Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

Bruno Valentin Seifert

A 1.05 NEF stacked cascode inverter-based amplifier for neural signal recording

Master's thesis in Electronic Systems Design
Supervisor: Trond Ytterdal, Erwin Habibzadeh Tonekabony Shad
June 2021

NTNU
Kunnskap for en bedre verden

Bruno Valentin Seifert

A 1.05 NEF stacked cascode inverter-based amplifier for neural signal recording

Master's thesis in Electronic Systems Design
Supervisor: Trond Ytterdal, Erwin Habibzadeh Tonekabony Shad
June 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

NTNU
Norwegian University of
Science and Technology

Abstract

Low noise amplifiers are an important component of modern wearable biosensors. These sensor systems typically have a heavily limited battery capacity which makes power efficiency an important factor when designing their parts.

This work presents a stacked cascode inverter-based amplifier as a feasible topology for current-efficient neural signal amplifiers. It is a highly noise efficient design with low power usage that still can provide high gain and a multiple kHz bandwidth. To minimize the supply voltage some of the input transistors also serve as current mirrors. Special care has also been taken to be able to present a robust amplifier despite a V_{ds} of around 100 mV for most of the transistors. Negative DC-feedback has been incorporated as well as especially robust CM-feedback to ensure the stability of the output nodes. The cascode transistors are biased by diode-connected transistors in a compromise between overall complexity and robustness.

With a supply voltage of 1.35 V and a total bias current of 0.5 µA, the designed amplifier can provide a closed-loop gain of 40 dB on two channels. This gives a power consumption of only 337.5 nW per channel. Simultaneously the system is also highly noise efficient with an NEF of 0.900 when excluding flicker noise. When including flicker noise the NEF is 1.052 which is still a low value compared to other state-of-the-art systems. The noise floor is at 35.7 nV/$\sqrt{\text{Hz}}$, despite a parasitic input capacitance of nearly 2.5 pF diminishing the effectiveness of the amplifier. The bandwidth of the amplifier is from 3.5 Hz to 10 500 Hz with a load capacitance of 1.5 pF.

Sammendrag

Strømeffektive lavstøy forsterkere er en viktig del av moderne bærbare biosensorer på grunn av deres liten størrelse og lav batterikapasitet.

Denne rapporten presenter en stablet forsterker med en cascode-inverterer basert topologi som et veldig strømeffektivt design. Forsterkeren er meget støyeffektiv, bruker lite energi og leverer uansett høy forsterkning med en båndbredde på flere kiloHz. For å minimere forsyningsspenningen brukes noen av inntakstransistorene også som strømspeil. For å kunne presentere en robust forsterker til tross for at de fleste transistorer har en V_{ds} av rundt 100 mV bruker designet negativ feedback og spesielt hensyn ble lagt på stabiliteten av common-mode tilbakemeldingskretsen. Forspenningen av cascode-transistorene blir levert av diode-tilkoblete transistorer for å balansere kompleksheten og robusthet.

Med en forsyningsspenning av 1.35 V og en forspenningsstrøm på 0.5 µA er den designede forsterkeren i stand til å gi en lukket sløyfeforsterkning på 40 dB på to kanaler. Dette gir et strømforbruk på bare 337.5 nW per kanal. Samtidig er systemet også svært støyeffektivt med en NEF på 0.900 når det ikke er flimmerstøy. Hvis vi inkluderer flimmerstøy, er NEF 1.052, noe som fremdeles er en lav verdi sammenlignet med andre toppmoderne systemer. Termisk støygulv er på 35.7 nV/$\sqrt{\text{Hz}}$, til tross for en parasittisk inngangskapasitans på nesten 2.5 pF som reduserer effektiviteten til forsterkeren. Forsterkerens båndbredde varierer fra 3.5 Hz til 10.5 kHz med en lastekapasitans på 1.5 pF.

Preface

This master thesis is part of a degree of master of science in Electronic Systems Design at the department of electronics and telecommunications, Norwegian University of Science and Technology in the spring semester of 2021, written under the supervision of Trond Ytterdal and Erwin Habibzadeh Tonekabony Shad.

I would like to thank both of them most kindly for their help, knowledge, and not least patience while supporting me throughout this final semester. You have taught me a great deal in both, theoretical knowledge and practical skills, while also providing me with the motivation needed to persevere through this complex assignment.

Trondheim, Norway, 17. June 2021

Bruno Seifert

Problem description

This thesis sets out to describe the design of a low noise amplifier compatible with EEG and neural signal recording applications. The amplifier is supposed to be a fully differential, multi-channel, AC-coupled topology implemented in a commercially available 180 nm CMOS technology.

Table 1: The proposed specifications of the amplifier

Output capacitance	1.5 pF
Power	1 µW
CL-gain per channel	40 dB
Bandwidth	10 kHz
CMRR	60 dB
PSRR	60 dB
Input-referred noise	40 nV/$\sqrt{\text{Hz}}$

The required specifications of the system are noted in Table 1. These were set to ensure that the amplifier can work reliably with a range of different neural signals. To achieve the specified gain of 40 dB and the AC-coupling without using a very large area the input capacitance is set to 10 pF with a feedback capacitance of 0.1 pF.

The main focus of this design is to be noise efficient, measured by its noise efficiency factor (NEF) and power efficient. To optimize for noise efficiency the input referred noise and the current consumption are the most important factors to minimize. Therefore, the proposed amplifier should use current reuse techniques like an inverter-based topology and a structure where multiple channels are stacked on top of each other.

The finished design is to be tested for robustness with corner and Monte Carlo simulations including process variation and stochastic mismatch. A layout for the amplifier is to be designed and post layout simulations are also to be done.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Electroencephalography (EEG) is a common tool used for many different purposes in medicine and science: Whether for neuroscience research [1], development of novel means to control prosthetics [2] or to diagnose and study illnesses [3]. One of the current obstacles for increased use is the size and practicability of these systems which is why there is a focus on developing more easily wearable and unobtrusive solutions [3]. Being able to use an EEG over long time periods while going about one's daily life opens up new possibilities in monitoring patients and the use of brain-machine interfaces. However, current wearable EEG systems still face technical challenges as they are limited by battery life. These batteries usually take up roughly half of the available space [4], so to improve upon this more power-efficient designs are necessary [5]. For the amplifiers, the most important factor next to power efficiency is low noise operation[6].

Therefore, state-of-the-art neural signal amplifiers have recently begun pushing further towards and beyond a noise efficiency factor (NEF) of 1 [7]. NEF values as low as 1.07 have been achieved [8].

This thesis will, in detail, describe the topology of a stacked cascode inverter-based amplifier optimized for noise efficiency and the use in portable EEG and neural signal recording systems. This amplifier is a dual-channel, fully differential, AC-coupled design in 180 nm CMOS technology. The cascode topology allows this design to have a high gain despite a very low bias current.

This design has two major challenges. The first is due to the capacitive coupling: The limited area of wearable biosensors means that the coupling capacitance cannot be large and therefore the Miller capacitance at the input transistors gates can cause significant problems. It both reduces the amplifier's gain and increases the input-referred noise, therefore the total size of all input transistors has to be kept at a minimum. But two other factors also play a role in this: MOSFETs are most efficient when in weak inversion [9] but this requires a larger W/L ratio of the transistors and the amount of flicker noise generated by a MOSFET is inversely proportional to its size. Balancing the size of the input transistors to satisfy all of these factors is an important part of the design. The second is in achieving a minimal power consumption because Stacking input stages requires an increase

in supply voltage [6]. To nonetheless improve the power consumption while creating a robust system the V_{ds} and bias voltages of all transistors have to be set carefully. Despite these shortcomings the amplifier achieves an NEF of 1.052 with a power consumption of 337.5 nW per channel.

1.1 Outline

The content of this thesis is organized as follows:

At first Section 2 gives an overview of the most important points for noise efficient design as well as inverter-based and stacked topologies together with a short overview of neural signals and their characteristics.

Section 3 lays out the exact topology of the proposed system and its characteristics. It also explains how to best minimize noise and what the biggest challenges of the design are.

Section 4 goes through the details of and reasoning for this specific implementation of the design.

Then, Section 5 presents the results of this implementation and in Section 6 important details of the results are examined further, and finally the amplifier is compared to similar, recent works.

Finally, Section 7 contains a summary of the results and proposes possible future continuations of this work.

Chapter 2

Theoretical background

2.1 Noise in MOSFETs

As with other solid-state components, the three noise sources, flicker, thermal, and shot noise, are also present in MOSFETs. In MOSFETs, flicker noise dominates at low frequencies and the white noise is largely made up of thermal noise originating from the channel resistance when in the strong inversion. In the subthreshold region, the dominant white noise mechanism is shot noise, though [10].

2.1.1 Thermal noise

Thermal noise is generated by the channel resistance of the transistor. This resistance is not homogeneous when in the active region and therefore the noise current has to be calculated by integrating over small regions of the channel which gives [11]

$$I_n^2, th = 4kT\gamma g_m \tag{2.1}$$

where k is the Boltzmann constant, T the temperature in Kelvin, γ an operating point and geometry dependent white noise parameter that can be approximated as $2/3$ and g_m the transconductance of the MOSFET. This shows that the intensity of the noise is dependent on the transconductance of the device, however, this can be misleading. When the device is also an input transistor the input-referred noise is usually inversely related to the transconductance as the output noise is divided by the gain squared. For example, the input-referred noise voltage of an input transistor in a common source stage is $4kT\gamma/g_m$. Therefore in most applications, one would want to maximize the overall transconductance of the system to minimize the effects of thermal noise.

2.1.2 Shot noise

Shot noise is created by the granular nature of electric current. Since a current is made up of many individual charge carriers the number of carriers passing

a pn-junction is not fixed and rather pulse-like[11]. This fluctuation generates additional noise which in MOSFETs is characterized as [10]

$$I_{n,sh}^2 = 2qI_{ds} \tag{2.2}$$

q stands for the elementary charge and I_{ds} is the DC bias current through the device.

This equation can be rewritten to resemble (2.1) by recognizing that the drain current of a transistor in the subthreshold region is dominated by the diffusion mechanism and therefore roughly equal to $kT/q\eta g_m$ [12]. So the shot noise is

$$I_{n,sh}^2 = 2kT\eta g_m \tag{2.3}$$

where η is equal to $(g_m + g_{mb})/g_m \approx 1.5$ [12]. Thus, while the mechanisms of shot noise and thermal noise are different, the resulting noise in the circuit is similar, no matter whether the transistor is operating in the superthreshold or the subthreshold region.

2.1.3 Flicker noise

The exact cause of flicker noise is still unclear. It is hypothesized to arise due to charge carriers getting temporarily stuck in traps between the channel and the oxide, or fluctuations in the mobility of the free carriers [13], impeding the flow of a DC current. The noise power quickly falls when frequency increases as AC current flow is less impeded by this effect. Because of this, flicker noise is also referred to as 1/f noise [11].

The power of flicker noise is highly dependent on details of the technology used as the number of traps in the semiconductor can vary wildly [13]. Flicker noise is commonly modeled as a voltage source in series with the gate instead of a noise current and in the saturation region given by [14]

$$V_{n,fl}^2 = \frac{K}{C_{ox}WLf} \tag{2.4}$$

where C_{ox} is the oxide capacitance, W and L are the device's width and length and K is a technology-dependent constant. Holes as charge carriers are less likely to become trapped so p-channel devices tend to have a much lower K and therefore exhibit much less flicker noise than n-channel devices [14].

From this, we see that flicker noise is largely independent of the biasing of a transistor and instead dependent on its size. A larger transistor generates less flicker noise, therefore, to minimize its effect one would want as large transistors as possible.

Because flicker noise is dependent on the frequency it can, however, be removed with chopping and filtering techniques. Therefore it is less of a concern when looking at the noise efficiency of an amplifier than white noise.

2.2 Figure of merit

To be able to accurately and fairly rate the performance of the amplifier the noise efficiency factor (NEF) and the power efficiency factor (PEF) figure of merits will be used.

The NEF was first introduced by [15] to be able to compare the noise of a system with the thermal noise produced by an ideal bipolar transistor with the same bandwidth and current drain and is defined as

$$NEF = V_{n,i,rms}\sqrt{\frac{2I_{tot}}{\pi * V_T * 4kT * BW}} \tag{2.5}$$

where V_T is the thermal voltage, k is Boltzmann's constant, T is the temperature, I_{tot} is the total current drawn by the amplifier, BW is its bandwidth, and $V_{n,i,rms}$ is the root-mean-square of the total input-referred noise. $V_{n,i,rms}$ is obtained by integrating the noise over the noise bandwidth which is equal to a brickwall-equivalent of the real bandwidth, i.e. $\pi/2 * BW$ [11]. A typical CMOS differential pair has a theoretical NEF of 2.02 [16]. However, modern ultra low-power amplifiers have broken this barrier by reusing the same current for multiple, gain-providing transistors. In recent years CMOS differential amplifiers using this technique that are even more efficient than an ideal bipolar transistor have been developed, like the one presented in [17] with an NEF of 0.84.

The PEF developed is an extension of the NEF to take into account the power consumption which is calculated by [18]

$$PEF = NEF^2 * V_{DD} \tag{2.6}$$

2.3 Current efficient amplifiers

To achieve highly noise-effective amplifiers different techniques have been developed to reuse the amplifier's bias current. When a current is used multiple times for amplification the input-referred noise is reduced while keeping the current use constant, thus reducing the NEF.

2.3.1 Inverter-based amplifiers

An inverter-based amplifier is similar to a simple common-source amplifier, but the load is instead replaced by another common source connected transistor as shown in Fig. 2.1, building an inverter cell. The gates of these two amplifiers are both connected to the input signal which means the signal is amplified by both of them, doubling the amplifier's transconductance. The output resistance is made up of both transistor's small-signal resistance in parallel which gives an amplification of

$$A = \frac{2g_m}{2g_{ds}} = \frac{g_m}{g_{ds}} \tag{2.7}$$

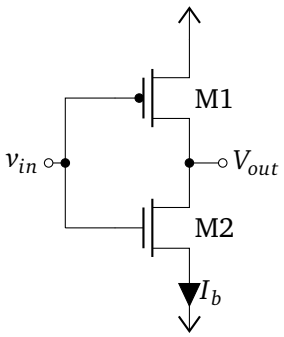


Figure 2.1: An inverter-based amplifier

So this technique doubles the gain without changing the power consumption of the system. In addition, the input-referred noise is more than halved compared to a common-source amplifier. The second input transistor replaces the load and thus eliminates the noise generated by it, though this noise is usually negligible. With just the two input transistors the total input-referred noise becomes

$$v_{n,in}^2 = \frac{2kT\gamma}{g_m} \tag{2.8}$$

From this can be deducted that an inverter-based topology improves noise efficiency by a factor of more than $\sqrt{\frac{1}{2}}$ when used instead of a common source amplifier.

2.3.2 Stacked amplifiers

Another method for current reuse is to stack input stages like in Fig. 2.2 [19]. This way multiple amplifiers work with the current of one. The gain and output referred noise is simply that of every single amplifier added together which means the input-referred noise is, assuming both amplifiers being the same, calculated with

$$v_{n,in}^2 = \frac{2 * v_{n,o,single}^2}{\left(2 * A_{single}\right)^2} = \frac{1}{2}\frac{v_{n,o,single}^2}{A_{single}^2} = \frac{1}{2}v_{n,in,single}^2 \tag{2.9}$$

Again, effectively halving the input-referred noise without changing the bias current.

But to be able to accommodate the additional transistors this topology usually needs a larger voltage headroom, so the improvement in power efficiency would not be as great as in noise efficiency.

Another challenge to stacking infinitely many amplifiers on top of each other is the increasing body effect of the transistors. The increased threshold voltage quickly makes designing a functioning amplifier difficult and it becomes necessary to utilize deep n-wells instead [17].

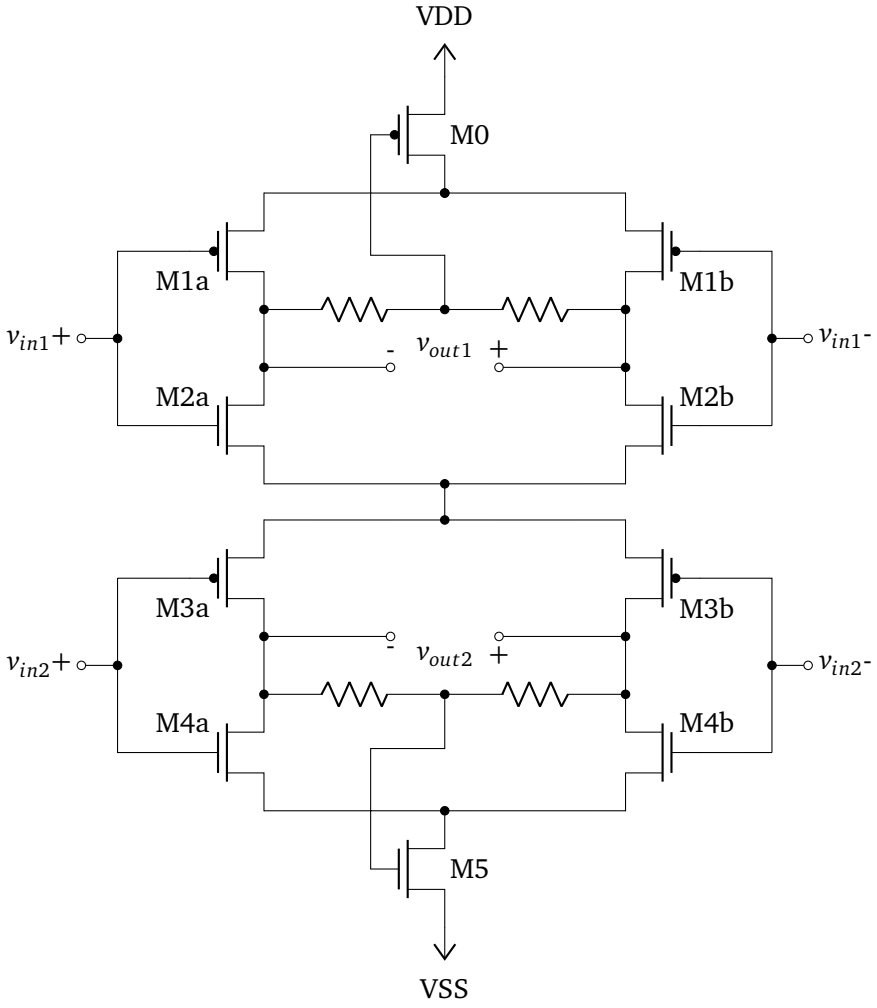


Figure 2.2: A stacked inverter-based amplifier

Anyhow, this topology has a different advantage as well: Each amplifier in the stack can function as a separate amplifier enabling the amplification of different signals with the same current. This can be useful for systems that do not need very high gain but have many channels, like in neural signal recording. In this case, the input-referred noise of every stacked input stage is unchanged, but the bias current of each stage is effectively the total bias current divided by the number of stages stacked. This has the same effect on the NEF as halving the input-referred noise.

2.4 Brain signal measurement

There are multiple types of signals of interest in the brain with different bandwidths and intensities. The most important of these are the brainwaves, the local field potentials (LFPs) and the action potentials (APs) [20].

Brainwaves are a collection of different types of activity patterns in a person's brain with intensities ranging from 0.5 µV to 100 µV in amplitude and have frequencies between 0.5 Hz and 140 Hz [20]. These are the signal measured most often measured with an EEG.

But other types of neural signals are also of interest, even if not measured with an EEG but different techniques.

The LFPs are low-frequency signals that indicate the levels of general activity in an area of the brain [21]. They have an amplitude of around 1 mV and occur in a range of 1 Hz to 200 Hz [22].

The APs on the other hand are weaker signals at higher frequencies, these are the potential differences generated by the activity of a single neuron [20]. The amplitude of these is about 100 µV with a frequency of up to 10 kHz [22].

Most measurements of the brain require multiple electrodes and signal channels to record simultaneously to create a full picture of the activity.

Chapter 3

Amplifier design

When used with a capacitively-coupled input the inverter-based amplifier shown in Fig. 2.1 has a total gain that is susceptible to large parasitic C_{gd}. The Miller effect tells us that the equivalent Miller capacitance, C_M from gate to ground is $C_{gd}(1+A_{gd})$ [23], creating a capacitive voltage divider attenuating the input signal strength at the gates. This is problematic because MOSFETs are most efficient in the subthreshold region and to operate in this region a relatively large W/L ratio is necessary [9]. Additionally, to minimize the necessary supply voltage for the amplifier the gate-to-source voltages have to be minimized which necessitates an even larger W/L ratio [14]. The resulting input signal attenuation logically also affects the input-referred noise inversely. Henceforth making this problem highly detrimental to the aim of creating a noise-efficient amplifier.

To alleviate that flaw and further decrease the NEF this work combines multiple different topologies: A cascode amplifier, an inverter-based amplifier, and a stacked amplifier. The resulting system has the potential for very high gain and efficiency, but also multiple difficulties for the design. The details of this will be explored in this chapter.

3.1 Cascode inverter-based amplifier

Fig. 3.1 shows the schematic of a cascode inverter-based amplifier. The topology and characteristics are similar to the simple inverter-based amplifier but each input stage consists of a cascode instead of a single transistor to provide a larger output resistance. It presents an improvement over an inverter-based design by providing higher gain and less Miller capacitance.

3.1.1 Gain

The transconductance of this cascode inverter-based amplifier is the same as that for a simple inverter-based amplifier, two input transistors amplify the incoming signal giving a G_m that is twice the g_m of an input transistor. The higher gain is

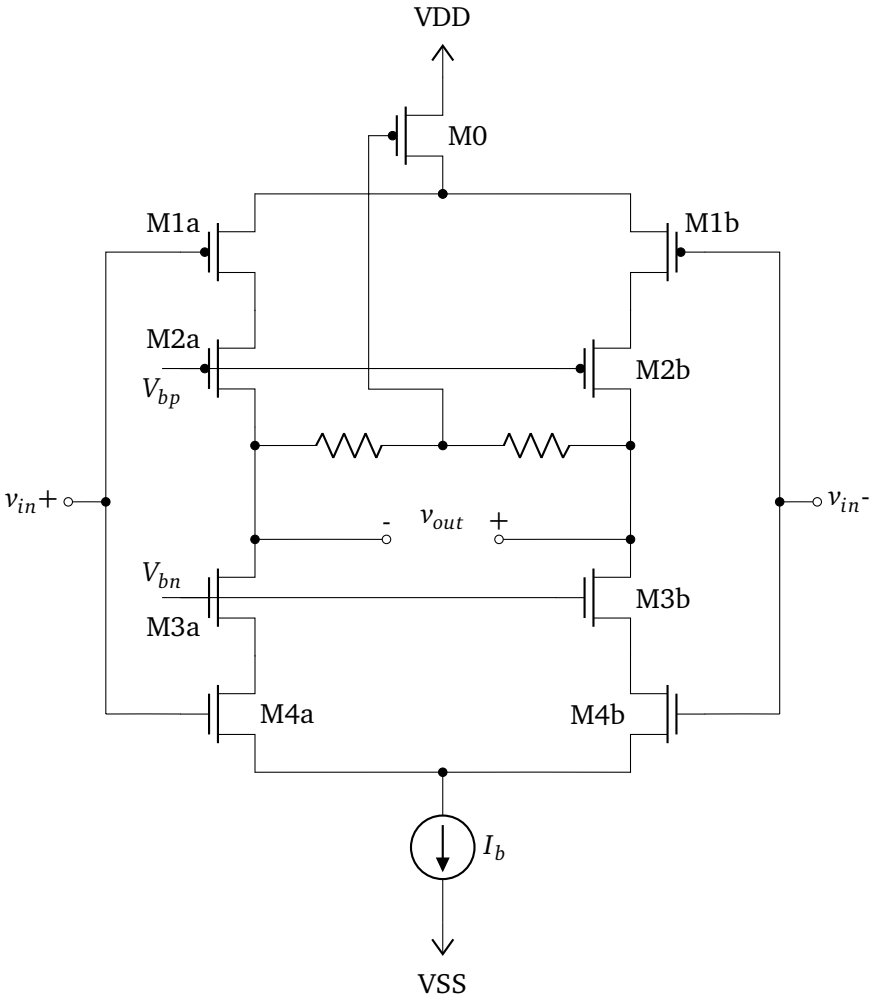


Figure 3.1: A cascode inverter-based amplifier

achieved by a higher output resistance. This resistance is made up of two cascodes in parallel, therefore calculated with

$$R_o = \frac{1}{2}\frac{g_{m,2,3}}{g_{ds,2,3} * g_{ds,1,4}} \tag{3.1}$$

where $g_{m,2,3}$ refers to the transconductance of either of the transistors M2a, M2b, M3a, and M3b as all of these should have the same transconductance. The same applies to the notation of $g_{ds,2,3}$ and $g_{ds,1,4}$. which means the gain of the cascode inverter-based amplifier is

$$A = \frac{g_{m,2,3} * g_{m,1,4}}{g_{ds,2,3} * g_{ds,1,4}} \approx \frac{g_m^2}{g_{ds}^2} \tag{3.2}$$

where g_m and g_{ds} assume that these values are approximately the same. Nevertheless, (3.2) does not characterize the total gain of the system with an ac-coupled input signal. When taking into account this input capacitance and the amplifier's Miller capacitance the gain is

$$A_{tot} = \frac{C_i}{C_i + C_M}\frac{g_m^2}{g_{ds}^2} \tag{3.3}$$

Where C_M is the Miller capacitance from the gate of the input transistors to ground, created by the transistor's parasitic C_{gd} and the Miller effect, and C_i is the AC coupling capacitor between the input signal and the gate. This shows the Miller capacitance plays a significant role in the total gain. The cascode topology has the advantage that the C_{gd} of the input transistors is not directly connected to the output, meaning it is not amplified by the whole gain of the amplifier resulting in a smaller C_M than the simple inverter-based design in Section 2.3.1 for the same amplifier gain.

3.1.2 Frequency response

In a cascode stage the output node is the dominating pole [11] which in this circuit is calculated as

$$\omega_p = \frac{1}{R_o * \left(C_L + C_{db,2,3} + C_{dg,2,3}\right)} \tag{3.4}$$

This means that the dominant pole is heavily limited by the output resistance. Nonetheless, it is critical that this resistance is high to achieve very high gain while having a very low bias current.

The unity-gain frequency of the amplifier should be relatively high due to the inverter-based input providing twice the transconductance of a single transistor, but g_m generally will be very low in power-efficient designs due to the transistors operating in the subthreshold region [9]. The transconductance can be controlled directly by the bias current of the input transistors and the minimum transconductance necessary can be calculated with

$$g_{m,min} = f_{ug} * \pi * C_L \tag{3.5}$$

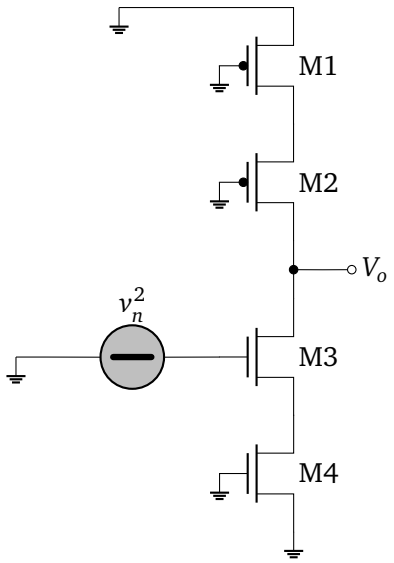


Figure 3.2: A simplified model of the noise generated by one of the cascode transistors with all other independent sources set to 0

3.1.3 Noise

Higher gain via an increase in g_m leads to a decrease in input-referred white noise, if the increase in gain is achieved by higher output resistance, however, it does not affect the noise. The higher gain in (3.2), compared to that of the simple inverter in (2.7) is only achieved by means of higher resistance and thus does not lessen input-referred noise. That means this design's thermal noise is twice that of the single inverter cell in (2.8) due to the differential topology. But only when neglecting the noise generated by the cascode transistors.

Modeling the noise produced by the cascode transistors as a voltage source at the gate of the transistor and setting all other independent sources to 0, like in Fig. 3.2, makes it obvious that the cascode transistors effectively operate as source degenerated common-source amplifiers with a cascode as load. This load resistance is simply $2 * R_o$. Therefore the input-referred noise generated by the cascode transistors is

$$v_{n,i,2,3}^2 = v_{n,2,3}^2 * \frac{g_m^2 (2R_o)^2}{\left(1 + \frac{g_m}{g_{ds}}\right)^2 A^2} = v_{n,2,3}^2 * \frac{1}{\left(1 + \frac{g_m}{g_{ds}}\right)^2} \tag{3.6}$$

In contrast, the input-referred noise one of the input transistors, when also modeled as a source at the transistor's gate, is the noise voltage multiplied by the transconductance of the single input transistor but divided by that of both i.e

$$v_{n,i,1,4}^2 = v_{n,1,4}^2 * \frac{1}{4} \tag{3.7}$$

Since $(1 + g_m/g_{ds})^2 \gg 4$ the noise produced by the cascode transistors can indeed be safely neglected when estimating the noise performance of the amplifier.

Thermal and shot noise

Standard input-referred noise equations assume the signal source sits at the gate of the input transistors. This is not the case here, the source is at the input to the feedforward capacitor. This capacitor forms a capacitive divider with the input capacitance, therefore the input-referred gain is amplified by this divider and the full input-referred noise is written as:

$$v_{n,i} = \frac{C_M + C_i}{C_i} * v_{n,g} \tag{3.8}$$

Where $v_{n,g}$ is the noise referred to the gates of the input transistors. This also means that an increase in gain due to a longer channel's higher resistance can significantly reduce noise efficiency when it substantially increases the input transistor size.

Even though thermal noise tends to be the dominating white noise source in MOSFETs shot noise contributes the majority of the white noise because the amplifier is designed for efficiency and all noise-critical transistors are operating in weak inversion.

For practical purposes, however, this makes little difference because both noise types originate in the channel of the transistor and their input-referred magnitude is inversely related to the transconductance.

Flicker noise

As stated in Section 2.1.3 flicker noise is reduced by increasing the size of the transistors and p-channel devices produce less noise than n-channel, so the most important parameter for the amount of flicker noise in the circuit is the size of M4 and it should be maximized. But (3.8) also applies here which leads to a trade-off between white noise and flicker noise in the circuit as well as diminishing returns. However, while flicker noise scales linearly with transistor size, (3.8) does not. When $C_M \gg C_i$ it is a roughly linear relation but this should not be the case, so for low levels of parasitic capacitance an increase in transistor size is beneficial. Nonetheless, at low frequencies flicker noise will be non-negligible due to this so it is advocated to implement other techniques to fully eliminate its effects.

3.2 Stacked topology

To increase efficiency further the cascode inverter-based amplifier is used as a double stacked amplifier as seen in Fig. 3.3. This, of course, comes at a cost of heightened supply voltage but as explained in Section 2.3.2 halves the total input-referred noise of the design and gives us the possibility to either have double

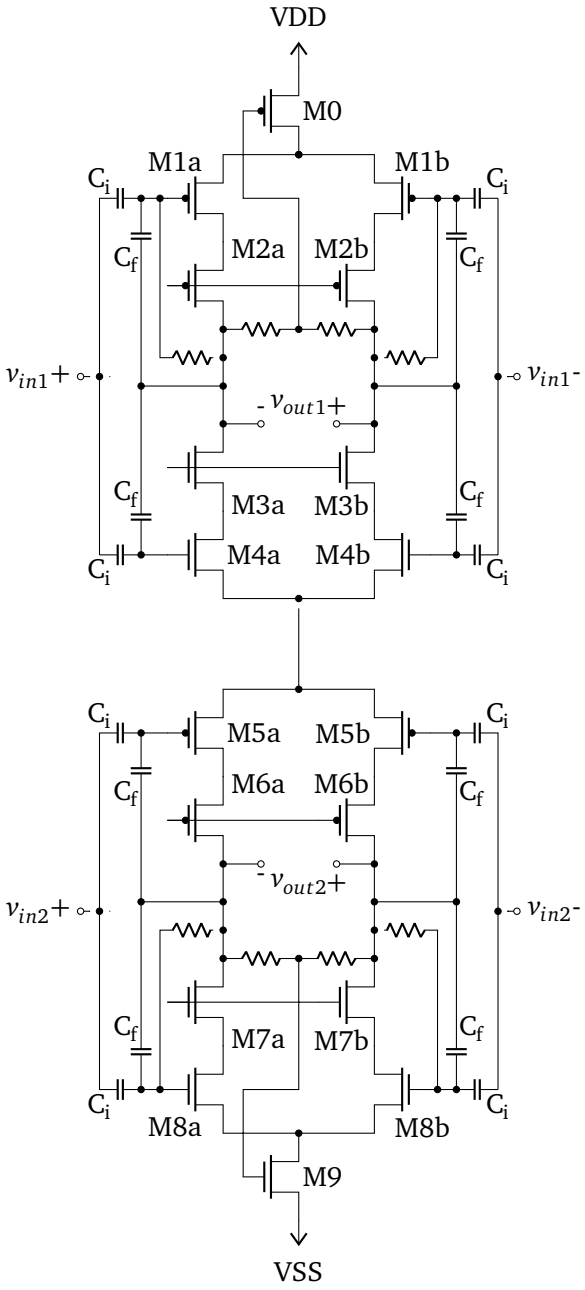


Figure 3.3: The basic structure of the stacked cascode inverter-based amplifier

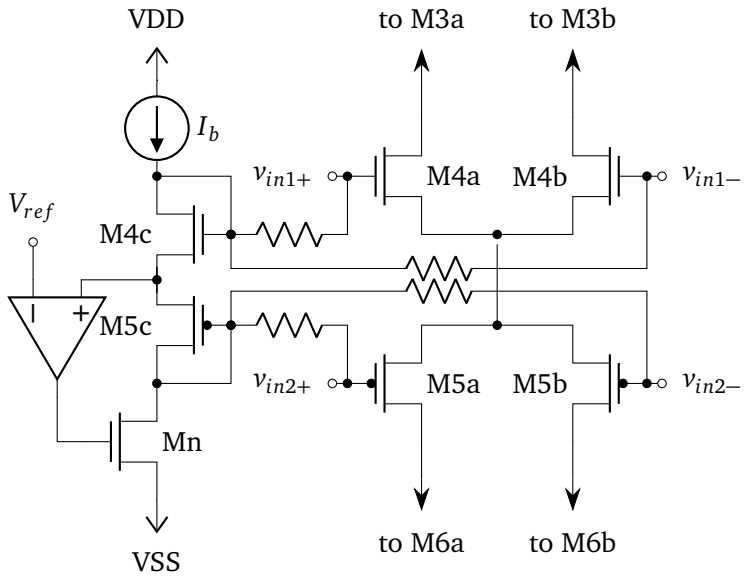


Figure 3.4: The current mirror biasing mechanism of the M4 and M5 in close-up

the gain or serve two channels simultaneously. But this topology introduces an additional challenge to the design due to fewer tail transistors per amplifier in the stack.

3.2.1 Current source

Both of the tail transistors of this design, M0 and M9, serve as common-mode feedback devices. To set the bias current another transistor is necessary. The simple solution would be to place a transistor in between the two channels to act as a common mode feedback amplifier and then use one of the tail transistors as a current source. But this would increase the minimum supply voltage so a more optimal solution is to instead use the central input transistors as current sources as well as inputs. Their DC-bias can be set by a current mirror structure as in Fig. 3.4, separated by a high resistance. The sources are all at virtual ground which is at a predefined voltage and can therefore be set by a reference voltage and a feedback amplifier controlling the source node of the diode-connected part of the current mirrors. Being able to finely control this node's voltage makes the amplifier more reliable as well - the central node in the amplifier controls the V_{gs} of both of the central differential pairs. But the V_{gs} of the pairs are affected inversely by changes in the source node which has the effect that small changes in it together with the effect on one pair can quickly push one of the pairs out of its operating region. This configuration prevents that from happening because their gate bias is set by diode-connected devices and the source node is controlled by a feedback amplifier.

The only downside is that this requires separating each of the inverter inputs.

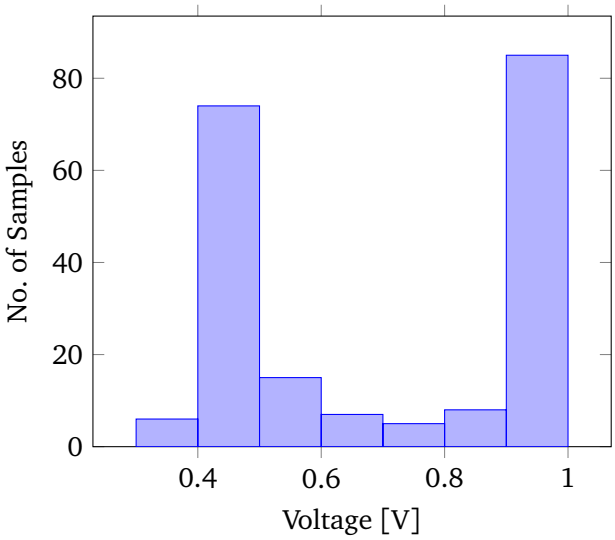


Figure 3.5: The distribution of 200 samples of voltages at the output node of a cascode inverter-based amplifier when simulated with mismatch. Under ideal conditions the node voltage is expected to be 700 mV

They are still capacitively connected to the same input signal, but each input requires its own capacitor instead of one for both inputs together. Thus this solution results in an increased area consumption.

3.3 Process and Mismatch Robustness

To ensure the performance of the physical amplifier corresponds to that of the ideal simulation adjustments have to be included to account for the effects of process and mismatch variation. Due to the high optimization for efficiency in this design, many of the transistor's operating points are highly sensitive to small changes and thus it is important to minimize the effect of variations on these.

3.3.1 Negative DC-feedback

The addition of the cascodes to an inverter-based amplifier has a side effect not seen in ideal simulations which becomes clear when adding stochastic mismatch to the simulations: The resulting voltages at the output nodes are not distributed in a bell curve but instead in the two topped curve seen in Fig. 3.5. Closer inspection of this phenomenon is necessary to fully explain it, but likely the cascode configuration and its high resistance lead to even tiny differences in bias current having catastrophic effects on the node voltage and the operating points of the transistors. So when introducing mismatch a cascode inverter-based amplifier without any feedback mechanisms nearly always drives some of its cascode transistors into the triode region to achieve a stable operating point. The common-mode feedback

provided by the tail transistor is not sufficient to prevent this phenomenon, since it depends on differences in bias current between the two branches each branch diverges equally in a different direction from the ideal operating point, keeping the gate of the feedback transistor at a stable equilibrium.

To counteract this effect negative from the outputs to at least one of the input transistors can be incorporated. This way the bias current cannot diverge in an uncontrolled manner anymore and the output nodes are stable even with mismatch applied.

3.3.2 Cascode transistor biasing

To be able to run with a minimal supply voltage the V_{ds} of all transistors should be kept close to their $V_{ds,sat}$, but this makes them more susceptible to variations as they can more easily drive V_{ds} into too low regions. Critical to the V_{ds} of all amplifying transistors is the voltage at the node between each input transistor and its cascode transistor. This voltage is directly set by the gate bias voltage of the cascode transistor and its V_{gs} and the largest source of voltage variations at this node are process variations in the threshold voltage. These can be mitigated by using the same type of transistor in a diode configuration to set the bias voltage as done in Fig. 3.6. Now any change in threshold voltage is also applied to the transistor creating the bias voltage so that the bias voltage adjusts and the source voltage of the cascode transistor stays stable.

However, more stable solutions are possible as well. The former solution does little against other variations. The biasing transistor will have a very different size than the cascode transistors and is therefore affected differently by many variations. To improve robustness further a replica can be used with a feedback amplifier and a reference voltage that controls the source voltage of the replica like done for the current mirror gate biasing of the central input transistors in Fig. 3.4. This has the advantages that the source voltage of the bias transistor is very stable and process variations affect the bias transistor in the same way as the cascode transistors leading to their sources being at the same voltage as the bias transistor when neglecting the effects of random mismatch.

Table 3.1: Voltage variations at the source node of the cascode transistors at 100 mV V_{ds} with different biasing methods

Node	Biasing method	Std. dev. [mV]
M2 source	Replica	1.78
	Simple diode connection	11.33
M3 source	Replica	1.04
	Simple diode connection	7.86

Table 3.1 shows that this produces even more robust results than the first solution, but also that with both solutions the voltage variation is within reasonable limits. Due to the very high gain of the design somewhat non-ideal operating

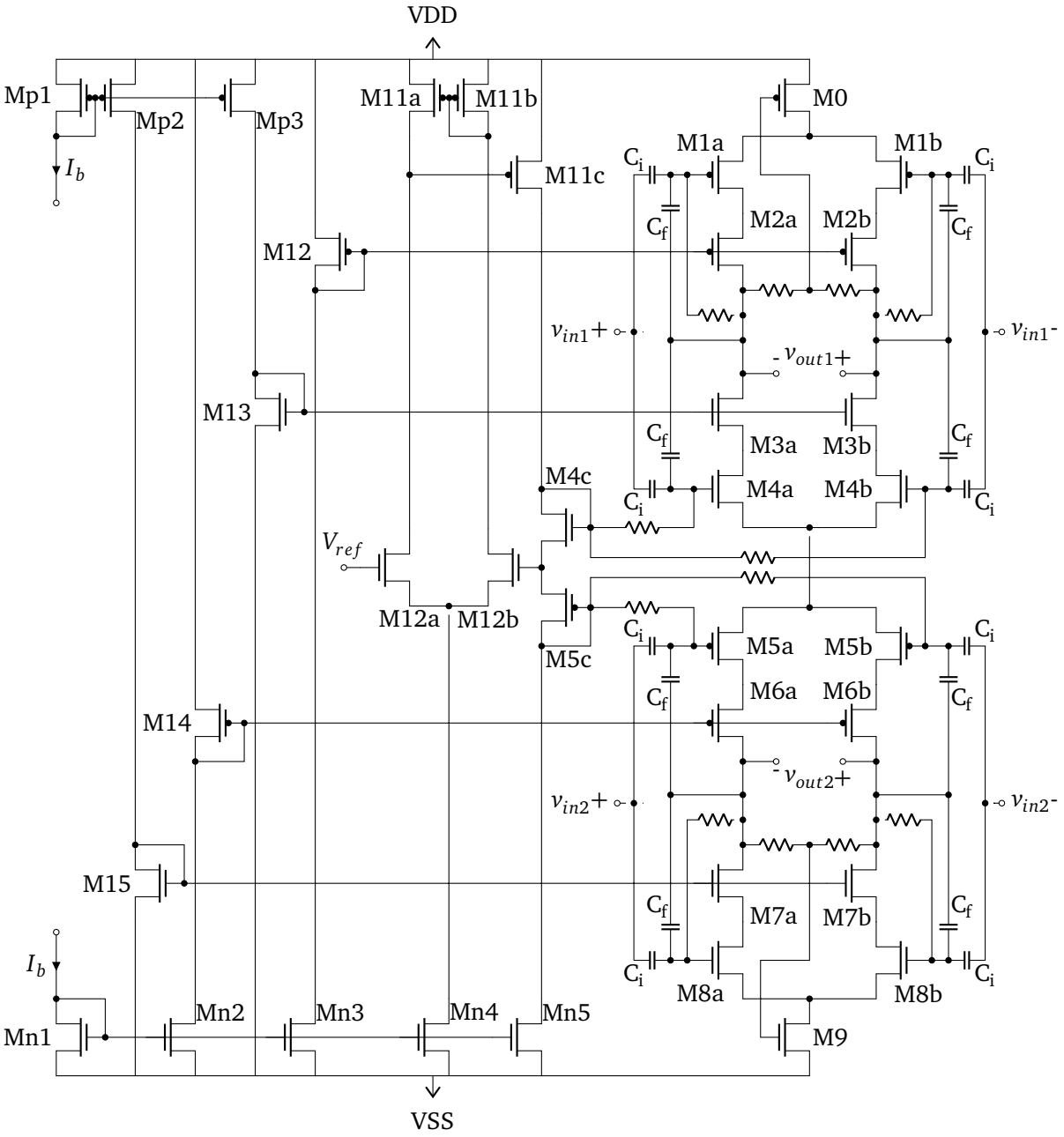


Figure 3.6: The complete double stacked cascode inverter-based amplifier

points can be tolerated - the minimum closed-loop gains achieved only differ by 0.18 dB. So while the replica biasing works better it is not a significant change and the replica bias actually has significant drawbacks: in addition to a higher total bias current due to the feedback amplifier, the feedback amplifier is also reliant on an exact and stable reference voltage. Providing one for each cascode transistor pair, likely using a bandgap reference, would add substantial power consumption and complexity to the final system and is therefore not justified.

3.3.3 Common-mode feedback

Equally as important are the output node voltages which are set by the V_{gs} of the tail transistors. This V_{gs} is equal to $V_{ds,tail} + V_{ds,input} + V_{ds,cascode}$. If the amplifier would be designed to use the minimum supply voltage possible, i.e. $10 * V_{ds,sat}$ this would result in the tail transistors operating in weak inversion or medium inversion with a low V_{ds} margin for variations. To have a stable gate voltage, however, it is preferable to operate in strong inversion with a larger margin. Therefore the supply voltage should be raised to accommodate enough voltage headroom to set the DC output voltage so that the tail transistors can operate in strong inversion and allocate this additional voltage headroom so that its V_{ds} is reliably above $V_{ds,sat}$. This guarantees a much more stable gain, albeit at the cost of a higher supply voltage.

3.4 Supply voltage

The supply voltage of the full design is then characterized by the saturation voltage needed for the central transistors and the bias voltage needed for the tail transistors to be in strong inversion and is calculated as

$$VDD = 4 * V_{ds,sat} + (V_{tn} + 100\,\text{mV}) + \left(V_{tp} + 100\,\text{mV}\right) \tag{3.9}$$

The voltage over the outer cascode pairs and the tail transistors does not influence the minimum supply voltage because their saturation voltage is lower than the V_{gs} of the tail transistors which covers the same nodes in the circuit.

3.5 Resistances

To separate the DC gate biases from the rest of the circuit high resistances are necessary. But these resistances introduce new challenges.

The resistances at the gates of the input transistors create a high-pass filter together with the input capacitance that creates a lower cutoff frequency for the amplifier's bandwidth at the frequency

$$f_{c,l} = \frac{1}{2\pi C_{in} R} \tag{3.10}$$

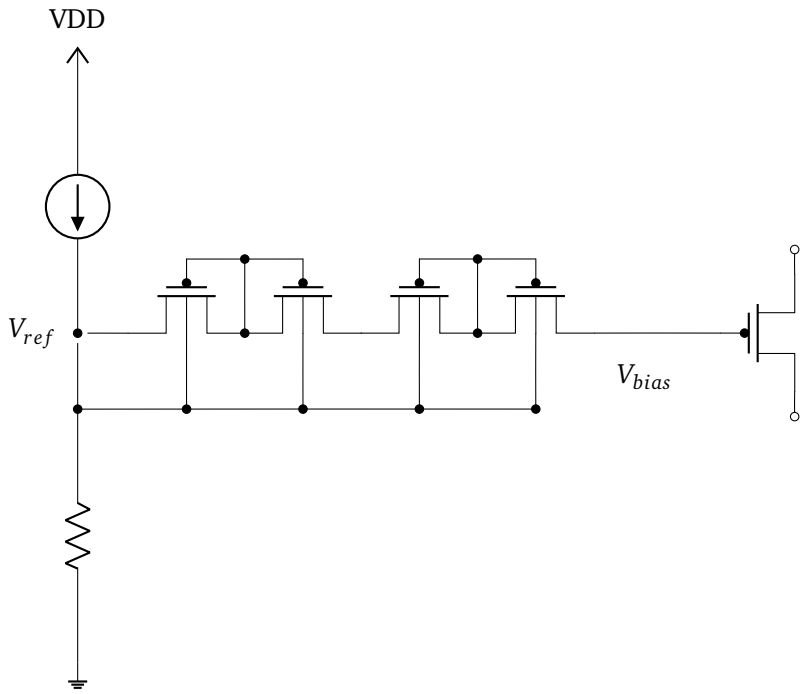


Figure 3.7: The pseudo resistances used in this design

where R corresponds to the resistances R1, R4, R5, and R8. As explained in Section 2.4, the amplifier's lower cutoff frequency should be below 1 Hz, so it follows that both the input capacitance and resistance should be large. However, the input capacitance is responsible for the majority of the amplifier's area, and therefore having a very large input capacitance is infeasible. Thus the feedback resistance has to be extremely large.

Building these out of actual resistors would be impractical due to their large size and therefore it is recommended to use pseudo-resistors instead.

The resistances used in this design are shown in Fig. 3.7. They are made up of two pseudo-resistors in series with pseudo-resistor consisting of two diode-connected pmos transistors connected to each other at their drains. This way at least one of the two transistors is always in the cutoff region and creating a large resistance.

Their long length creates a significant amount of gate-to-source capacitance, though, and a single pseudo-resistor would exhibit half of that capacitance in parallel with its resistance. This is detrimental, especially in the resistances used as negative DC feedback as that capacitance adds to the parasitic capacitance between the input transistor's gate and drain. By using two pseudo-resistors in series this capacitance is halved once more.

Due to the large resistance, even small currents create a significant voltage drop over the resistance which is unwanted. Thus all bulks of the transistors are

connected to the same node, V_{ref} which is created by a current and provides the voltage that serves gate bias V_{bias}. This way the reverse leakage current through the diode built by the border between the n-well and the p-well is fed directly by the current creating V_{ref}. This current is a negligible fraction of the bias current but if each bulk was connected to the source terminal of the transistor parts of this leakage current would flow through the resistor creating a voltage drop. So this topology helps minimize the voltage drop over the resistance.

Chapter 4

Implementation

4.1 Transistor sizing and operating points

4.1.1 Tail transistors

Table 4.1: Sizes and operating point of the tail transistors

	W/L (µm/µm)	V_{eff} (mV)	$V_{ds,sat}$ (mV)	V_{ds} (mV)
M0	10/12	−119.1	−120	−225
M9	10/38	102	107	250

Table 4.1 shows the size and operating point of the tail transistors M0 and M9. The sole function of these is to set the common-mode feedback of the output nodes via their gate voltage. The voltage has to be especially stable, so they should be safely in the saturation region and strong inversion. To satisfy this they were sized to be as large as possible without impacting the total amplifier area significantly to better tolerate variations.

Their V_{ds} is set by the V_{gs} of M1 and M8 and aimed to be at least 100 mV above $V_{ds,sat}$ to avoid going into the triode region because of small variations.

4.1.2 Input transistors

Table 4.2: Sizes and transconductances of the input transistors

	W/L (µm/µm)	g_m (µS)	g_{ds} (µS)
M1	100.8/0.4	6.983	0.1622
M4	150.08/0.5	7.512	0.189
M5	80/0.6	6.967	0.1795
M8	199.68/0.4	7.203	0.1503

The sizes of all input transistors are seen in Table 4.2, but the reasons for their sizing are slightly different for the outer input transistors and the inner ones as they serve different secondary functions next to being input transistors. They are, however, all operating in weak inversion for optimum current efficiency.

Outer input transistors

The outer input transistors, M1 and M8 have their gate voltage set by the output common-mode voltage, thus their V_{gs} directly controls the voltage at their source node, controlling the V_{ds} of both, the tail transistors and the outer input transistors. Thus the geometry of these transistors has to be set so that all these can be in saturation simultaneously. The magnitude of V_{ds} of these transistors is dictated by the tail transistor and is calculated $(|V_{gs,0/9}| - (|V_{ds,sat,0/9}| + 100\,\text{mV}))/2$ as the voltage headroom remaining after the tail transistors are well within saturation is simply divided equally among the transistors of the outer cascode. This comes out as roughly 115 mV.

Otherwise, it also controls the g_{ds} of the transistors which is a part of the output resistance, but as explained in Section 3.1.1 keeping the total size of the amplifier low is more important for gain and noise efficiency. When increasing the length to improve the output resistance the resulting parasitic capacitance, therefore, has to be monitored closely.

Current mirror input transistors

The central input transistors, M4 and M5, are also current mirrors. So these are three and not just two transistors which should all be the same size for the sake of matching during layout. Technically, it is possible to fashion them from unit transistors and make the differential pair's size a multiple of that of the diode-connected transistor, M4c and M5c. But when optimizing as far as possible and fashioning the layout interdigitized as explained in Section 4.2.1 this can be challenging to achieve.

The priorities when determining the size of these are the same as those for the other input transistors pertaining to g_{ds} and the parasitic capacitance. But in contrast to these the voltage at the sources of the current mirrors is fixed and set by a reference voltage, though, while the gate voltage is independent. Also, V_th is significantly higher for these transistors because of a higher source voltage causing a stronger body effect. The main concern is, therefore, to set W/L high to keep V_{gs} low enough for the current mirror setting the bias current for M4c and M5c to be in saturation. The V_{ds} of these transistors is not limited by any outer factor apart from the supply voltage and has been set as 100 mV. Not for M4c and M5c though, since these are diode-connected.

Flicker noise

In addition to these concerns, both nmos input transistors should be as large as possible to reduce their flicker noise as explained in Section 2.1.3. But in practice, there is not much room to increase their size due to all other constraints. M4 and M5 are less constrained in size than M1 and M8 since their V_{gs} is less important to the overall circuit, but even their maximum size is very limited by the parasitic capacitance.

4.1.3 Cascode transistors

Table 4.3: Sizes, overdrive voltage, and transconductances of the cascode transistors

	W/L (µm/µm)	g_m (µS)	g_{ds} (µS)
M2	32.48/1.1	6.318	0.122
M3	200.16/2	7.44	0.109
M6	100.24/2	6.593	0.108
M7	30/1	6.857	0.1045

The cascode transistors increase the gain by effectively increasing the output resistance which is dependent on both, their g_m and g_{ds} as laid out in (3.1). Thus the main concern for these is to maximize both. In weak inversion g_m is independent of geometry, so the main concern is achieving a high small-signal resistance without being significantly bigger than other transistors.

In addition, their V_{gs} controls the voltage at their source node which is crucial to the operation of the amplifier. However, the gate voltage is independent and set as needed by the biasing transistors M12-15. The gate voltage can introduce when instability in the biasing point when it becomes too high, though. This can happen especially to M3 and M6 because their V_{th} is increased due to the body effect. The gate voltage should not be so high that the current mirror providing the bias current to their respective biasing transistor has a V_{ds} too low to be kept in saturation since its drain is directly connected to the cascode transistor's gate. The V_{ds} used for the cascode transistor is the same as for their corresponding input transistor. Table 4.3 shows the geometry and transconductances of the cascode transistors.

4.1.4 Cascode biasing transistors

The diode-connected cascode biasing transistors only serve to set a stable gate bias for the cascode transistors that also adjusts with process variation in the V_th of the transistors. Therefore, the only concern for their sizing is to set the voltage as needed while remaining a reasonable size. A larger size of both, W and L, means they are less susceptible to variations, but due to the large V_{gs} needed of especially

Table 4.4: Sizes of the cascode bias transistors

	W/L (µm/µm)
M12	0.5/6.5
M13	0.3/86
M14	0.5/33.7
M15	0.5/25

M13 and M14 they tend to be large transistors already anyway, so this can take a lower priority compared to minimizing their total size. Their final sizes are shown in Table 4.4.

4.1.5 Current mirrors and feedback amplifier

Table 4.5: Sizes of the current mirrors and the feedback amplifier

	W/L (µm/µm)
Mp	3.6/4
Mn	2/10
Mn4	4/10
M11	0.5/33.7
M12	0.5/25

The circuit contains two current mirrors, Mn1-5 and Mp1-3, both set currents of the same magnitude but one is made up of nmos transistors and one of pmos transistors. Both nmos and pmos are necessary to have an as simple setup for the biasing of the cascode transistors as possible. The feedback amplifier made up of M11 and M12 sets the voltage at the central node of the main amplifier by controlling the voltage at the source of M4 and M5.

Therefore these structures only need to provide as stable biasing as possible and little else, so they are quite simple and their sizes, noted in Table 4.5, are chosen to keep them in strong inversion and saturation.

4.2 Layout

For the physical layout matching the input and cascode transistor pairs and little parasitic input capacitance were prioritized over area usage to ensure optimal functioning of the amplifier. The full layout is displayed in Fig. 4.1. All matching transistors are interdigitized to improve matching. The layout is similar to the schematic in Fig. 3.6 as M0 is on top with the amplifying transistors below it descending from M1-M8 ending in M9. The feedback amplifier transistors M11

and M12 are rotated by 90 degree to take up less space and fit vertically on the side of the other transistors. The current mirrors and pseudo-resistances are distributed throughout the leftover free space. The amplifier layout uses an area of $0.026\,\mathrm{mm}^2$.

4.2.1 Current mirror-differential pair

The most complex part of the layout are the center pairs of input transistors. They simultaneously serve as differential pairs and as current mirrors setting the bias current of the whole amplifier. However, their function as a differential pair is paramount as the amplifier functions well even when the current deviates from the ideal bias current by 20 %.

Therefore, the a and b input transistors making up the differential pair are laid out interdigitized as a normal differential pair with one axis of symmetry and overlapping source terminals, but not wholly enclosed by a guard ring. Instead of one dummy transistor on each end there are two which then transition to the gate biasing transistors c. This whole structure is then enclosed by another dummy transistor and by the guard ring. Fig. 4.2 shows this structure with the example of the M5 transistors.

The two dummy transistors instead of one serve to better isolate the sources of the input from the biasing transistors. These dummy transistors' gates and their drains should be connected to VDD/VSS by connecting them to the guard ring for the same reason. But this means that the metal connection layout on the top and bottom on the cannot be the same for all transistors as the connectinos of the dummy transistors break it up. Since the matching of the input transistors takes precedence over the current mirror these inconsistencies are put closer to the biasing transistors than to the input transistors.

4.2.2 Limitations

The development kit provided for the technology used does not calculate the gate resistance of a transistor. Thus no statement can be made about the effect of transistor finger lengths on the circuit. This also means the finger lengths have been chosen rather with geometry in mind than gate resistance.

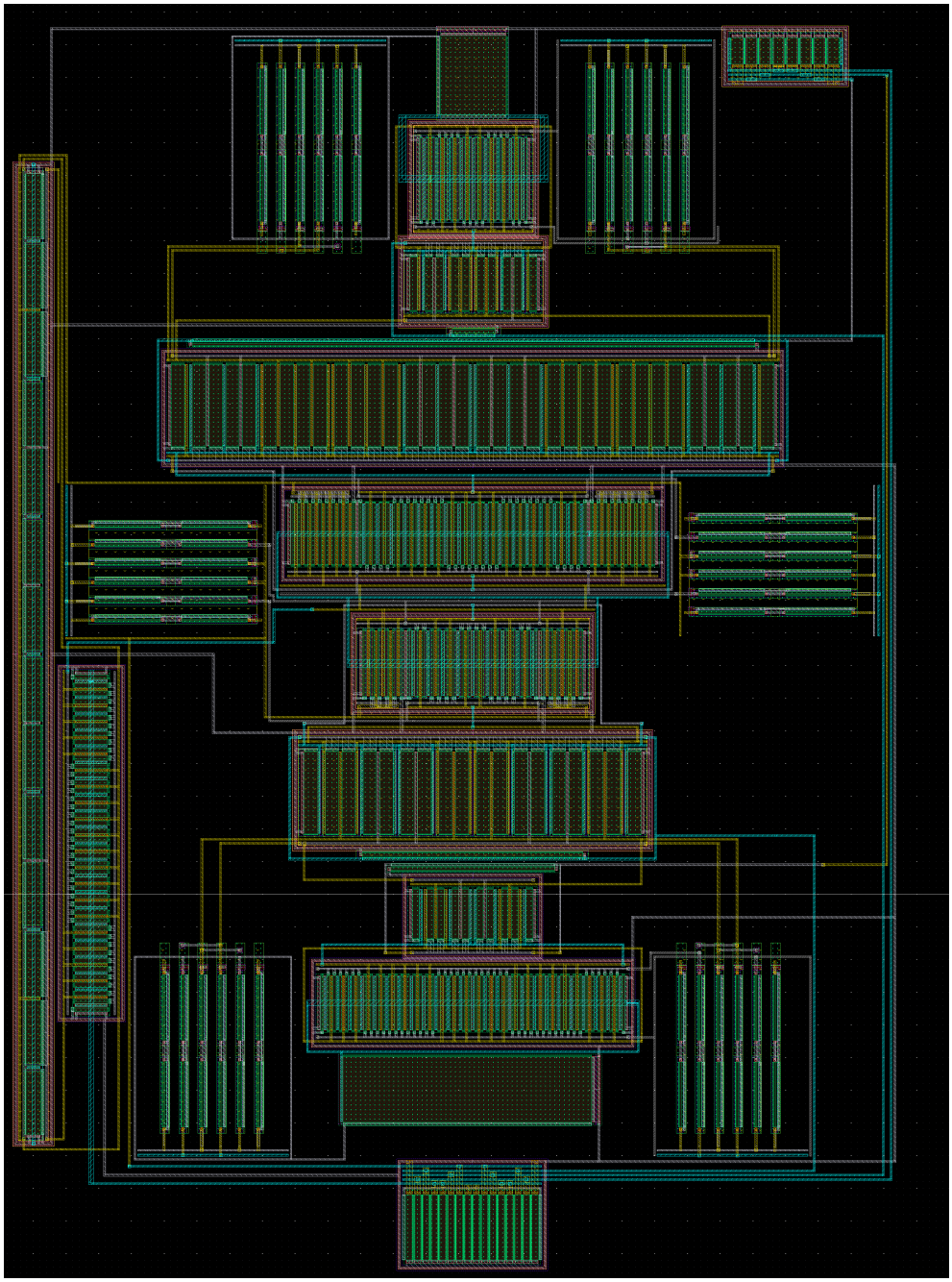


Figure 4.1: The layout of the complete design, without capacitors.

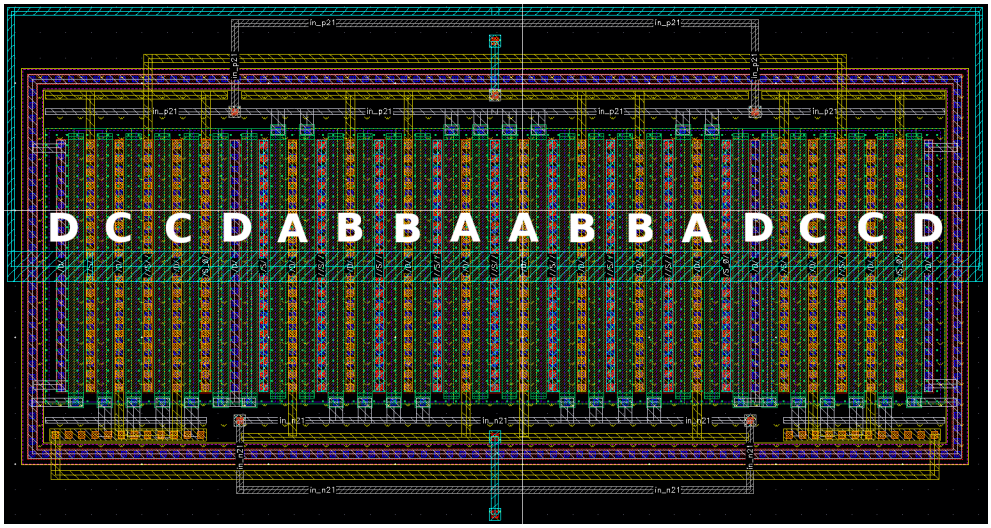


Figure 4.2: The layout of the M5 transistors. The letters A, B, and C correspond to the transistor of the same letter in the layout while D denotes the dummy transistors. Each transistor is made up of 4 equal parts distributed interdigitized along an axis of symmetry

Chapter 5

Simulation results

This chapter presents the performance of the proposed amplifier in Fig. 3.6 under different stimuli and simulations. The test bench code used for the simulation can be found in the appendix in Section A. It supplies the reference voltage, bias current, and supply voltage. The input signal is configured in such a way that each input receives a separate signal which can nonetheless be traced to a single source via voltage-controlled voltage sources and the same is used to collect all output signals to a single differential signal. This way the noise of the whole amplifier can be easily evaluated without side effects. The circuit has been simulated in ideal situations and with statistical models of process variation and mismatch in Monte Carlo simulations. All components of the amplifier, including the input and feedback capacitances, had these models of variations applied to them for these simulations and each Monte Carlo simulation encompassed 200 iterations.

5.1 Results

Table 5.1: Performance of the stacked cascode inverter-based amplifier under ideal conditions

	channel 1		channel 2	
	open loop	closed loop	open loop	closed loop
Supply Voltage (V)	1.35			
Current draw/channel (µA)	0.25		0.25	
Power/channel (nW	337.5		337.5	
Gain (dB)	69.5	39.61	70.03	39.63
Bandwidth (Hz)	700	2.361 to 11 900	694	2.507 to 12 250
Phase margin (°)	–	92.35	–	92.27

In Table 5.1 the nominal performance of the design at the typical process corner and room temperature are presented. Channel 1 refers to the top channel made up of M1-M4 and the bottom channel of M5-M8 is channel 2. It indeed is

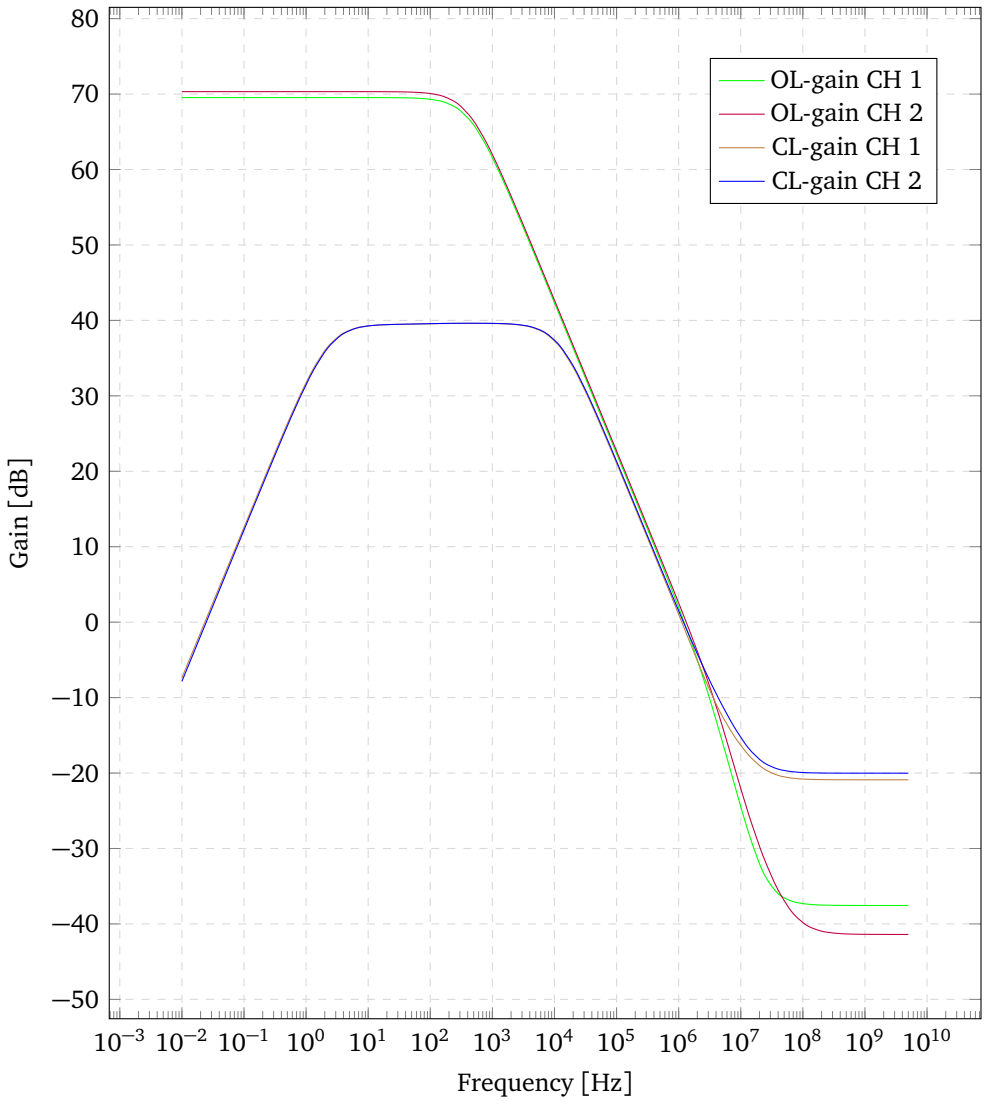


Figure 5.1: The open- and closed-loop gain of both channels

able to run while using minimal power, at only 337.5 nW per channel. The open-loop small-signal gain, which can also be seen in Fig. 5.1, is as expected high and at over 69.5 dB enough to reliably produce a closed-loop gain of 40 dB. It also is very close to the approximation in (3.2). These results show that the system is functioning, only the bandwidth falls short of the specification in the lower cutoff frequency, 2.5 Hz instead of 0.5 Hz.

Table 5.2: Peak Miller capacitances at of the input transistors

	Miller capacitance (pF)
M1	2.22
M4	2.813
M5	1.415
M8	3.61

Looking at just the open- and closed-loop gain is hiding the effect of the Miller capacitance of the input transistors. This capacitance is frequency-dependent as it varies with the gain, the figures for all four different capacitances are shown in Fig. 5.2 (Technically eight capacitances but they are the same on both transistors of a differential pair). Frequency-wise we can see that they are divided into two groups, the outer transistors M1 and M8 peak at 100 Hz while M4 and M5 peak at 280 Hz. This is because the peak capacitance is not actually dependent on the overall gain but the gate-to-drain gain of the specific MOSFET. Due to the topology, the outer and inner pairs have slightly different operating points as explained in Section 4.1.2. As a simplification, just the peak values are listed in Table 5.2. Compared to the input capacitance all these are significant values and overall the capacitive divider attenuates the input signal by nearly 25 %.

Looking at the gain for the different process corners in Fig. 5.3 and Fig. 5.4 it is also obvious that the gain is stable, even the worst-case still shows a closed-loop gain of 38 dB. And the lower cutoff frequency varies from 1.142 Hz to 5.579 Hz which is not optimal, but also not outright disqualifying the amplifier for EEG applications as only delta and theta waves occur at such low frequencies[20].

The results of Monte Carlo simulations with process variation and stochastic mismatch are shown in Table 5.3 and Figures 5.5 and 5.6 show some of the worst case gain results for both channels. Crosstalk is defined as the gain from the input of the other channel to the output of the specified channel.

The deviation from the ideal values for gain and the bandwidth is small. There is also little crosstalk between the channels, but the minimum PSRR and CMRR are slightly below specification. They also show high standard deviations, so while the average values are good they are not very robust. This high volatility can be traced to mismatch in the pseudo-resistors and the four central transistor pairs of the amplifier, M3-M6. All of these have a V_{ds} of only 100 mV which makes them susceptible to even slight variations in their operating point.

Table 5.3: Performance of the stacked cascode inverter-based amplifier when simulated with process and mismatch in closed loop

	Channel 1			Channel 2		
	worst	μ	σ	worst	μ	σ
Gain (dB)	38.77	39.54	0.1156	38.26	39.54	0.1559
Upper cutoff frequency (kHz)	11.08	11.96	0.3884	11.32	12.32	0.4329
Lower cutoff frequency (Hz)	4.335	2.482	0.518	4.772	2.584	0.553
CMRR @50 Hz (dB)	61.49	77.15	8.13	59.26	74.92	7.746
PSRR @50 Hz (dB)	63.27	78.58	6.135	59.18	72.14	6.132
Crosstalk (dB)	−65	−82.81	7.687	−53.77	−74.5	8.606

5.1.1 Noise performance

The focus of this design was to maximize its noise efficiency, therefore a closer look at the results of noise simulations and efficiency is taken in Table 5.4. The input-referred noise requirement of 40 nV/$\sqrt{\text{Hz}}$ is fulfilled by both channels. But this value does not give away the whole picture. The noise figures in Fig. 5.7 show that there is a significant amount of flicker noise in the circuit with a corner frequency of about 1 kHz.

The flicker noise also has a significant impact on the NEF, there is a difference of 0.157 between the NEF of the system and the NEF when excluding flicker noise.

Table 5.4: Noise efficiency of the stacked cascode inverter-based amplifier

	Channel 1	Channel 2	Channels combined
Input-referred noise (nV/$\sqrt{\text{Hz}}$)	35.03	35.76	24.71
$V_{n,i,rms}$ (µV)	6.076	6.288	4.371
NEF	1.031	1.052	1.042
PEF	1.435	1.495	1.465
NEF excl. flicker noise	0.870	0.900	0.885
PEF excl. flicker noise	1.023	1.094	1.058

5.2 Post-layout results

In Table 5.5 the results of post-layout simulations are presented. These results differ little from the pre-layout simulations, indicating that only small amounts of parasitics have been added to the circuit. The most significantly worse result is for crosstalk which increased more than tenfold. Luckily, these values are still negligible compared to the gain. The upper cutoff frequency has been diminished slightly which was to be expected when adding the effect of the parasitic capacitances at the output and in the input high-pass filter but has not been worsened significantly

Table 5.5: Post-layout results of the stacked cascode inverter-based amplifier

	channel 1			channel 2		
OL-gain (dB)	69.85			70.54		
CL-gain (dB)	39.57			39.61		
Bandwidth (Hz)	3.285 to 10 460			3.484 to 11.61		
Input-referred noise (nV/$\sqrt{\text{Hz}}$)	35.16			35.475		
$V_{n,i,rms}$ (µV)	6.014			6.195		
	worst	μ	σ	worst	μ	σ
CMRR @50 Hz (dB)	61.83	75.15	7.43	60.63	73.91	6.311
PSRR @50 Hz (dB)	56.32	71.83	6.355	54.17	67.56	6.724
Crosstalk (dB)	−35.24	−41.74	1.559	−31.9	−38.34	1.438

either. The PSRR, however, has been halved - to about 54 dB for channel 2. This is now significantly below the specification of 60 dB.

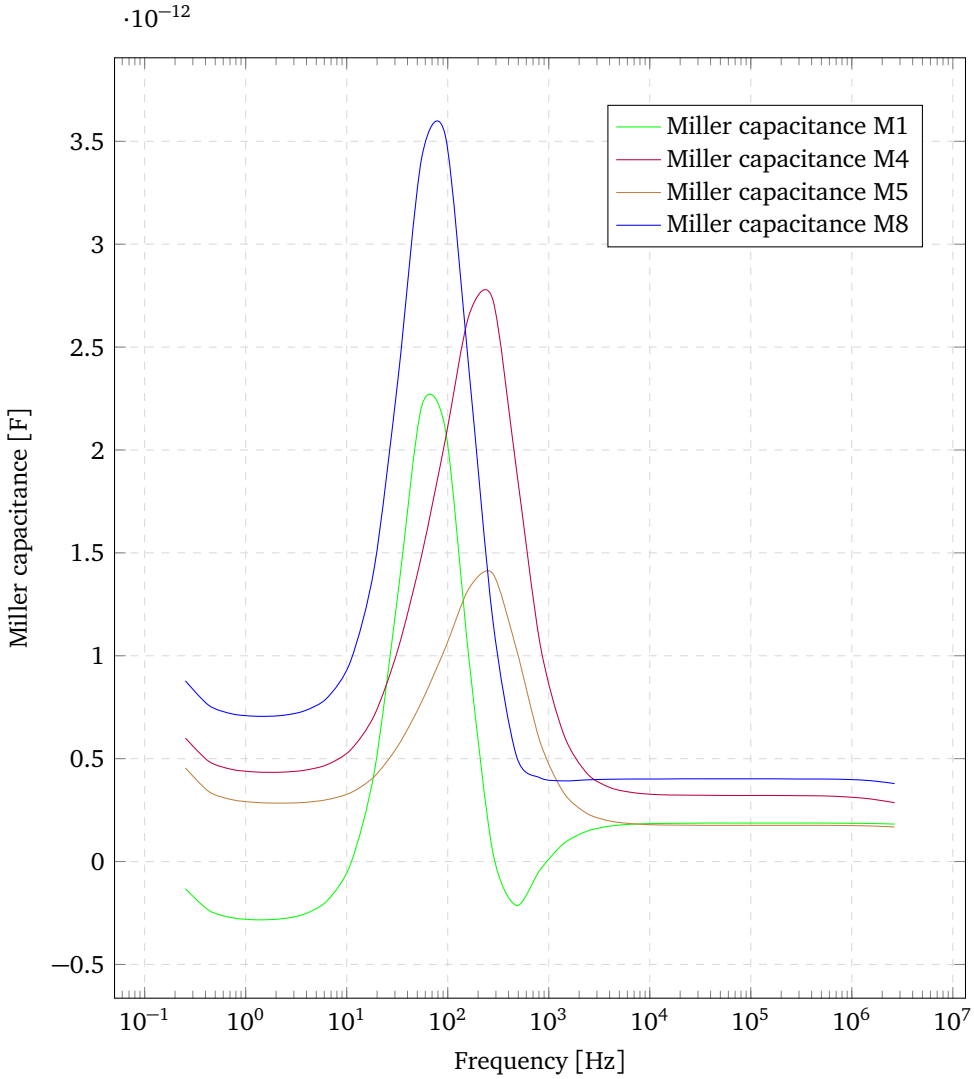


Figure 5.2: The Miller capacitances at the gate of each input transistor

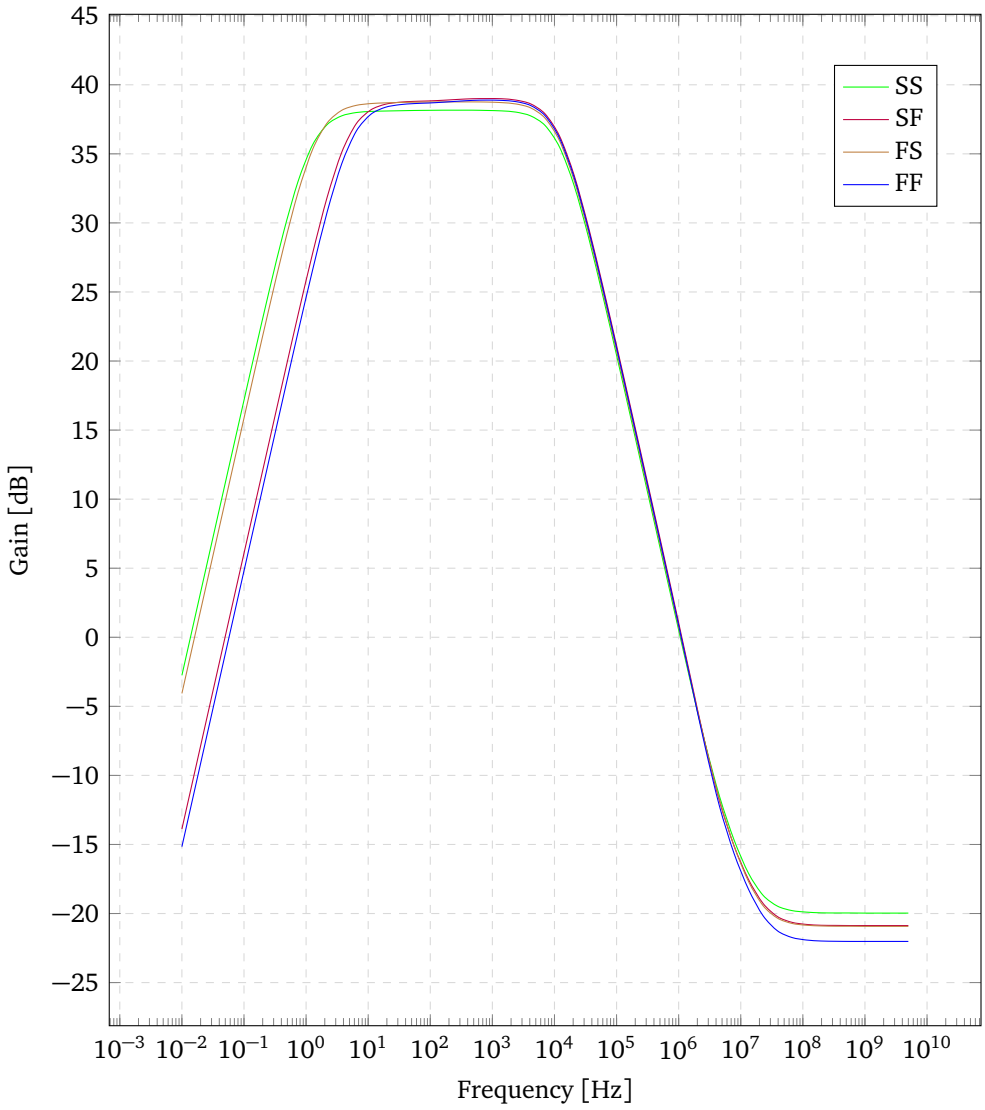


Figure 5.3: The gain of channel 1 when simulated with different process corners

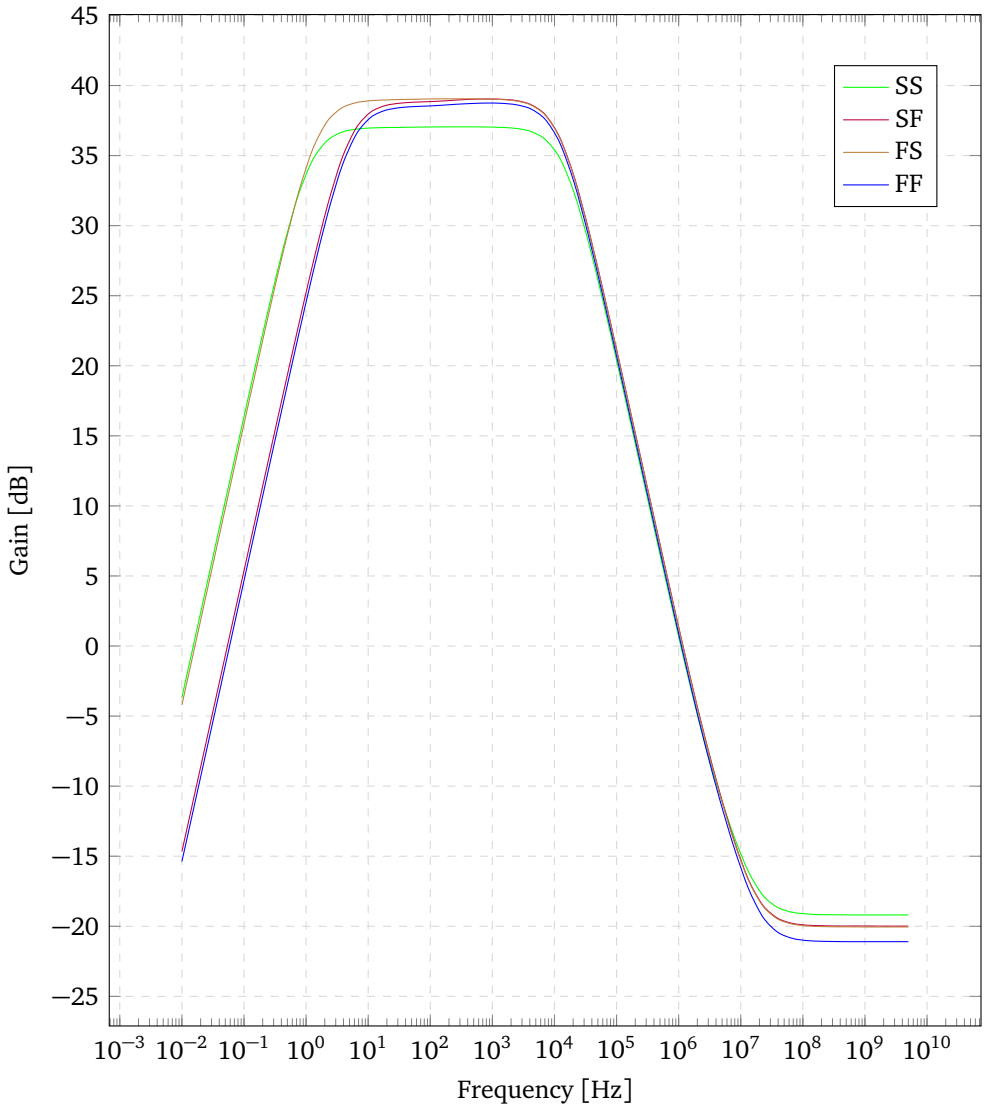


Figure 5.4: The gain of channel 2 when simulated with different process corners

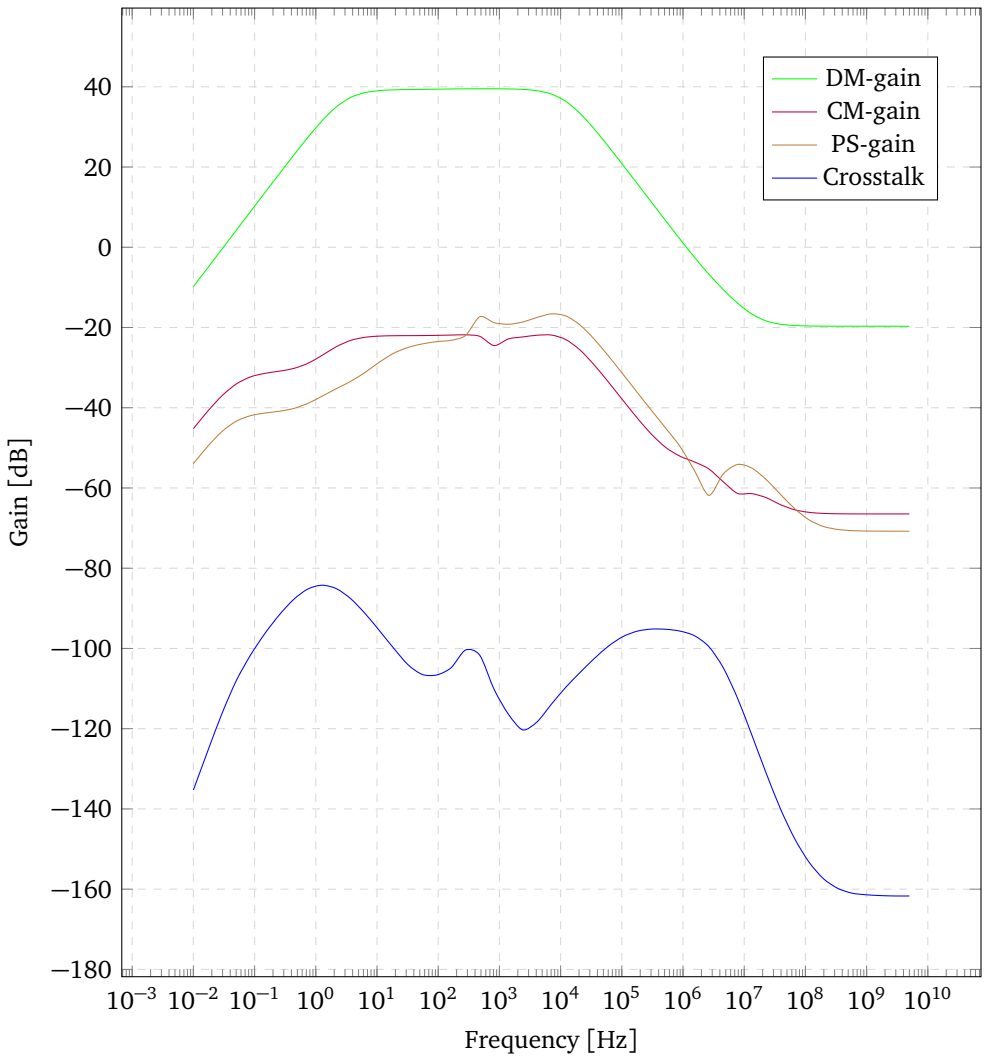


Figure 5.5: The differential, common-mode, power source and crosstalk gain of channel 1

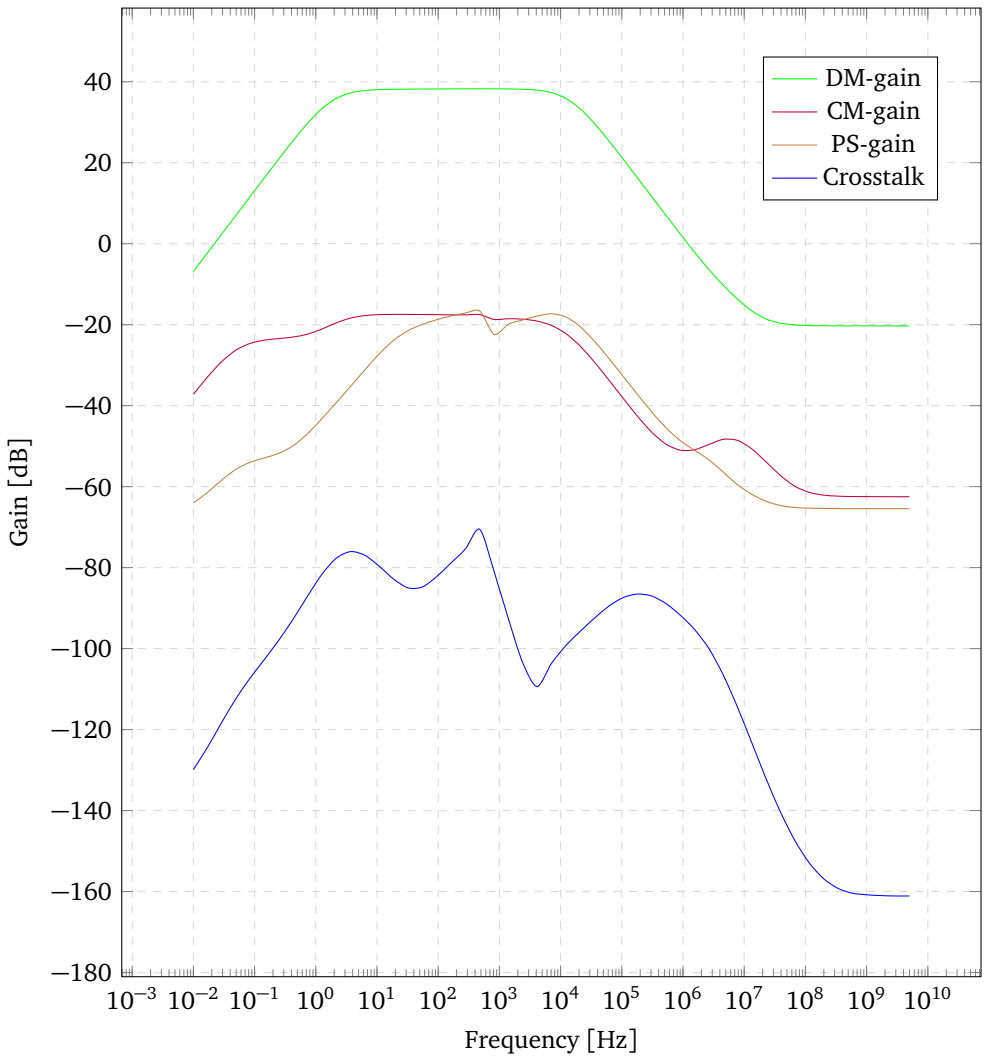


Figure 5.6: The differential, common-mode, power source and crosstalk gain of channel 2

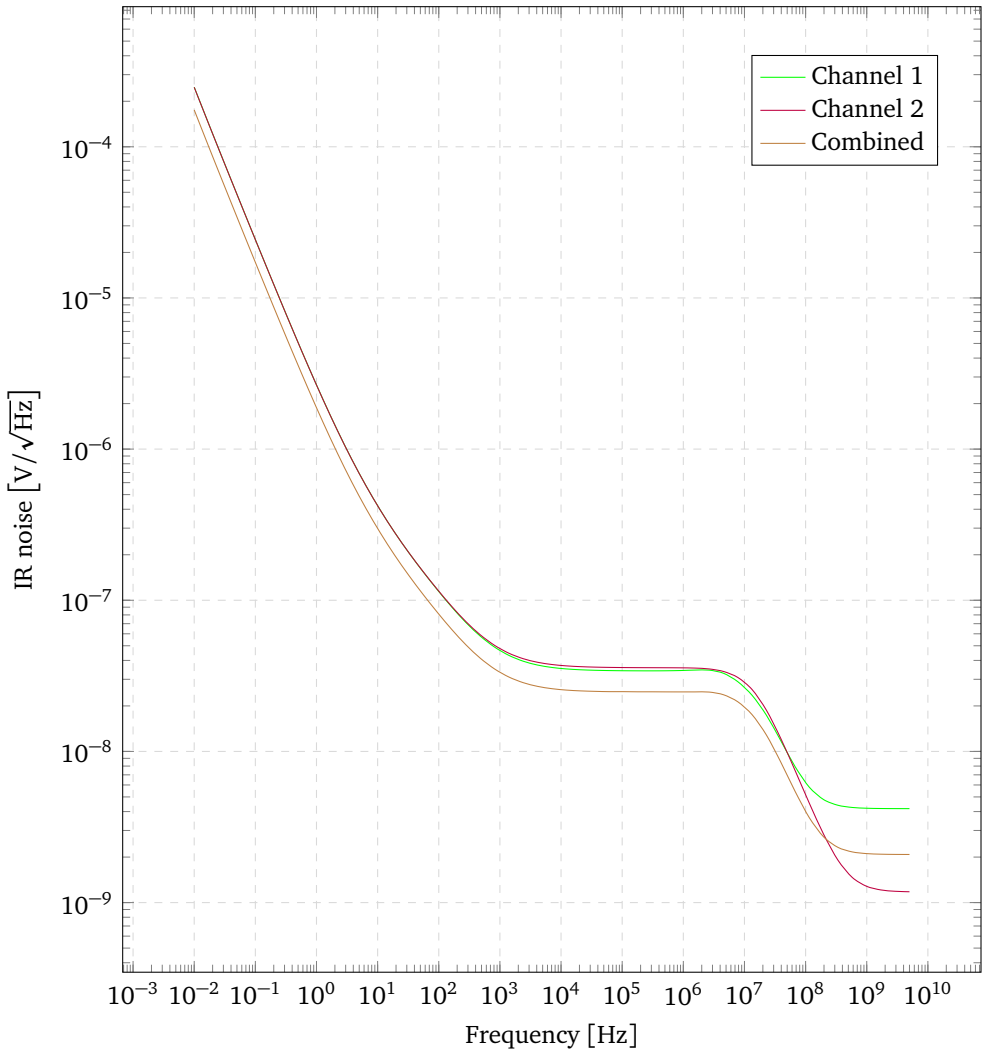


Figure 5.7: The full noise figures of both channels and the whole amplifier combined

Chapter 6

Discussion

All in all, the results show a well-functioning amplifier that is highly noise-efficient. But only when disregarding flicker noise, otherwise, while still having good NEF and functioning, the amplifier is much less noise efficient at lower frequencies.

The lower frequency brainwaves can be at around 50 μV[20] while the input-referred noise at 1 Hz is 2.5 $\mu V/\sqrt{Hz}$ meaning that the output signal will have an SNR of 26 dB. This is likely not good enough for most ADCs used in portable biosensors. Therefore the amplifier should be used together with choppers to be able to filter out the flicker noise.

The main sources of flicker noise are the nmos input transistors, M4 and M8, but increasing their size is not a viable solution because their Miller capacitance is already very large and having a significant impact on the performance. Filtering out flicker noise, therefore, also has the added side effect of enabling the use of smaller input transistors to lessen the Miller capacitance. This will enhance the gain and more importantly also diminish all input-referred noise - improvements of more than 10 % will easily be possible.

The other drawback of this design is the too high lower cutoff frequency mentioned in Section 5.1. Even with choppers 1 Hz signal cannot be measured reliably because it is outside of the bandwidth of the amplifier and the gain has fallen to 31 dB. This lower cutoff frequency occurs due to the high pass filter at the input, explained in Section 3.5. The pseudo-resistors designed for this work do not have a high enough resistance to fulfill the requirements for the lower cutoff frequency. Increasing the resistance while keeping the design otherwise the same would worsen the voltage drops seen across the pseudo resistors in the post-layout results and is therefore not a feasible solution either. Possible alternatives would be to reduce the closed-loop gain of the amplifier to 26 dB where the new lower cutoff frequency would be 0.35 Hz or combining both channels into a single channel.

The CMRR and PSRR values are just below the specified minimum value of 60 dB. This can cause problems for some systems. Therefore the requirements of all other components of an EEG need to be considered carefully before using exactly this implementation. If higher values are determined to be necessary they

can be achieved by increasing the supply voltage and the V_{ds} of especially M3-M6. In addition, special care has to be given to the layout of the pseudo-resistors to ensure they are well-matched.

Table 6.1: Comparison with other state-of-the-art neural signal recording amplifiers

	Shen2018 [8]	Han2013 [24]	Xu2021 [7]	Shad2020 [25]	This work
Technology (nm)	180	180	180	180	180
No. of channels	1	100	4	1	2
Supply (V)	1	0.45	0.9	1.2	1.35
Current/channel (µA)	0.25	1.62	1.400	3.7	0.25
Power/channel (nW)	250	730	1820	4440	337.5
Gain (dB)	25.6	52	40	40	40
Bandwidth (Hz)	4 to 10k	0.25 to 10k	0.02 to 7.2k	0.05 to 10k	3.5 to 10k
CMRR (dB)	84	73	54	94	60
PSRR (dB)	76	80	46	80	54
IR noise (nV/$\sqrt{\text{Hz}}$)	40	29	22	35	35
$V_{n,i,rms}$ (µV) (Integrated BW(Hz))	5.5 (250 to 10k)	3.2 (1 to 10k)	2.39 (1 to 50k)	2.9 (0.05 to 10k)	4.371 (3.5 to 10.5k)
NEF	1.07[1]	1.57[1]	1.28[1]	2.2[2]	1.052[1] / 0.900[3]
PEF	1.145	1.12	1.47	5.81	1.495 / 1.094

[1] Including flicker noise
[2] With chopping
[3] Ignoring flicker noise

In Table 6.1 the results are compared to other noise efficient amplifier designs. Clearly, the final NEF value is comparable to other state-of-the-art systems, arguably better, even when including flicker noise in the figure. This advantage is not gained by lower input-referred noise, but by a much lower bias current. Through this, the power consumption per channel is also comparably low. Only [8] has an equally low bias current and NEF, but also a slightly higher noise figure and lower gain. But this design also has a comparatively high supply voltage, low PSRR and CMRR values, and an insufficient lower cutoff frequency. Most other designs have a bandwidth that starts significantly lower than the specified 0.5 Hz.

Nonetheless, the aim of this design was to be highly noise efficient and use little power. That has certainly been achieved even with a relatively high amount of flicker noise. If that were to be reduced in future iterations, the amplifier will be among the most noise efficient designs with a bandwidth of several kHz developed to date.

Chapter 7

Conclusion

This work has shown that a cascode inverter-based amplifier is a feasible topology for current-efficient neural signal amplifiers which can even be used as a stacked amplifier.

The cascode inverter-based design proved to be a way of achieving very high gain even in the subthreshold region, resulting in a high current efficiency through the better g_m/I_d. The cascode structure provides a higher output resistance and lower Miller capacitance compared to a regular inverter-based design while adding a negligible amount of noise.

To able to provide a sufficient CMRR of 60 dB, however, the amplifier has to run at a supply voltage higher than just the $V_{ds,sat}$ of all devices. As the DC output voltages become too unstable otherwise.

Stacking this topology creates a challenge for the creation of the bias current which can be solved by using input transistors as current mirrors. The stacking has the expected effect of providing a second channel without having to raise the total bias current necessary, effectively halving the creating input-referred noise. The crosstalk created by this is negligible and the NEF is decreased by a factor of $\sqrt{2}$.

With a supply voltage of 1.35 V and a bias current of 0.5 µA, the amplifier can provide a closed-loop gain of 40 dB on two channels. This gives a power consumption of only 337.5 nW per channel. Simultaneously the system is also highly noise efficient with an NEF of 0.88 but only when excluding flicker noise.

With the added parasitic capacitance of the finished layout the bandwidth comes out to just above the 10 kHz but the lower cutoff frequency falls short of the specification with 3.5 Hz.

Nevertheless, in comparison with other state-of-the-art amplifiers, it is clear that this design is extremely noise efficient and uses little power per channel.

7.1 Future work

The work done for this thesis can be further improved and expanded upon in multiple ways.

First of all the design of the pseudo-resistors should be improved to be able to expand the bandwidth to encompass all possible brainwaves. This will likely not just require developing pseudo-resistors with a higher resistance, but also to further minimize the leakage currents present in the amplifier as these will cause more problems in conjunction with larger resistances.

Another future improvement on this technique that should be explored is to expand the stacking to more channels as each channel increases noise efficiency further as the overall NEF decreases by $\sqrt{N}$ compared to a single-channel where N is the number of channels stacked.

To be able to fully realize the potential of the design flicker noise has to be reduced. Thus an obvious expansion before implementing the amplifier in a real design will be to add choppers at the inputs and outputs which enable one to then filter out the created flicker noise. Otherwise, low-frequency signals will have a significantly worse SNR than high-frequency signals.

Also, this thesis has purely relied on simulations to evaluate the design of this amplifier and as such this cannot be seen as a complete evaluation of the topology. Further work will have to measure its performance on a fabricated chip before using the amplifier in a larger system. The corner, Monte Carlo, and post-layout simulations should have contributed to a realistic evaluation of the performance, but these cannot substitute any real measurements.

Bibliography

[1] N. Srinivasan, "Cognitive neuroscience of creativity: EEG based approaches," *Methods*, vol. 42, no. 1, pp. 109–116, 2007, Neurocognitive Mechanisms of Creativity: A Toolkit, ISSN: 1046-2023. DOI: https://doi.org/10.1016/j.ymeth.2006.12.008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1046202306003094.

[2] A. B. Schwartz, X. T. Cui, D. J. Weber, and D. W. Moran, "Brain-controlled interfaces: Movement restoration with neural prosthetics," *Neuron*, vol. 52, no. 1, pp. 205–220, 2006, ISSN: 0896-6273. DOI: https://doi.org/10.1016/j.neuron.2006.09.019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0896627306007264.

[3] J. Askamp and M. J. van Putten, "Mobile eeg in epilepsy," *International Journal of Psychophysiology*, vol. 91, no. 1, pp. 30–35, 2014, Towards mobile EEG, ISSN: 0167-8760. DOI: https://doi.org/10.1016/j.ijpsycho.2013.09.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167876013002523.

[4] A. J. Casson, D. C. Yates, S. J. Smith, J. S. Duncan, and E. Rodriguez-Villegas, "Wearable electroencephalography," *IEEE Engineering in Medicine and Biology Magazine*, vol. 29, no. 3, pp. 44–56, 2010. DOI: 10.1109/MEMB.2010.936545.

[5] A. J. Casson, "Wearable EEG and beyond," *Biomedical Engineering Letters*, vol. 9, no. 1, pp. 53–71, Feb. 2019, ISSN: 2093-985X. DOI: 10.1007/s13534-018-00093-6. [Online]. Available: https://doi.org/10.1007/s13534-018-00093-6.

[6] S. Mondal and D. A. Hall, "A 13.9-nA ECG amplifier achieving 0.86/0.99 NEF/PEF using AC-coupled OTA-stacking," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 2, pp. 414–425, 2020. DOI: 10.1109/JSSC.2019.2957193.

[7] Q. Xu, K. Liao, J. Yang, P. Qu, F. Luo, W. Hu, S. VanDuyne, and Z. Xiao, "An inverter-based amplifier structure for neural signal recording with an nef of 1.28 and area-per-channel of 0.06mm2," in *2021 IEEE Custom Integrated Circuits Conference (CICC)*, 2021, pp. 1–2. DOI: 10.1109/CICC51472.2021.9431491.

[8] L. Shen, N. Lu, and N. Sun, "A 1 v 0.25 μW inverter stacking amplifier with 1.07 noise efficiency factor," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 3, pp. 896–905, 2018. DOI: 10.1109/JSSC.2017.2786724.

[9] F. Silveira, D. Flandre, and P. Jespers, "A g_m/I_D based methodology for the design of cmos analog circuits and its application to the synthesis of a silicon-on-insulator micropower OTA," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 9, pp. 1314–1319, 1996. DOI: 10.1109/4.535416.

[10] H.-F. Teng, S.-L. Jang, and M. Juang, "A unified model for high-frequency current noise of MOSFETs," *Solid-State Electronics*, vol. 47, no. 11, pp. 2043–2048, 2003, ISSN: 0038-1101. DOI: https://doi.org/10.1016/S0038-1101(03)00248-X. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S003811010300248X.

[11] T. Carusone, D. Johns, and K. Martin, *Analog Integrated Circuit Design*, ser. Analog Integrated Circuit Design. Wiley, 2011, ISBN: 9780470770108. [Online]. Available: https://books.google.no/books?id=1OIJZzLvVhcC.

[12] H. Lee and S. Mohammadi, "A Subthreshold Low Phase Noise CMOS LC VCO for Ultra Low Power Applications," *IEEE Microwave and Wireless Components Letters*, vol. 17, no. 11, pp. 796–798, 2007. DOI: 10.1109/LMWC.2007.908057.

[13] Y. Nemirovsky, D. Corcos, I. Brouk, A. Nemirovsky, and S. Chaudhry, "1/f noise in advanced cmos transistors," *IEEE Instrumentation Measurement Magazine*, vol. 14, no. 1, pp. 14–22, 2011. DOI: 10.1109/MIM.2011.5704805.

[14] B. Razavi, *Design of Analog CMOS Integrated Circuits*, ser. IRWIN ELECTRONICS & COMPUTER E. McGraw-Hill Education, 2016, ISBN: 9781259255090. [Online]. Available: https://books.google.no/books?id=YXtqjwEACAAJ.

[15] M. S. J. Steyaert and W. M. C. Sansen, "A micropower low-noise monolithic instrumentation amplifier for medical purposes," *IEEE Journal of Solid-State Circuits*, vol. 22, no. 6, pp. 1163–1168, 1987. DOI: 10.1109/JSSC.1987.1052869.

[16] W. Wattanapanitch, M. Fee, and R. Sarpeshkar, "An energy-efficient micropower neural recording amplifier," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 1, no. 2, pp. 136–147, 2007. DOI: 10.1109/TBCAS.2007.907868.

[17] M. R. Miguez, J. Gak, A. Arnaud, A. R. Oliva, and P. Julián, "A Current-Reuse Biomedical Amplifier with a NEF < 1," *Analog Integr. Circuits Signal Process.*, vol. 95, no. 2, pp. 283–294, May 2018, ISSN: 0925-1030. DOI: 10.1007/s10470-018-1175-8. [Online]. Available: https://doi.org/10.1007/s10470-018-1175-8.

[18] R. Muller, S. Gambini, and J. M. Rabaey, "A 0.013 mm^2, 5 μW, DC-coupled neural signal acquisition IC with 0.5 V supply," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 1, pp. 232–243, 2012. DOI: 10.1109/JSSC.2011.2163552.

[19] B. Johnson and A. Molnar, "An orthogonal current-reuse amplifier for multi-channel sensing," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 6, pp. 1487–1496, 2013. DOI: 10.1109/JSSC.2013.2257478.

[20] G. Adelman and B. Smith, *Encyclopedia of Neuroscience*, ser. A pro scientia viva title v. 1. Elsevier, 1999, ISBN: 9780444816122. [Online]. Available: https://books.google.no/books?id=pGBqAAAAMAAJ.

[21] A. D. Legatt, J. Arezzo, and H. G. Vaughan, "Averaged multiple unit activity as an estimate of phasic changes in local neuronal activity: Effects of volume-conducted potentials," *Journal of Neuroscience Methods*, vol. 2, no. 2, pp. 203–217, 1980, ISSN: 0165-0270. DOI: https://doi.org/10.1016/0165-0270(80)90061-8. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0165027080900618.

[22] H. Chandrakumar and D. Marković, "An 80 mvpp linear-input range, 1.6 GΩ input impedance, low-power chopper amplifier for closed-loop neural recording that is tolerant to 650-mvpp common-mode interference," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 11, pp. 2811–2828, 2017. DOI: 10.1109/JSSC.2017.2753824.

[23] R. Spencer and M. Ghausi, *Introduction to Electronic Circuit Design*. Prentice Hall, 2003, ISBN: 9780201361834. [Online]. Available: https://books.google.no/books?id=J31GAAAAYAAJ.

[24] D. Han, Y. Zheng, R. Rajkumar, G. Dawe, and M. Je, "A 0.45 V 100-channel neural-recording IC with sub-μW/channel consumption in 0.18μm CMOS," in *2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 2013, pp. 290–291. DOI: 10.1109/ISSCC.2013.6487739.

[25] E. Shad, M. Molinas, and T. Ytterdal, "A low-power and low-noise multi-purpose chopper amplifier with high CMRR and PSRR," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, pp. 3998–4001. DOI: 10.1109/EMBC44109.2020.9176224.

Appendix A

Simulation testbenches

A.1 Closed loop testbench

A.2 Open loop testbench

A.3 Amplifier

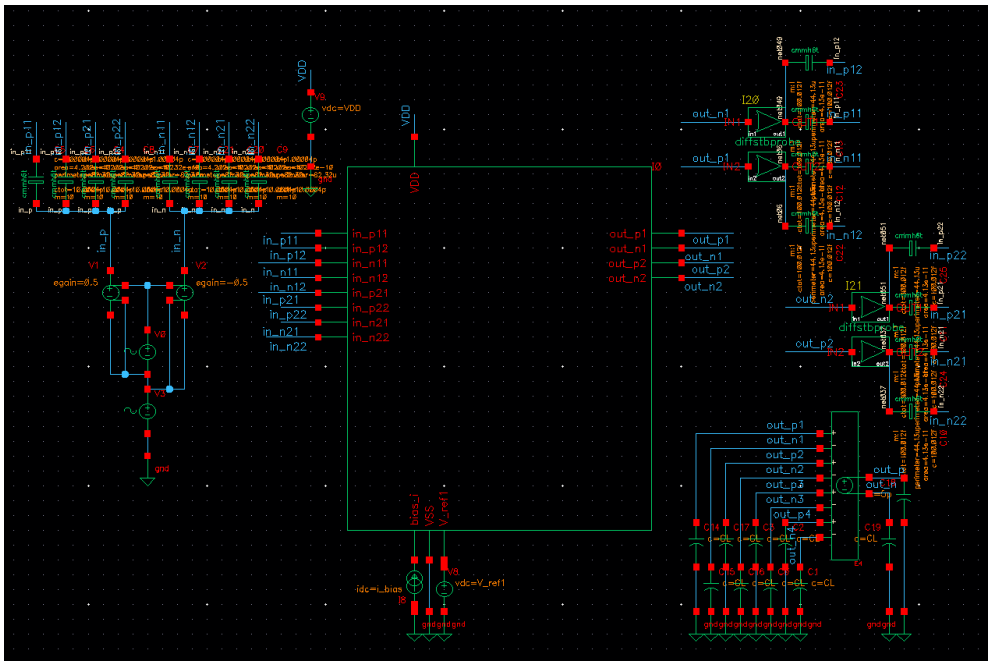


Figure A.1: The closed loop testbench

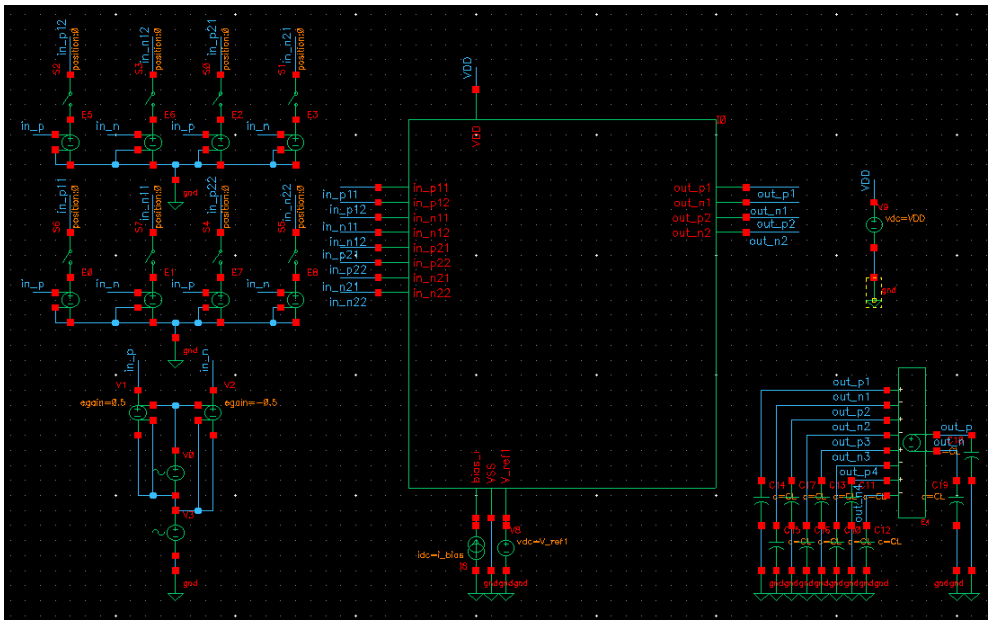


Figure A.2: The open loop testbench

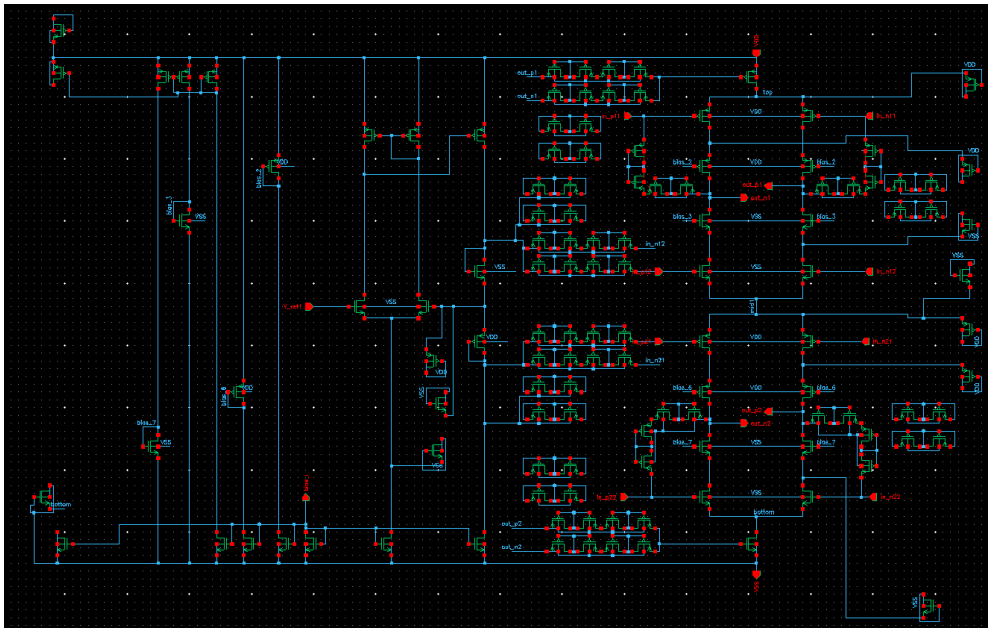
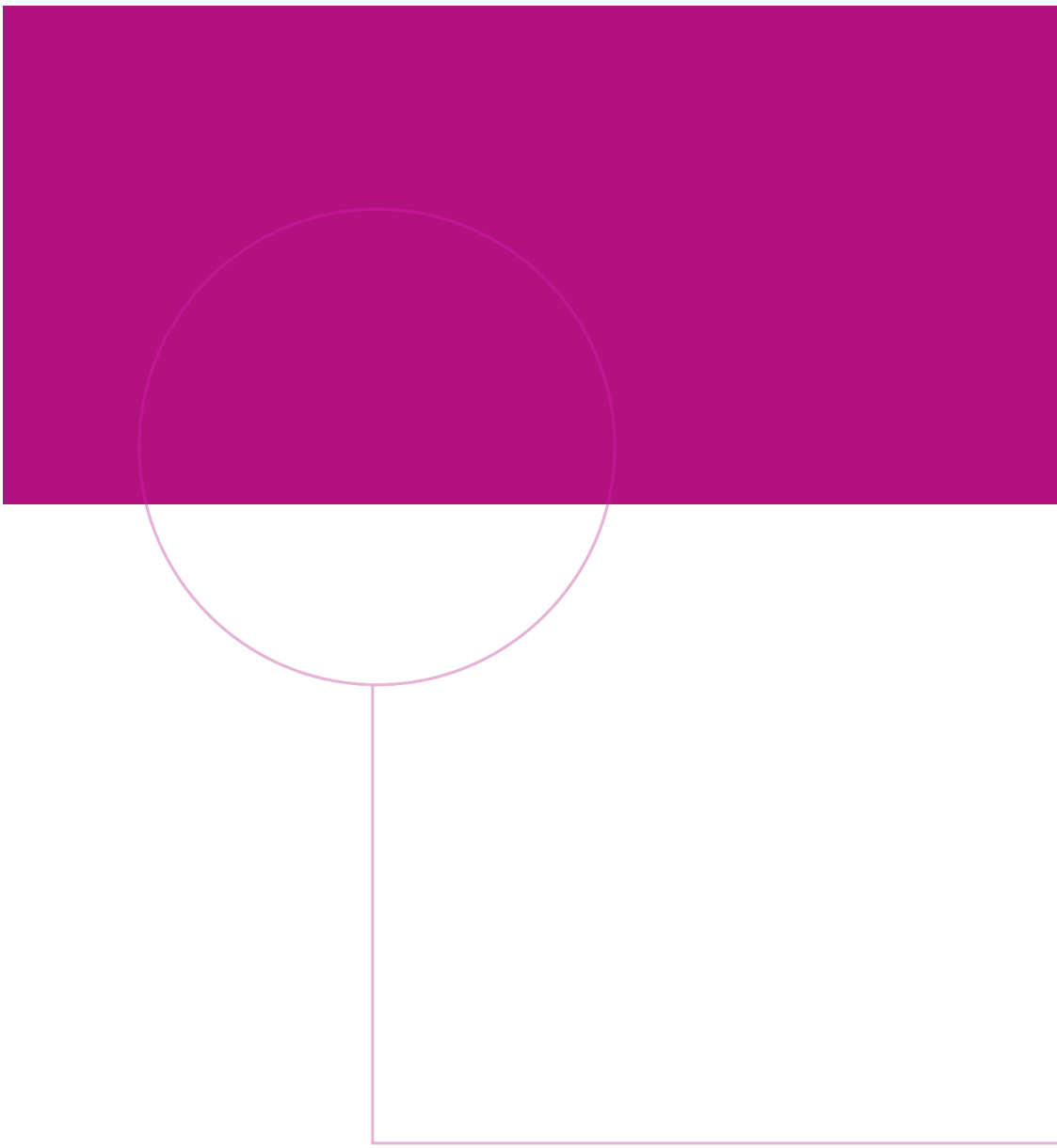


Figure A.3: The amplifier schematic used for virtuoso simulations, including layout dummy transistors



NTNU

Kunnskap for en bedre verden