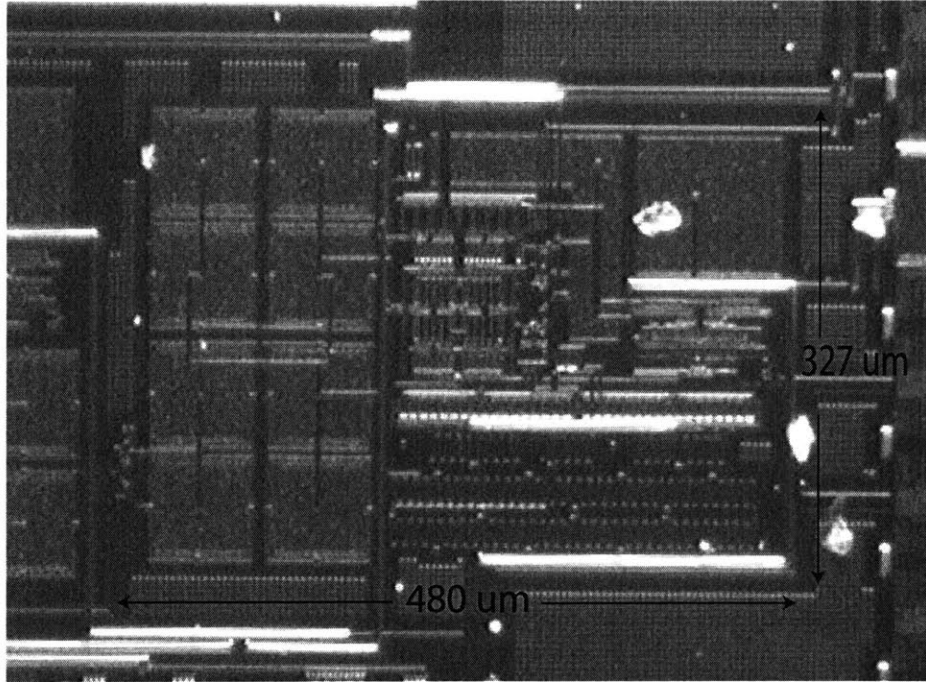


Figure 4-8: A die micrograph of our neural amplifier.

assuming a typical value of $\kappa = 0.7$. We now show that our experimental NEF is near this value, and our theoretical NEF was computed to be 2.47 from Section 4.2.2.

4.3 Measurement Results

The amplifier was fabricated in a $0.5 \mu\text{m}$ CMOS process through the AMI foundry. It was designed to give a gain of approximately 110 (40.8 dB) by setting the value of C_{in} to 14 pF and C_f to 120 fF. The OTA in the bandpass filter stage is a wide common-mode range OTA to reduce signal distortion in the case of large input amplitudes. The amplifier occupies a chip area of 0.16 mm^2 . A chip micrograph of our amplifier is shown in Fig. 4-8.

Four chips were tested on the lab bench and they exhibited very similar performance characteristics, indicating that the severe current-scaling scheme worked robustly. The measured transfer function of one of our neural amplifiers is shown in Fig. 4-9. The amplifier consumes $2.7 \mu\text{A}$ including the current from the bias circuit

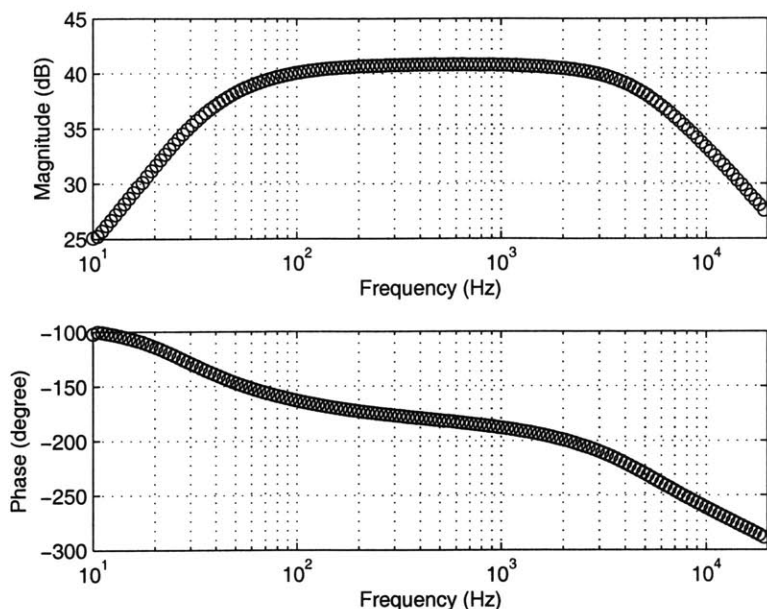


Figure 4-9: Measured transfer function of the neural amplifier configured for recording neural spikes.

(M_{b2} , M_{c2} and M_{c3}) from a 2.8 V supply. We do not include the current I_{bias} shown in Fig. 4-3 since it can be shared by many amplifiers in the array. The -3 dB cutoff frequencies are adjusted to be at 45 Hz and 5.32 kHz. The amplifier is configured as an inverting amplifier, thus the phase is approximately -180° near the midband frequency.

Fig. 4-10 shows the measured input-referred noise spectrum together with a circuit simulation of the noise spectrum with a similar noise model to the theoretical calculations (the smooth curve). There is a good agreement between the measured and simulated curves. The measured thermal noise level is $31 \text{ nV}/\sqrt{\text{Hz}}$. Integrating under the area of the measured curve from 10 Hz to 98 kHz yields a total input-referred noise of $3.06 \mu\text{V}_{rms}$, while the simulated result is $3.1 \mu\text{V}_{rms}$. With a high-pass cutoff frequency at 45 Hz, $1/f$ noise is filtered out and is not noticeable in the passband.

The NEF of this amplifier is calculated from the achieved experimental measurements to be 2.67. This value is close to 2.02 which is the theoretical NEF limit that has been calculated in 4.2.4 and also near our expected theoretical calculation of 2.47

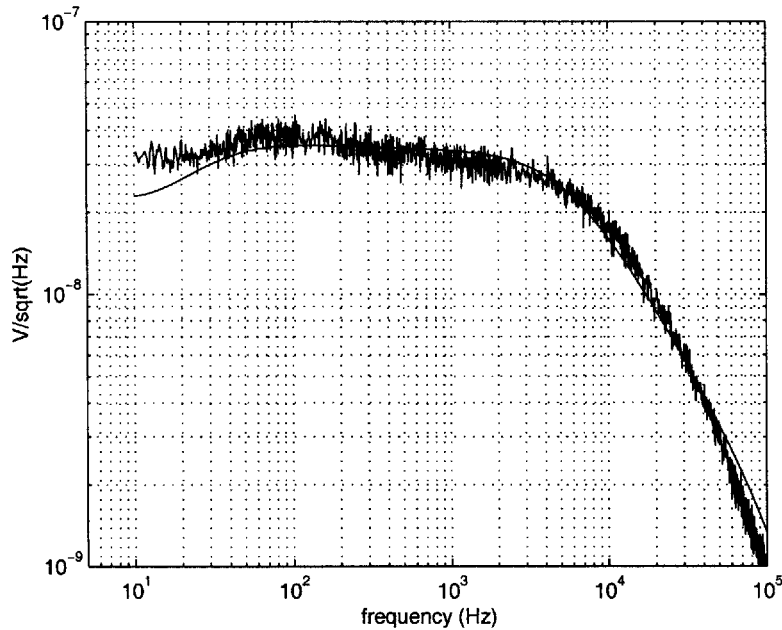


Figure 4-10: Measured and simulated (smooth curve) input-referred noise spectra of the neural amplifier configured for recording neural spikes.

in (4.21). The good power-noise tradeoff of this amplifier is a result of minimizing the effective number of transistors that contribute noise. Moreover, almost all the power is consumed by the input-differential pair. Therefore, little power is wasted in less critical parts of the amplifier. Fig. 4-11 compares NEF of previously reported neural amplifiers as a function of the total supply current. Our amplifier exhibits the best NEF and lowest power consumption reported to date.

The measured CMRR and PSRR are shown in Fig. 4-12. CMRR is calculated as the ratio of the differential-mode gain to the common-mode gain. PSRR is calculated as the ratio of the differential-mode gain to the gain from power supply to the output. The measured CMRR and PSRR exceed 66 dB and 75 dB (over the range of 45 Hz to 5.32 kHz) respectively. The measured characteristics of the neural amplifier are summarized in Table 4.2.

We verified that this neural amplifier works in a real recording environment by using it to record action potentials in the RA region of a zebra finch's brain. Data were taken with a Carbostar electrode that had an impedance of approximately 800 k Ω .

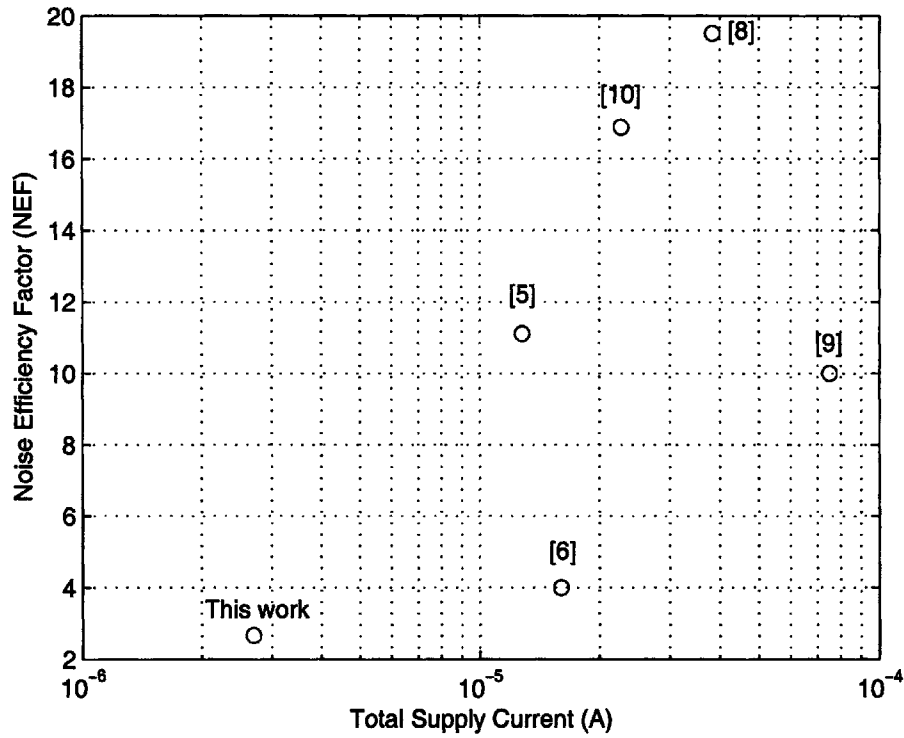
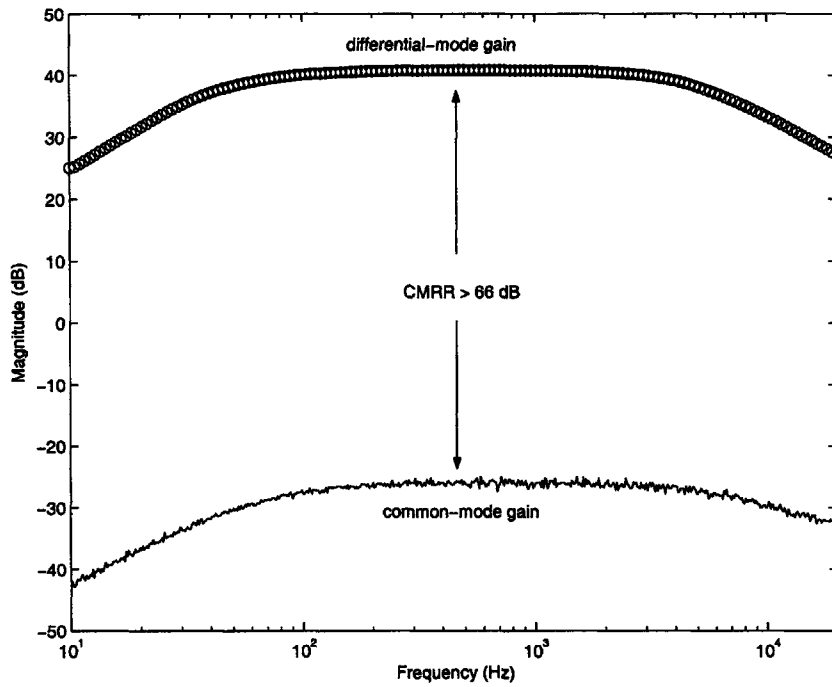


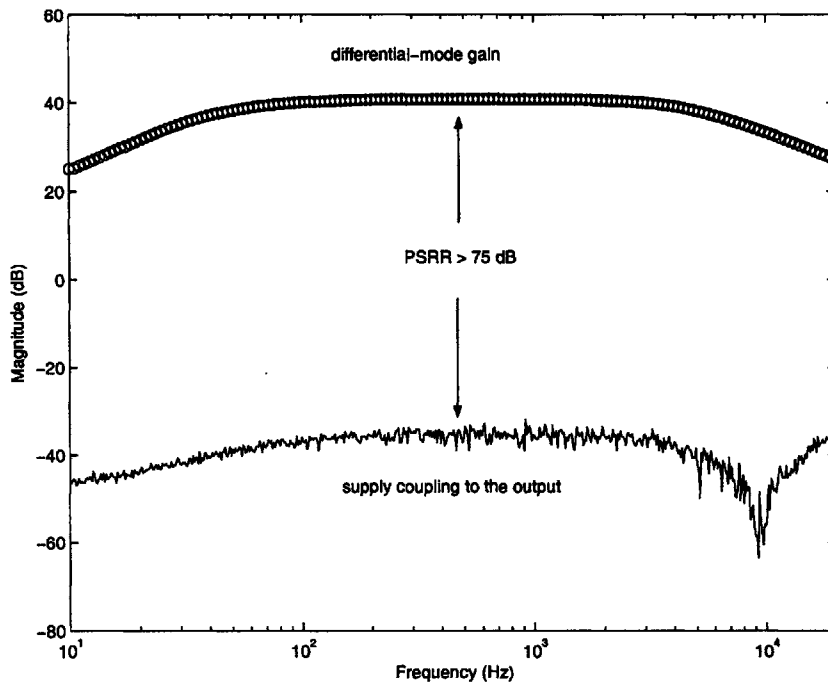
Figure 4-11: NEF of published neural amplifiers as a function of the total supply current.

Table 4.2: Measured Performance Characteristics

Parameter	Measured
Supply voltage	2.8 V
Total current	2.7 μ A
Gain	40.85 dB
Bandwidth	45 Hz-5.32 kHz
Input-referred noise	3.06 μ V _{rms}
Noise efficiency factor	2.67
Max. signal (1% THD @ 1.024 kHz)	7.3 mV _{pp}
Dynamic Range (1% THD)	58 dB
CMRR (45 Hz-5.32 kHz)	66 dB
PSRR (45 Hz-5.32 kHz)	75 dB
Area (in 0.5 μ m CMOS)	0.16 mm ²



(a) CMRR



(b) PSRR

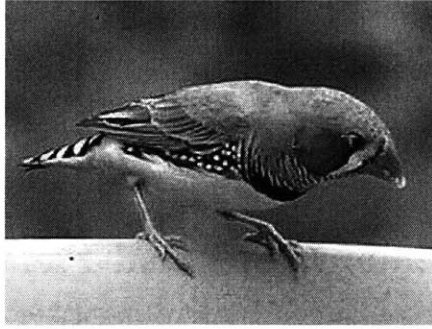
Figure 4-12: CMRR and PSRR measurements of the neural amplifier configured for recording action potentials.

A long extracellular trace and a short extracellular trace recorded from our amplifier normalized by the gain are shown in Fig. 4-13. They were found to be identical to that recorded by a commercial neural amplifier.

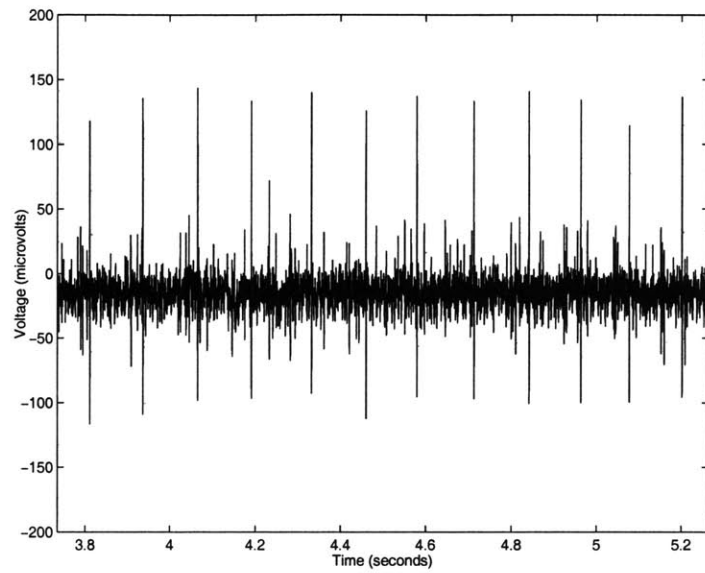
4.4 Measurements of Local Field Potentials

Local Field Potentials (LFPs) instead of action potentials are often used in some brain-machine interfaces, e.g, those used in paralysis prosthetics [12]. Therefore, we also measured the performance characteristics of our amplifier configured with lower bandwidth (and power) for such applications. Since the LFP contains energy in the frequency range of 1 Hz to 100 Hz, we can simply lower the -3 dB lowpass cutoff frequency of our amplifier by lowering the supply current of the OTA in the bandpass filter stage. The highpass cutoff frequency can also be lowered to be below 1 Hz by adjusting V_{tune} . If we just change the bandwidth in this manner, the input-referred noise of the amplifier becomes excessively low. From a hand-analysis, if we adjust the bandwidth of the amplifier to be 0.5 Hz-300 Hz while maintaining the same supply current for the gain-stage OTA, the input-referred noise of the amplifier is less than $1 \mu V_{rms}$. Such low input-referred noise is unnecessary and is wasteful of power. From (4.20), the input-referred noise power is inversely proportional to g_{m1} , therefore inversely proportional to I_1 . Thus, we can save more power by lowering the current in the gain-stage OTA as well.

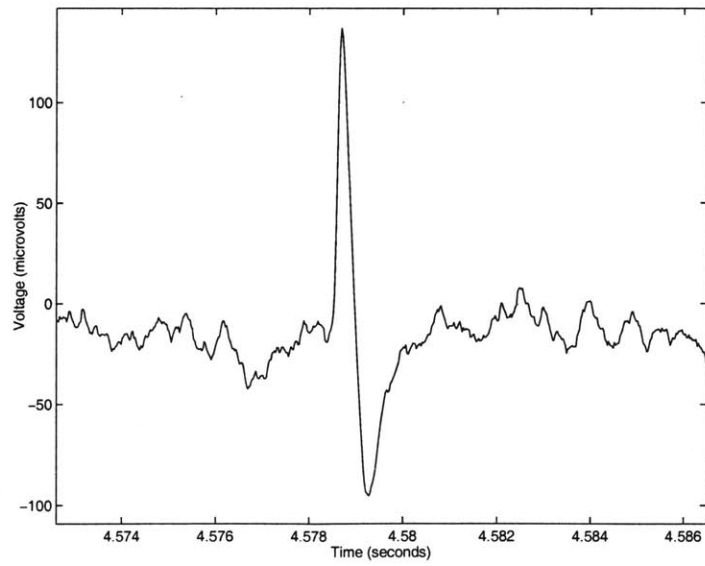
The amplifier was adjusted to have a highpass cutoff frequency of 392 mHz and a lowpass cutoff frequency of 295 Hz for LFP-suitable configuration. The total current of our amplifier was measured to be 743 nA, corresponding to a power consumption of $2.08 \mu W$ from a 2.8 V supply and $1.66 \mu V_{rms}$ total input-referred noise integrated from 0.2 Hz to 1 kHz. The measured transfer function for the amplifier configured for recording LFP is shown in Fig. 4-14. The measured input-referred noise spectrum and expected noise curve from simulation are shown in Fig. 4-15. The measured NEF for LFP recording is then 3.21, still better than any other amplifier to date. Note that the NEF is worse than that of the amplifier configured to record neural spikes.



(a) A zebra finch



(b) Spike train



(c) Single spike

Figure 4-13: Neural recording from a zebra finch's brain: (a) A zebra finch (b) Long time trace (c) Short time trace.

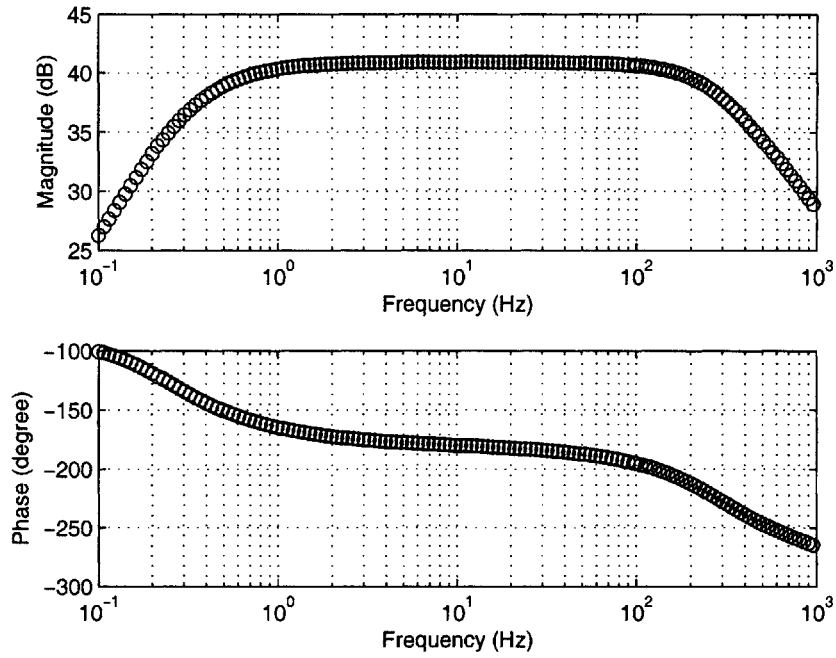


Figure 4-14: Transfer function of the amplifier configured for recording LFP.

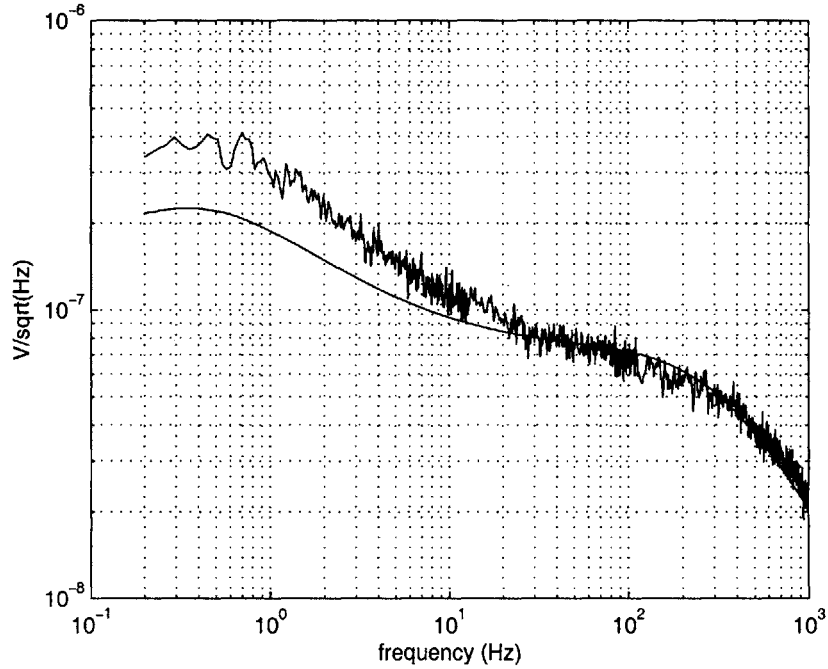


Figure 4-15: Measured and simulated (smooth curve) input-referred noise spectra for the amplifier configured for recording LFP.

This degradation in NEF is due to the fact that the thermal noise from the resistors R_1 and R_2 becomes more significant once the current in the input differential pair is low. Moreover, $1/f$ noise becomes significant as well since the highpass cutoff has been decreased to 395 mHz. The other measured performance characteristics of the LFP amplifier are summarized in Table 4.3 and similar to those shown in Table 4.2.

Table 4.3: Measured Performance Characteristics of LFP Amplifier

Parameter	Measured
Supply voltage	2.8 V
Total current	743 nA
Gain	40.9 dB
Bandwidth	392 mHz-295 Hz
Input-referred noise	1.66 μV_{rms}
Noise efficiency factor	3.21
Max. signal (1% THD @ 1.024 kHz)	7.2 mV _{pp}
Dynamic Range (1% THD)	63.7 dB
CMRR (392 Hz-295 Hz)	66 dB
PSRR (392 Hz-295 kHz)	75 dB
Area (in 0.5 μm CMOS)	0.16 mm ²

4.5 Conclusion

This chapter presented a micropower low-noise neural recording amplifier. Many low-noise design techniques were employed to enable the amplifier to achieve an input-referred noise near the theoretical limit of two devices of an input differential pair. The design was done carefully and all the implementation issues encountered in the amplifiers of previous chapters were solved. The amplifier appears to be the lowest power and most energy-efficient neural amplifier reported to date. It can be configured to record either action potentials or local field potentials. We obtained successful recordings of action potentials with our amplifier from a zebra finch's brain. This amplifier may thus be useful in brain-machine interfaces for paralysis prosthetics, visual prosthetics, or experimental neuroscience systems for chronic monitoring.

Chapter 5

Variable Input-Referred Noise Amplifier and Adaptive-Biasing Technique

As seen from the probability distribution of the background noise in Fig. 1-2 which is collected from a multi-electrode array, there is only a small probability that the background noise will be smaller than $10 \mu V_{rms}$. According to this probability distribution, the probability that the background noise at a particular recording site is lower than $10 \mu V_{rms}$ is approximately 0.05. The amplifier presented in Chapter 4 was designed to have an input-referred noise less than $4 \mu V_{rms}$ which corresponds to the lowest noise level expected in any recording situation according to Fig. 1-2. I have shown in Chapter 1 that the total power consumption of a thermal-noise limited amplifier scales as $1/v_n^2$ where v_n is the total input-referred noise of the amplifier. Even though the amplifier in Chapter 4 achieves the best power-noise efficiency of any neural amplifier ever reported, the total power consumption of such multi-amplifier array will be far from optimum since most of the amplifiers in the array still have an input-referred noise lower than necessary. A further power saving can be achieved if the knowledge of the background noise is known and the power consumption of the amplifier could be adjusted such that its input-referred noise is just low enough for the amplifier to get a clean recording from that particular recording site. In this

chapter, I will discuss an example of a neural recording amplifier whose input-referred noise can be adjusted by varying its bias current. Furthermore, a feedback calibration scheme that is capable of sensing the input-referred noise of an amplifier and adjusting its bias current based on the information of the background noise at the recording site will be presented.

5.1 A Variable Input-Referred Noise Amplifier

The feedback amplifier discussed in Chapter 3 is a good example of an adjustable input-referred noise amplifier. Recall that if the pole due to the loading effects at the output of the gain stage of the amplifier in Fig. 3-1 is at a much higher frequency than the pole due to the bandwidth-limiting stage, the bandwidth of the amplifier is determined by the bandwidth-limiting stage alone (given that the current in the gain stage is not excessively low). Since the input-referred noise per unit bandwidth of the amplifier is proportional to $1/g_{m1}$ where g_{m1} is the g_m of one of the input differential-pair transistors, the total input-referred noise of the amplifier is proportional to $1/\sqrt{g_{m1}} \propto 1/\sqrt{I_{tot}}$ where I_{tot} is the total current of the amplifier. A plot showing the noise spectrum corresponding to three levels of the amplifier's total current is shown in Fig. 5-1. The total input-referred noise values corresponding to the total current levels of $1 \mu\text{A}$, $2 \mu\text{A}$, and $3.5 \mu\text{A}$ are $8.4 \mu\text{V}_{rms}$, $6.4 \mu\text{V}_{rms}$, and $5.5 \mu\text{V}_{rms}$ respectively. It can be seen that these total input-referred noise values are approximately proportional to $1/\sqrt{I_{tot}}$ as expected.

5.2 Feedback Calibration Scheme for Adjusting the Amplifier's Input-Referred Noise

An example of a feedback calibration loop capable of adjusting the input-referred noise of the amplifier is shown in Fig. 5-2. In this scheme, each neural amplifier in the array has two modes operation; the normal operation mode and the input-referred noise calibration mode. During the normal operation mode, the neural amplifier just

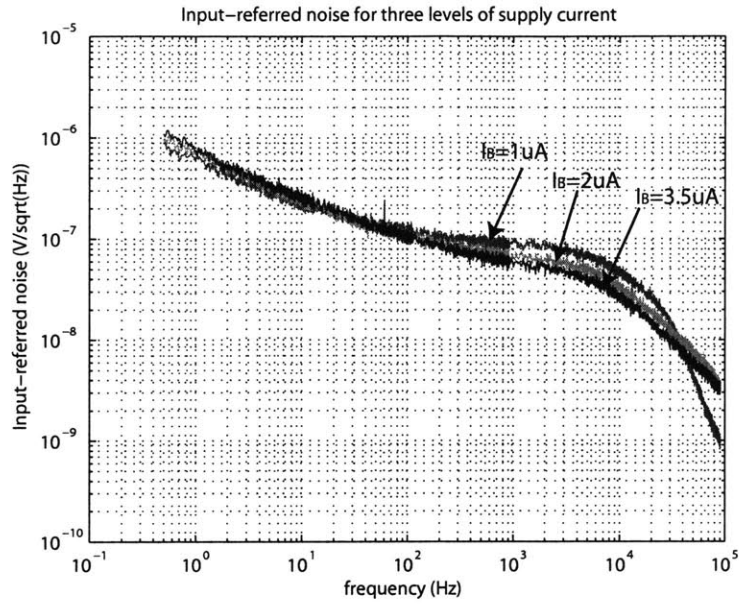


Figure 5-1: Noise power spectrum for three different supply currents.

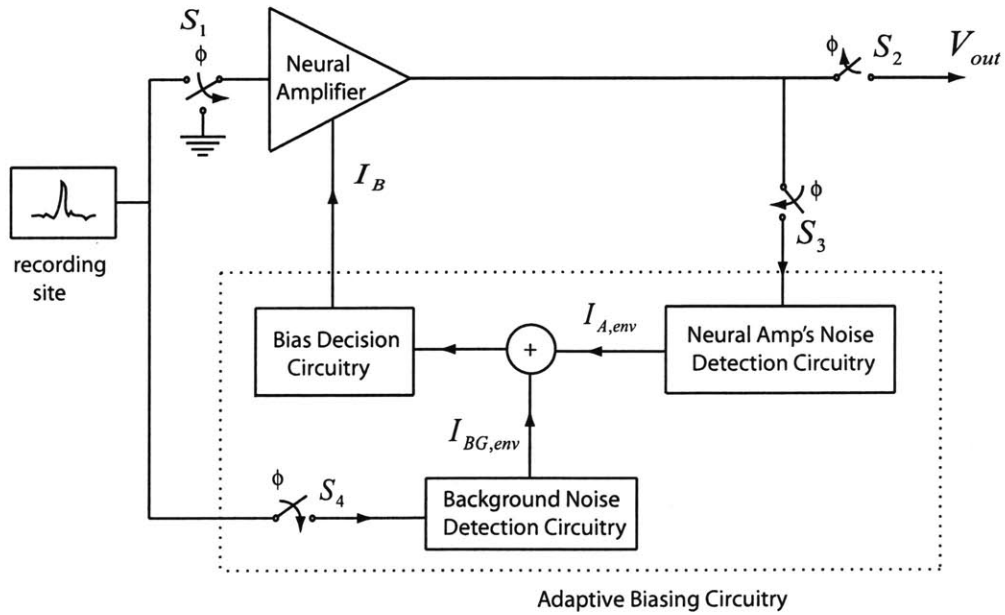


Figure 5-2: Overall architecture of adaptive biasing loop.

operates normally by amplifying the neural signals at the recording site and the amplified signals are obtained at the output V_{out} . During this mode, the switch control signal Φ is not asserted. The switch S_1 connects the input of the amplifier to the recording site and S_2 connects the output of the amplifier to V_{out} while S_3 and S_4 are open, disconnecting the amplifier from the adaptive biasing circuitry. Once the amplifier has operated normally for a period of time, the input-referred noise of the amplifier may need to be calibrated since the background noise at the recording site may have changed. The external control unit (not shown) will configure the feedback calibration loop to be ready for the input-referred noise calibration mode. During the input-referred noise calibration mode, the switch control signal Φ is asserted. The switch S_1 disconnects the amplifier's input from the recording site and then connects it to ground. The switches S_3 and S_4 are closed, connecting the output of the neural amplifier to the adaptive biasing circuitry. The switch S_2 is open, making the output of the amplifier invalid during the input-referred noise calibration.

During input-referred noise calibration mode, the "Background Noise Detection Circuitry" extracts the envelope of the background's noise at the recording site. At the same time, the "Neural Amp's Noise Detection Circuitry" extracts the envelope of the input-referred noise of the amplifier. The magnitudes of these two noise envelopes are compared and the difference is sensed by the "Bias Decision Circuitry". The "Bias Decision Circuitry" can then adjust the bias current of the neural amplifier, and subsequently the input-referred noise of the amplifier based on the difference between the two noise envelopes. The feedback loop stabilizes where the difference in the noise envelopes is negligible. Once the loop is stabilized, the input-referred noise of the amplifier is already close to the background noise at the recording site. After the calibration is finished, the control signal Φ is deasserted and the amplifier resumes its normal operation. Note that this calibration scheme can be applied to other neural amplifiers in a multi-electrode array and that the adaptive biasing circuitry can be shared by all the amplifiers in the array. Therefore, the power overhead per amplifier due to the addition of the adaptive biasing circuitry is negligible if the adaptive biasing circuitry is shared by a large number of neural amplifiers.

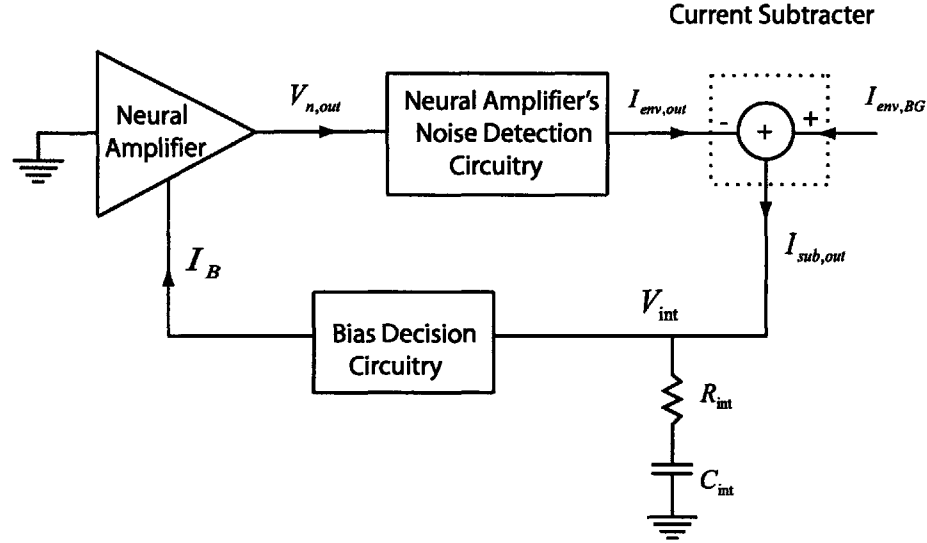


Figure 5-3: Schematic of noise calibration feedback loop implemented in this design.

In this thesis project, I have designed and built a feedback calibration loop for calibrating the input-referred noise of a neural amplifier. The purpose is to study the dynamic of the loop during input-referred noise calibration mode. Therefore, only one neural amplifier is incorporated into the calibration loop. In this chapter, I will explain the system that was built and will analyze its behavior during input-referred noise calibration.

5.3 Input-Referred Noise Calibration Loop's Building Blocks

The schematic of the input-referred noise calibration loop that I designed for this thesis project is shown in Fig. 5-3. The voltage $V_{n,out}$ is the output noise of the amplifier (input-referred noise multiplied by the amplifier's midband gain) when its input is connected to ground. The Neural Amplifier's Noise Detection Circuitry extracts the envelope of $V_{n,out}$ in the form of the current $I_{env,out}$. In this implementation, I have assumed that the background noise at the recording site is extracted in the same manner as that of $V_{n,out}$ and that the background noise envelope is represented by

$I_{env,BG}$. The current subtracter subtracts $I_{env,out}$ from $I_{env,BG}$ to create an error current $I_{sub,out}$. This error current is then integrated by a lead compensation network composed of R_{int} and C_{int} to create a voltage V_{int} . The Bias Decision Circuitry creates the bias current I_B for biasing the neural amplifier based on the value of V_{int} . The feedback calibration loop will servo until V_{int} is at a value such that $I_{env,out}$ is equal to $I_{env,BG}$ and the error current $I_{sub,out}$ is zero.

In order to analyze the dynamic of the loop during input-referred noise calibration, I will explain the implementation of each building block and derive the approximated transfer function for each block.

5.3.1 Neural Amplifier's Noise Detection Circuitry

During input-referred noise calibration mode, the input terminal of the neural amplifier is connected to ground, thus its output signal is the input-referred noise amplified by the gain of the amplifier. We can then extract the magnitude information of this output noise with an envelope detector. Since the amplifier has a gain of approximately 40 dB with the input-referred noise of less than $20 \mu V_{rms}$, the noise at the output of the amplifier has a magnitude of less than $2 mV_{rms}$ which might be too small for an envelope detector. Therefore, another preamplifier is needed before the noise signal is passed to the envelope detector. Since the input-referred noise of the neural amplifier is gained up by more than $100\times$, the input-referred noise requirement of the preamplifier can be relaxed. Therefore, the power consumption in the preamplifier can be much smaller than the power in the neural amplifier. The high-level schematic of the neural amplifier's noise detection circuitry is shown in Fig. 5-4. The preamplifier only needs to provide moderate gain to exceed the dead-zone of the envelope detector. Furthermore, when the bias current of the neural amplifier is changing during the input-referred noise calibration, its DC output voltage also varies at a frequency comparable to the loop bandwidth. Thus, the preamplifier should incorporate a highpass cutoff to reject this DC voltage variation. The topology of the preamplifier used in this design is similar to the folded-cascode design discussed in Chapter 3, except that the weak-inversion PMOS transistors are used to realize the

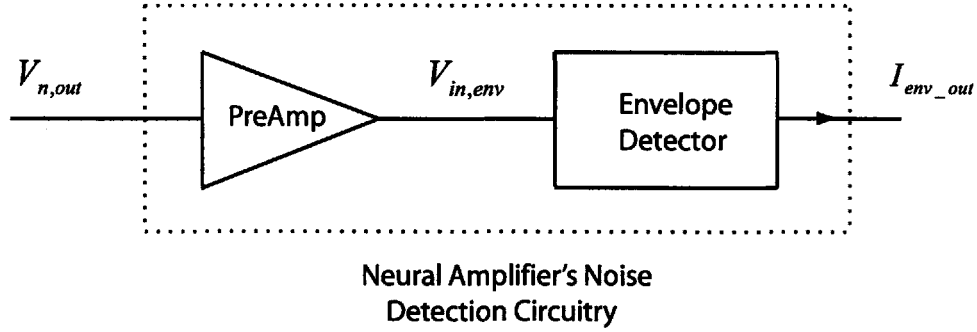


Figure 5-4: High-level schematic of the neural amplifier's noise detection circuitry.

highpass cutoff of the amplifier instead of the MOS-bipolar pseudo-resistor elements. In this way, the highpass cutoff of the preamplifier can be tuned to about 100 Hz to safely reject the neural amplifier's DC output voltage variation during input-referred noise calibration. Note that using the weak-inversion PMOS transistors to realize the highpass cutoff frequency would introduce a low frequency noise that rolls off as $1/f^2$ in power units due to the filtering of the thermal noise in the transistors as discussed in Chapter 2. Nevertheless, due to a gain of 100 of the neural amplifier, this noise would be insignificant compared to the output noise of the neural amplifier.

The envelope detector was realized using the topology similar to that discussed in [16]. The schematic of the envelope detector is shown in Fig. 5-5. It composes of a half-wave rectifier and a current-mode peak detector. The amplifier A is included to reduce the dead zone of the envelope detector. The operation of the half-wave rectifier can be described as follows. When I_{in} is positive, the voltage at the negative terminal of A increases, thus $V_{out,BOT}$ and $V_{out,TOP}$ decrease, quickly shutting down M_n while turning on M_p . During this phase, the NMOS current mirror formed by M_1 and M_2 conducts and sinks the current of $I_{out} = -I_{in}$ from the capacitor C . On the contrary, when I_{in} is negative, the voltage at the negative terminal of A decreases, thus $V_{out,BOT}$ and $V_{out,TOP}$ increase, quickly shutting down M_p while turning on M_n . Therefore, the PMOS current mirror formed by M_3 and M_4 conducts and sources the current of $I_{out} = -I_{in}$ into the capacitor C . Therefore, we have the relationship $I_{out} = -I_{in}$.

The voltage across the capacitor which is also the voltage at the positive terminal

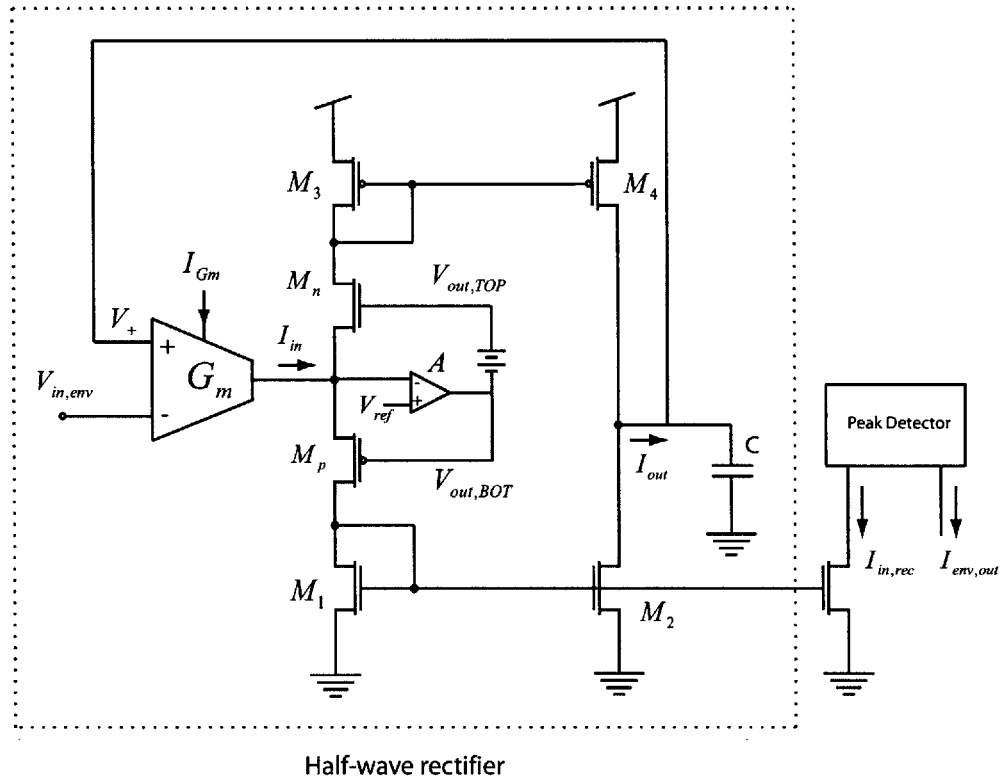


Figure 5-5: A schematic of the envelope detector used in this design.

of the G_m OTA is given by

$$V_+ = \frac{1}{sC} \cdot I_{out}. \quad (5.1)$$

We can then find the transfer function from $V_{in,env}$ to I_{out} as

$$-G_m \cdot (V_{in,env} - V_+) = I_{in} = -I_{out}. \quad (5.2)$$

Combining the results in (5.1) and (5.2), we obtain the transfer function from $V_{in,env}$ to I_{out} to be

$$\frac{I_{out}}{V_{in,env}}(s) = \frac{sC}{1 + sC/G_m}. \quad (5.3)$$

The transfer function in (5.3) exhibits a highpass cutoff at $G_m/(2\pi C)$ Hz. In order for the envelope detector to operate at low-frequency, the capacitor C needs to be large. In this design, I use an off-chip capacitor of 20 nF for the capacitor C . For this value of C and with the bias current of $I_{Gm} = 40$ nA, the highpass cutoff frequency of the envelope detector is below 5 Hz. Half-wave rectification is performed by taking a copy of I_{out} for one polarity. In this implementation, $I_{in,rec}$ is a copy of I_{out} when it is negative (when the NMOS current mirror sinks the current from C). The half-wave rectified current $I_{in,rec}$ is peak detected by the peak detector to obtain the amplitude information of the neural amplifier's input-referred noise.

The current-mode peak detector that is used in this design is shown in Fig 5-6. The detailed operation of the peak detector is described in [16]. The peak detector is a current-mode log-domain low-pass filter with asymmetric attack and release time constants. All transistors have the same size and are biased in subthreshold to achieve logarithmic I-V characteristic for implementation of a log-domain filter. To ensure the stability of the feedback loop formed by M_3 and M_5 , the current I_a is made much larger than I_r . During the attack mode when I_{in} increases, the output current $I_{env,out}$ of the peak detector rises with a time constant of

$$\tau_a = \frac{C_a \cdot U_T}{\kappa \cdot I_a}. \quad (5.4)$$

During the release mode when I_{in} decreases, the output current $I_{env,out}$ of the peak

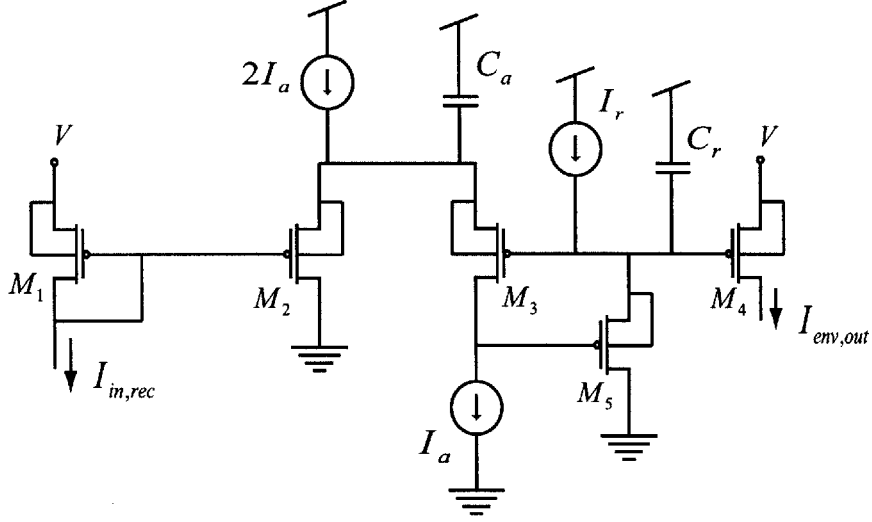


Figure 5-6: A schematic of the current-mode peak detector.

detector decreases with a time constant of

$$\tau_r = \frac{C_r \cdot U_T}{\kappa \cdot I_r}. \quad (5.5)$$

The transfer function of the overall envelope detector can be modeled as a first-order system by

$$H_{env}(s) = \frac{I_{out_{env}}(s)}{V_{in}} = \frac{G_{env}}{1 + \tau_{env}s} \quad (5.6)$$

where $\tau_{env} = \tau_a$ during the attack mode and $\tau_{env} = \tau_r$ during the release mode. Due to the asymmetric attack and release time constants, the input-referred noise calibration loop exhibits different dynamics when the input-referred noise of the neural amplifier is increasing and when it is decreasing. The time constant of the peak detector is very important in determining the dynamic of the calibration loop. Since the peak detector needs to smooth out the rectified noise current which is very spurious, its time constant needs to be slow enough to smooth out these spurious transients. Table 5.1 shows the designed parameters for the peak detector and the corresponding time constants at room temperature.

5.3.2 Current Subtractor

The output of the envelope detector is a DC current that is proportional to the magnitude of the neural amplifier’s input-referred noise. Similar circuitry can be used to extract the magnitude information of the background noise at the recording site. The background noise detection circuitry would need one very low-noise amplifier which has the gain equal to that of the neural amplifier but with much lower input-referred noise. Practically, the input-referred noise of the low-noise amplifier should be well below $4 \mu V_{rms}$, which is the lowest noise level expected at any recording site in the array. The low-noise operation is required in order to correctly sense the background noise at the recording site when the background noise is very low. This low-noise amplifier will be power-hungry, however, the power overhead per each neural amplifier in the array will be insignificant if the low-noise amplifier is shared among a large number of neural amplifiers in the array. As an example, if the low-noise amplifier consumes $50 \mu W$ of power, for the array of 100 neural amplifiers, this would add only $0.5 \mu W$ per neural amplifier. In this work, I assume that the current noise envelope of the background noise is already obtained and it is available to be compared to the current noise envelope of the neural amplifier.

The comparison of the noise envelopes can be done in a continuous-time fashion using a current subtractor. The schematic of a current subtractor is shown in Fig. 5-7. The transfer function of the current subtractor can be described by

$$I_{sub,out}(s) = \frac{1}{M} (I_{env,BG}(s) - I_{env,out}(s)). \quad (5.7)$$

Table 5.1: Parameters of the peak detector

Parameter	Value
I_a	2 nA
C_a	1 nF
τ_a	1.85×10^{-2} s
I_r	4 pA
C_r	20 pF
τ_r	1.85×10^{-1} s

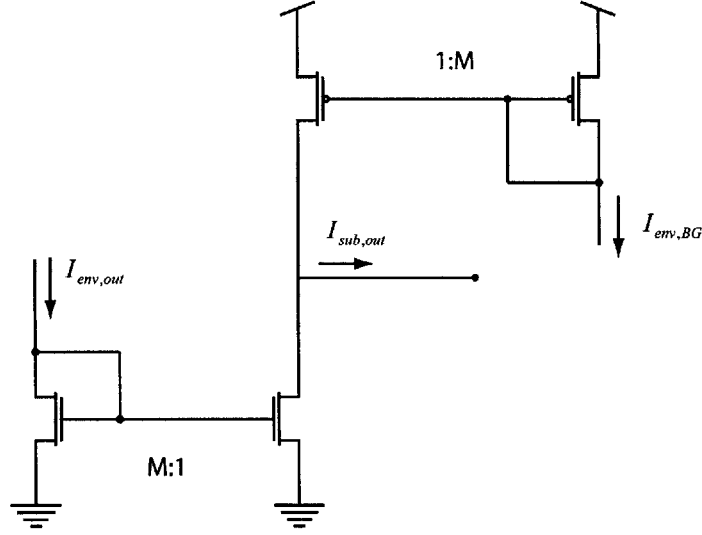


Figure 5-7: A schematic of current subtractor and lead compensation network.

5.3.3 Lead Compensation Network

The output of the current subtractor can be converted into a voltage by integrating $I_{sub,out}$ onto a capacitor C_{int} . The lead compensation for the feedback loop stabilization is implemented with an R_{int} added in series with C_{int} . The transfer function from $I_{sub,out}$ to the voltage V_{int} is described by

$$V_{int}(s) = \left(R_{int} + \frac{1}{sC_{int}} \right) \cdot I_{sub,out}(s) = \frac{1 + sR_{int}C_{int}}{sC_{int}} \cdot I_{sub,out}(s). \quad (5.8)$$

5.3.4 Bias Decision Circuitry

Once the error current $I_{sub,out}$ is obtained, the bias decision circuitry is needed to set the bias current of the neural amplifier based on the information provided by $I_{sub,out}$. The bias decision circuitry will adjust the neural amplifier's bias current until the neural amplifier's noise envelope $I_{env,out}$ is equal to the background noise's envelope $I_{env,BG}$. The block diagram schematic of the bias decision circuitry is shown in Fig. 5-8. The G_{mB} OTA converts V_{int} into a current I_{GmB} which is linearly proportional to V_{int} . The current inversion circuit calculates the inverse of I_{GmB} to produce the output current $I_{inv,out}$. The current inversion is necessary to ensure that the loop is

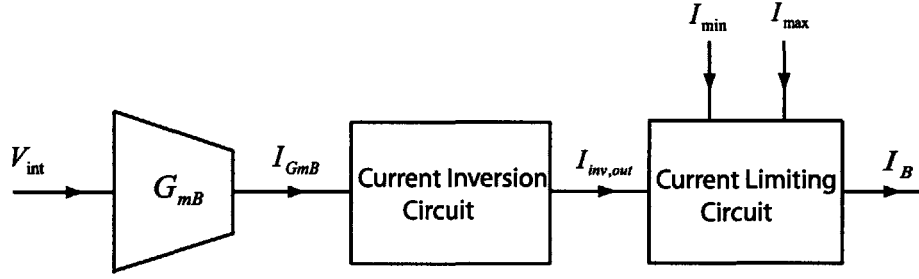


Figure 5-8: A schematic of the Biasing Decision Circuitry.

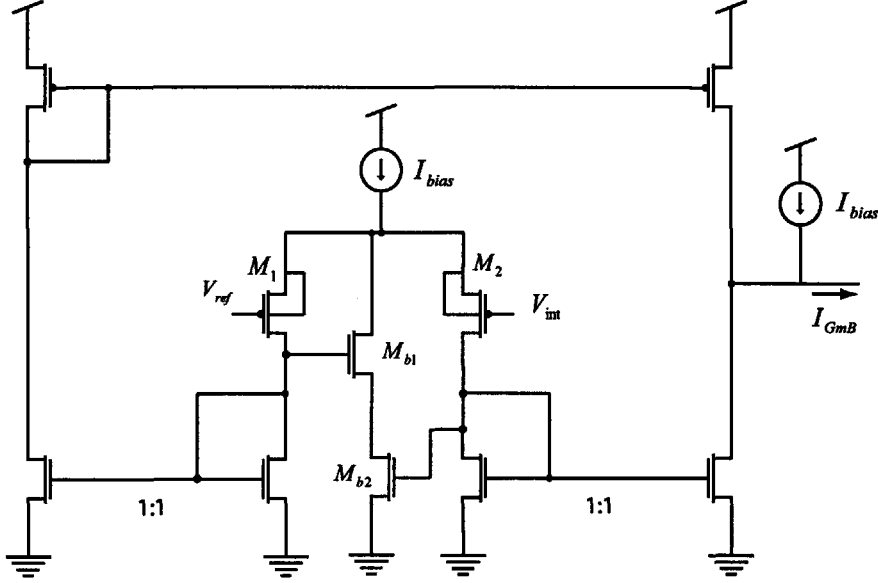


Figure 5-9: A schematic of the G_{mB} OTA.

in a negative feedback form. Then the current $I_{inv,out}$ is passed to the current limiting circuit to ensure that the bias current I_B of the neural amplifier is $I_{min} \leq I_B \leq I_{max}$.

The schematic of the G_{mB} OTA is shown in Fig. 5-9. The transistors M_{b1} and M_{b2} are included to implement the method of bump linearization [11] to increase the linear range of the OTA. Once the calibration loop is in lock, we can assume that $V_{int} - V_{ref}$ is within the linear range of the OTA. We can then express the output current I_{GmB} as

$$I_{GmB} = G_{mB} \cdot (V_{int} - V_{ref}) + I_{bias}. \quad (5.9)$$

A copy of I_{bias} is added to the output of the OTA to ensure that I_{GmB} is always positive since the current inversion circuit can only operate with one input current's

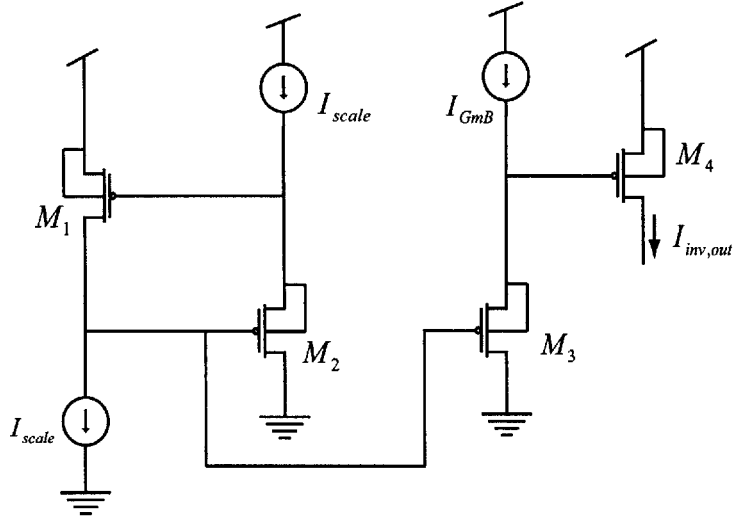


Figure 5-10: A schematic of current inversion circuit.

polarity.

The circuit schematic of a current inversion circuit is shown in Fig. 5-10. The transistors M_1 - M_4 are sized with the same W/L ratio and are biased in subthreshold to achieve an exponential I-V characteristic. In subthreshold, the current in each transistor M_i can be described by

$$I_i = I_{s,i} \cdot e^{\kappa V_{sg,i}/U_T}. \quad (5.10)$$

Since $I_{s,i}$ is a proportional constant and is the same for all transistors, we can denote it as I_s . Therefore, the source-gate voltage for each transistor can be derived from (5.10) to be

$$V_{sg,i} = \frac{U_T}{\kappa} \cdot \ln \left(\frac{I_i}{I_s} \right). \quad (5.11)$$

From the circuit of Fig. 5-10, using KVL we have

$$V_{sg,1} + V_{sg,2} = V_{sg,3} + V_{sg,4}. \quad (5.12)$$

Therefore, we have

$$\frac{U_T}{\kappa} \ln \left(\frac{I_1}{I_s} \right) + \frac{U_T}{\kappa} \ln \left(\frac{I_2}{I_s} \right) = \frac{U_T}{\kappa} \ln \left(\frac{I_3}{I_s} \right) + \frac{U_T}{\kappa} \ln \left(\frac{I_4}{I_s} \right). \quad (5.13)$$

Equation (5.13) can be reduced to

$$I_1 \cdot I_2 = I_3 \cdot I_4. \quad (5.14)$$

Thus, we have $I_{scale} \cdot I_{scale} = I_{GmB} \cdot I_{inv,out}$. Therefore, we get the output current $I_{inv,out}$ to be

$$I_{inv,out} = \frac{I_{scale}^2}{I_{GmB}}. \quad (5.15)$$

A scaled version of the current $I_{inv,out}$ can be used to bias the neural amplifier directly. However, during transient operation when the calibration loop is far from the lock position, the current $I_{inv,out}$ may be too large or too small. If a scaled version of $I_{inv,out}$ was used to bias the neural amplifier, it would cause the amplifier to stop functioning and the calibration loop would never return to the lock position. In order to prevent this problem, I have added the current limiting circuitry to ensure that the bias current of the amplifier is within a specified range that guarantees the correct operation of the neural amplifier. The schematic of the current limiting circuitry is shown in Fig. 5-11. The input current of the current limiting circuitry is $I_{inv,out}$ and its output current is I_B . The current I_B is used to bias the amplifier. The small current I_{min} is added to guarantee that the input current of the circuit is at least I_{min} . The transistors M_3 - M_5 form a Wilson current mirror. Instead of having one current input and one current output, this Wilson mirror has two current inputs which are the currents in M_2 and M_8 . First, let's consider when I_8 , the current in M_8 , is greater than I_2 , the current in M_2 . Originally, the mirror will try to equalize the currents in M_4 and M_5 . Since I_5 is greater than I_4 but M_4 and M_5 have the same gate-source voltage, the drain voltage of M_5 which is the gate voltage of M_3 will be pulled down until M_3 and M_8 enter linear region and shut down the current in M_8 . As a result, the current that is mirrored by M_6 is determined by M_4 , thus it is equal

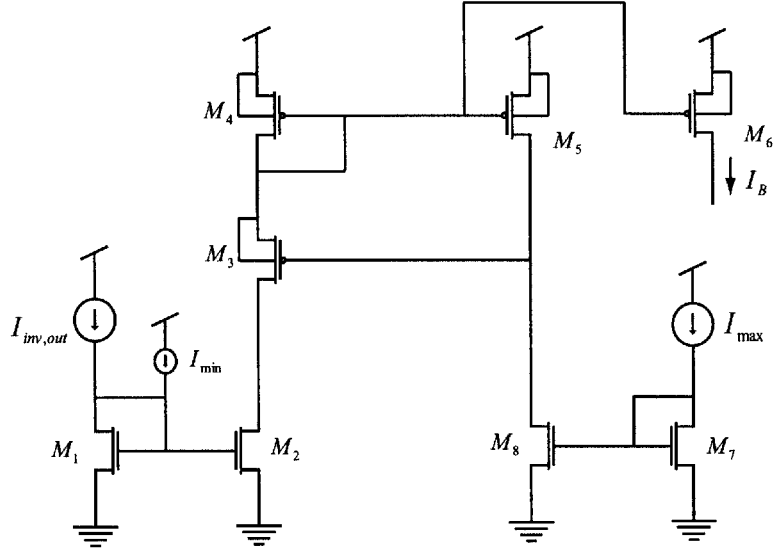


Figure 5-11: A schematic of the current limiting circuit.

to $I_{inv,out} + I_{min}$, assuming a 1:1 mirror ratio between M_1 and M_2 . Next, let's consider when $I_{inv,out} + I_{min} > I_{max}$. The gate-source voltage of M_4 and M_5 will be larger than needed for M_5 to pass I_{max} . Therefore, the gate voltage and the source voltage of M_3 will rise to shut down the current in M_4 . Since the current in M_2 is greater than the current in M_4 , the drain voltage of M_2 will be discharged until M_2 enters the linear region and has the same current as that of M_4 . As a result, the current in M_4 will be equal to the current in M_5 which is equal to I_{max} , assuming a 1:1 mirror ratio between M_7 and M_8 . Therefore, I_B which is a copy of the current in M_5 will be equal to I_{max} . Due to this current limiting process, we can write the current I_B as

$$I_B = \min(I_{min} + I_{inv,out}, I_{max}). \quad (5.16)$$

5.4 Feedback Analysis of the Input-Referred Noise Calibration Loop

During the input-referred noise calibration when the value of $I_{env,out}$ is far from $I_{env,BG}$, the feedback loop is very nonlinear. The analysis of this nonlinear behavior

is very complex and beyond the scope of this thesis. In this work, I shall consider a linearized model when the loop is near the lock position.

The dynamics of the envelope detector, the current subtractor and the lead compensation network have been described in Section 5.3. To model the biasing decision circuitry, we will assume that the loop is close to the lock position, therefore the current in each building block of the biasing decision circuitry does not change appreciably. As a result, we can linearize the behavior of each building block in Fig. 5-8 to obtain its transfer function. In this feedback loop system, the loop bandwidth is very small since it must be able to filter out the noisy transients of the neural amplifier's noise envelope. The time constants of the envelope detector are designed to be very low to smooth out the spurious transients in the amplifier's noise envelope. The lead compensation network provides a pole at the origin and a low-frequency zero for frequency compensation. The poles and zeros created by other circuit building blocks can be ignored since they are at much higher frequencies than the loop bandwidth.

First, let's model the relationship between the output noise of the neural amplifier and its bias current. Let's assume that the neural amplifier is thermal-noise limited and that it uses a subthreshold differential pair as an input stage. Furthermore, let's assume that the only significant noise sources in the amplifier are the two differential-pair transistors. Let $V_{n,in}$ be the total integrated noise of the neural amplifier. We know that

$$V_{n,in} \propto \sqrt{\frac{1}{g_m}} \propto \sqrt{\frac{1}{I_B}} \quad (5.17)$$

where g_m is the transconductance of one of the input differential-pair transistors and I_B is the total supply current of the neural amplifier. Therefore, we can approximate the output noise of the neural amplifier as a function of the total supply current, assuming that the output noise is linearly proportional to the input-referred noise, as

$$V_{n,out} = \frac{K_n}{\sqrt{I_B}} \quad (5.18)$$

where K_n is a proportionality constant. Therefore, we obtain the noise transfer

function of the neural amplifier to be

$$H_{n,amp} = \frac{\partial V_{n,out}}{\partial I_B} = -\frac{1}{2} \cdot \frac{K_n}{I_B \sqrt{I_B}}. \quad (5.19)$$

Next, let's find the transfer function of the bias decision circuitry. From (5.9), we can find the transfer function of G_{mB} OTA to be

$$H_{GmB} = \frac{\partial I_{GmB}}{\partial V_{int}} = G_{mB}. \quad (5.20)$$

The transfer function of the current inversion circuit is

$$H_{inv} = \frac{\partial I_{inv,out}}{\partial I_{GmB}} = -\frac{I_{scale}^2}{I_{GmB}^2}. \quad (5.21)$$

For the current limiting circuit, we assume that the output current does not stuck at either I_{min} or I_{max} since this only happens when the background noise is out of the range to which the neural amplifier's noise can tune. For this condition, we have $I_B = I_{inv,out} + I_{min}$ where $I_{inv,out} > 0$. Thus, we obtain the transfer function of the current limiting circuit to be

$$H_{limit} = \frac{\partial I_B}{\partial I_{inv,out}} = 1. \quad (5.22)$$

We can then draw a small-signal feedback block diagram as shown in Fig. 5-12. The loop transmission $L(s)$ of the input-referred noise calibration loop is given by

$$L(s) = \frac{1}{M} \cdot H_{GmB} \cdot H_{inv} \cdot H_{n,amp} \cdot H_{env}(s) \cdot \frac{1 + sR_{int}C_{int}}{sC_{int}} \quad (5.23)$$

$$= \frac{G_{mB}}{M} \cdot \left(\frac{I_{scale}^2}{I_{GmB}^2} \right) \cdot \left(\frac{1}{2} \frac{K_n}{I_B \sqrt{I_B}} \right) \frac{G_{env}}{1 + \tau_i s} \cdot \frac{1 + sR_{int}C_{int}}{sC_{int}}. \quad (5.24)$$

We can write (5.24) in a simplified form as

$$L(s) = A_0 \cdot \frac{1 + \tau_{int}s}{(1 + \tau_{env}s)s} \quad (5.25)$$

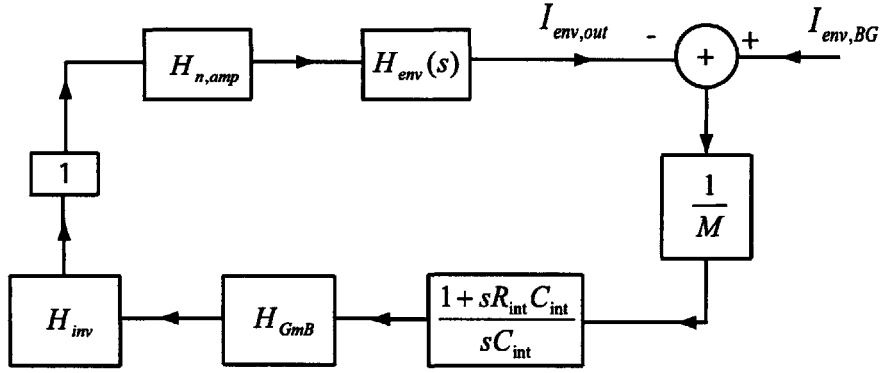


Figure 5-12: A small-signal feedback block diagram of the input-referred noise calibration loop.

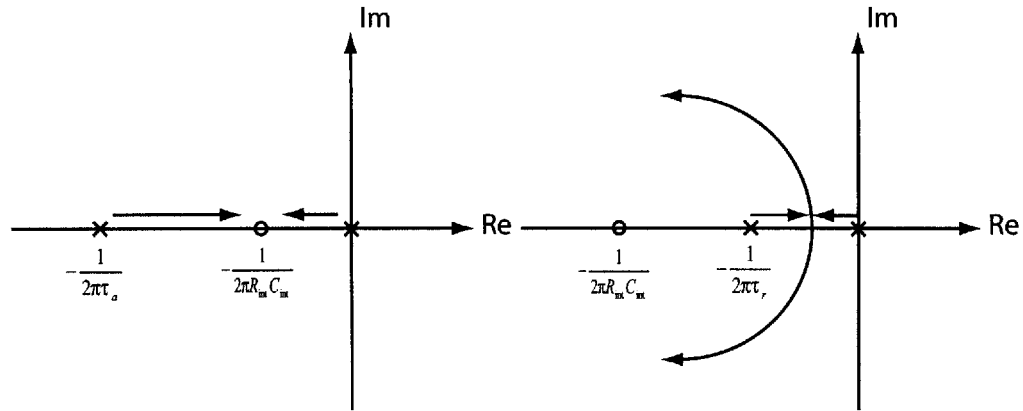
where

$$A_0 = \frac{G_{mB}}{M} \cdot \left(\frac{I_{scale}^2}{I_{GmB}^2} \right) \cdot \left(\frac{1}{2} \frac{K_n}{I_B \sqrt{I_B}} \right) \cdot \frac{G_{env}}{C_{int}} \quad (5.26)$$

is the loop gain and is frequency-independent.

From (5.25), the loop transmission has a pole at the origin from the lead compensation network and a low-frequency pole at $-1/\tau_{env}$ from the envelope detector. To ensure stability, the lead compensation network provided a left-half plane zero at $-1/\tau_{int}$ where $\tau_{int} = R_{int}C_{int}$. Due to asymmetric attack and release time constants of the peak detector, τ_{env} has two values depending on whether the neural amplifier's noise envelope is increasing or decreasing. When the noise envelope is increasing, $\tau_{env} = \tau_a$, whereas $\tau_{env} = \tau_r$ when the noise envelope is decreasing.

Let us consider the dynamic of the loop when the noise envelope is increasing. During this time, the envelope detector is in the attack mode and $\tau_{env} = \tau_a$. Using the approximated value of τ_a shown in Table 5.1, we have a pole due to the envelope detector at $1/(2\pi\tau_a) = 8.6$ Hz. This will create a left-half-plane zero at $1/(2\pi R_{int}C_{int}) = 3.18$ Hz. However, for the case when the neural amplifier's noise envelope is decreasing, the envelope detector is in the release mode and $\tau_{env} = \tau_r$. In this case, we have a pole due to the release time constant at $1/(2\pi\tau_r) = 0.86$ Hz. The position of the pole due to the release time constant makes it difficult to compensate the calibration loop for optimum performance. In order to have no overshoot, the left-half-plane zero at $1/(2\pi R_{int}C_{int})$ should be below the pole due to the release time



(a) Root locus plot (increasing noise envelope) (b) Root locus plot (decreasing noise envelope)

Figure 5-13: Root locus plots (a) Neural's amplifier's noise envelope is increasing (b) Neural's amplifier's noise envelope is decreasing.

constant. However, if such approach were used, the loop would settle very slowly since the closed loop poles will be on the real axis at the frequencies much less than 1 Hz. To make the loop settle faster, the left-half-plane zero is placed at a lower frequency than the pole due to τ_a but at a higher frequency than the pole due to τ_r . To achieve this R_{int} and C_{int} are made from off-chip components with the values of 10 M Ω and 5 nF respectively, corresponding to a left-half-plane zero at 3.18 Hz. The root-locus diagrams for the loop when the neural amplifier's noise envelopes are increasing and decreasing are shown in Fig. 5-13(a) and 5-13(b) respectively.

Bibliography

- [1] J.K. Chapin, K.A. Moxon, R.S. Markowitz, and M.L. Nicolelis. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nature Neuroscience*, pages 664–670, 1999.
- [2] T. Delbruck and C.A. Mead. Analog vlsi adaptive, logarithmic wide-dynamic-range photoreceptor. *Proc. Intl. Symp. Circuits and Systems. (ISCAS)*, 4:339–342, 1994.
- [3] C.C. Enz, F. Krummenacher, and E.A. Vittoz. An analytical mos transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications. *Analog Integrat. Circuits Signal Process*, 8:83–144, 1995.
- [4] K.S. Guillory and R. A. Normann. A 100-channel system for real time detection and storage of extracellular spike waveforms. *J. of Neuroscience Methods*, 91:21–29, 1999.
- [5] Reid R. Harrison, Paul T. Watkins, Ryan J. Kier, Robert O. Lovejoy, Daniel J. Black, Bradley Greger, and Florian Solzbacher. A low-power integrated circuit for a wireless 100-electrode neural recording system. *J. of Solid State Circuits*, 42(1):123–133, January 2007.
- [6] R.R. Harrison and C. Charles. A low-power low-noise cmos amplifier for neural recording applications. *IEEE J. Solid State Circuits*, 38(5):958–965, June 2003.

- [7] J. Mavoori, A. Jackson, C. Diorio, and E. Fetz. An autonomous implantable computer for neural recording and simulation in unrestrained primates. *J. of Neuroscience Methods*, 148:71–77, 2005.
- [8] P. Mohseni and K. Najafi. A fully integrated neural recording amplifier with dc input stabilization. *IEEE Trans. on Biomedical Engineering*, 51(5):832–837, May 2004.
- [9] Yevgeny Perelman and Ran Ginosar. An integrated system for multichannel neuronal recording with spike/lfp separation, integrated a/d conversion and threshold detection. *IEEE Trans. on Biomedical Engineering*, 54(1):130–137, January 2007.
- [10] Roy H. Olsson, Derek L. Buhl, Anton M. Sirota, Gyorgy Buzsaki, and Kendall D. Wise. Band-tunable and multiplexed integrated circuits for simultaneous recording and stimulation with microelectrode arrays. *IEEE Trans. on Biomedical Engineering*, 52(7):1303–1311, July 2005.
- [11] Rahul Sarpeshkar, Richard F. Lyon, and Carver Mead. A low-power wide-linear-range transconductance amplifier. *Analog Integrated Circuits and Signal Processing*, 13(1-2):123–151, 1997.
- [12] Hansjörg Scherberger, Murray R. Jarvis, and Richard Andersen. Cortical local field potential encodes movement intentions in the posterior parietal cortex. *Neuron*, 46:347–354, April 2005.
- [13] M.S.J Steyaert, W.M.C. Sansen, and C. Zhongyuan. A micropower low-noise monolithic instrumentation amplifier for medical purposes. *IEEE J. Solid-State Circuits*, SC-22:1163–1168, December 1987.
- [14] Yannis Tsividis. *Operation and Modeling of the MOS Transistor*. McGraw-Hill, 2nd edition, 1998.
- [15] J. Wessberg, C.R. Stambaugh, J.D. Kralik, P.D. Beck, M. Laubach, J.K. Chapin, J. Kim, S.J. Biggs, M.A. Srinivasan, and M.L. Nicolelis. Real-time prediction of

hand trajectory by ensembles of cortical neurons in primates. *Nature*, 408:361–365, November 2000.

- [16] Serhii M. Zhak, Michael W. Baker, and Rahul Sarpeshkar. A low-power wide dynamic range envelope detector. *IEEE Journal of Solid-State Circuits*, 38(10):1750–1753, October 2003.