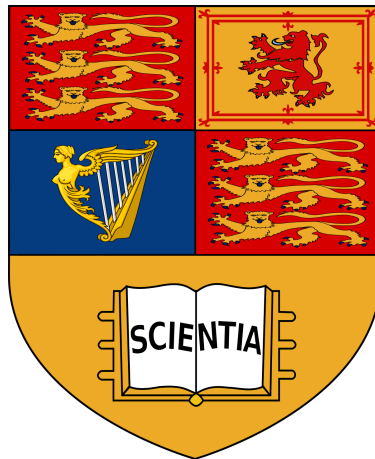Imperial College London

Department of Electrical and Electronic Engineering

Final Year Project Report 2020

Project Title:	**Designing CMOS Integrated Circuits for the Body Dust Project**
Student:	**Yaohua Zhang**
CID:	**01178478**
Course:	**4T**
Project Supervisor:	**Dr Pantelis Georgiou**
Second Marker:	**Dr Jesus Rodriguez Manzano**

Final Report Plagiarism Statement

I affirm that I have submitted, or will submit, electronic copies of my final year project report to both Blackboard and the EEE coursework submission system.

I affirm that the two copies of the report are identical.

I affirm that I have provided explicit references for all material in my Final Report which is not authored by me and represented as my own work.

Abstract

The Body Dust Project is a collaboration between Imperial College London and EPFL and presents a pioneering vision for future healthcare in which tiny pills can be ingested and be able to communicate useful bio-information wirelessly to the outside world. This Final Year Project investigated the design of potentiostats for the Body Dust Project, with particular emphasis on the inherent trade-offs when optimising the potentiostat circuit for low power dissipation, small area and high Signal-to-Noise Ratio. A figure of merit (FoM) using a weighted Euclidean distance in a 3-D space was proposed as a metric to compare op-amp performance based on its power, area and noise. A smaller FoM indicates a better performance.

The g_m/I_D methodology was analysed extensively and adopted in this Final Year Project as a systematic framework to aid the design of analogue ICs in challenging contexts such as the Body Dust Project. Existing potentiostat topologies were studied, simulated in Cadence Spectre and evaluated. The potentiostat topology proposed by Trakoolwattana and Thanachayanont was selected due to its superior performance and was the subject of further optimisation in this Final Year Project. The inherent trade-offs in Trakoolwattana and Thanachayanont's topology were analysed critically using the g_m/I_D methodology. Trakoolwattana and Thanachayanont's topology was redesigned using the g_m/I_D methodology and better results were obtained than that reported in their paper. Taking the key building block, the folded cascode op-amp, as an example, our design achieved a FoM of 0.85635, while Trakoolwattana and Thanachayanont's design had a FoM of 0.95440. This was largely due to the fact that the power dissipation had been drastically reduced from 409.641 nW to 161.674 nW, showing a 60.5% improvement.

Scaling the same topology to the more advanced TSMC 65 nm CMOS technology was attempted. Simulated results showed that the design suffered from a reduced gain as compared to that in TSMC 180 nm for a comparable power dissipation and area. Therefore, scaling to more advanced technology nodes calls for a different topology and advanced circuit design techniques.

Acknowledgements

Firstly, I would like to thank my project supervisor Dr. Pantelis Georgiou for the opportunity to work on the Body Dust project and Mr Daryl Mah for his advice and support. I cannot thank them enough for all the insights and motivation offered. This project would not have been possible without their encouragement and guidance.

Secondly, I would like to thank Dr. Sandro Carrara from EPFL for his high-level feedback during our Skype meetings.

Thirdly, I would like to thank Mr Miguel Cacho Soblechero, Mr Zeng Junming, Dr Nicolas Moser, Mr Shen Jiawei, and Mr Adil Malik for their help with Cadence and general analogue circuit design. In particular, I would like to thank Mr Michal Maslik for his stellar support as the Cadence system admin and his prompt responses. Mr Michal Maslik has been very patient and kind. I cannot thank him enough.

Lastly, I would like to thank my family and friends for their constant love and support, and for always believing in me.

List of Figures

List of Symbols

This report follows the IEEE standard for signal variable notation.

- Incremental (small-signal) quantities are represented by lowercase variable names and lowercase subscripts.

- DC quantities are represented by uppercase variable names and uppercase subscripts.

- Total quantities, which represent the sum of the DC quantity and small-signal quantity are represented by lowercase variable names and uppercase subscripts.

β	Current factor $(\mu C_{ox} W/L)$
γ	Backgate effect parameter
γ_n, γ_p	Thermal noise factor for n- channel and p- channel devices2
$\Re$	Ratio between load and input capacitances of a circuit
μ	Mobility
μ_0	Low-field mobility
ω	Angular frequency $(2\pi f)$
ω_c	Angular cutoff frequency $(2\pi f_c)$
ω_T	Angular transit frequency $(2\pi f_T)$
Ψ_S	Surface potential
ρ	Normalised transconductance efficiency
A_{intr}	Intrinsic gain
A_{v0}	Low frequency small-signal voltage gain
A_v	Small-signal voltage gain
C_C	Miller compensation capacitance
C_{gb}	Gate-to-bulk capacitance
C_{gd}	Gate-to-drain capacitance
C_{gs}	Gate-to-source capacitance

C_j	Junction capacitance
C_{ox}	Oxide capacitance per unit area
D	Diffusion constant
$DIBL$	Drain-Induced Barrier Lowering
EKV	Enz, Krumenacher and Vittoz compact model
f	Frequency in Hz
f_{co}	Flicker noise corner frequency
f_c	Cutoff frequency (-3dB frequency)
f_T	Transit frequency
g_{ds}	Output conductance
g_{mb}	Bulk transconductance
g_m	Gate transconductance
I_D	DC drain current
I_S	Specific current
IGS	Intrinsic Gain Stage
J_D	Drain current density (I_D/W)
k	Boltzmann constant
L	Gate length
N	Impurity concentration
n	Sub-threshold slope factor
q	Normalised mobile charge density
Q_i	Mobile charge density
q_S, q_D	Normalised mobile charge density at the source and drain
U_T	Thermal voltage kT/q
V_{Dsat}	Drain saturation voltage
V_{EA}	Early voltage
V_{GS}, V_{DS}	Gate and drain voltage with respect to the source (DC)
v_{gs}, v_{ds}	Incremental gate and drain voltage with respect to the source

v_{id}	Differential input voltage, AC component
V_I	DC component of input voltage
v_I	Total input voltage $v_I = V_I + v_i$
v_i	AC component of input voltage
V_{OV}	Gate overdrive voltage, $V_{GS} - V_T$
v_{sat}	Saturation velocity of mobile carriers
$V_S,\ V_G,\ V_D$	Source, gate and drain voltage with respect to bulk (DC)
V_T	Threshold voltage
V_X	DC voltage component at node x
v_x	AC voltage component at node x
W	Transistor width
WI, MI, SI	Weak, moderate and strong inversion

Contents

1 Introduction

1.1 Project Motivation

Figure 1: Body Dust particle concept diagram [1, Figure 1]

This Final Year Project (FYP) is a subset of the Body Dust project as proposed by Dr S. Carrara and Dr P. Georgiou of EPFL and Imperial College London respectively. In their pioneering paper (hereafter referred to as the Body Dust paper), a vision for wirelessly powered, drinkable CMOS integrated circuits (ICs) for the next generation of disease/tumour detection technologies was presented [1]. The idea is to develop tiny CMOS ICs, coated in a bio-compatible packaging with specific bio-molecules that are attracted to the source of a disease in organs and tissues like cancerous tumours. As shown in Figure 1, these CMOS ICs (dubbed Body Dust Particles) should be wirelessly powered and should also send back relevant data about the source of disease to medical professionals. For instance, transmitting the pH and glucose levels could indicate the nature of the tumour as well as any growth in size [5]. This method of detecting the source of a disease has the advantages of being minimally invasive compared to measures like biopsy, relatively small side-effects compared to X-ray and other radiation-based methods, and could be potentially cheap and affordable to the masses. An improvement in medical detection technology would ease the burden of the health sectors in economies.

The Body Dust Project is a massive research project with many research routes that need to be investigated such as designing (i) robust CMOS ICs with all the required functionalities, (ii) constrained wireless powering techniques, (iii) the coating functions that target a specific body mass, bio-compatible packaging etc [1]. This FYP focuses exclusively on research route (i), and more specifically on the design of CMOS ICs that perform amperometry with Signal to Noise Ratio (SNR)[1], power dissipation and area as the optimisation objectives. The basic target metabolite is glucose, a well-studied compound that provides a useful starting point in the research of electro-chemical sensing.

Regrettably, it is beyond our reach to realise a Body Dust particle using current technologies.

---

[1]A slight qualification must be raised here. The signal power in a potentiostat is largely dependent on the measured sensor current, which in itself is also dependent on the electrochemical sensor being used. The electrochemical sensor to be deployed in the Body Dust Project is still in development. Therefore, in this FYP optimising for SNR is actually focused on reducing the noise power instead.

The Body Dust paper has estimated that the Body Dust particles would need to be roughly the size of a red blood cell in order to reliably pass the gut-wall barrier and incur minimum hostile response from the host body [1]. With such a small form factor and an ultra-low power dissipation requirement, it is extremely difficult to develop the analogue CMOS ICs using current technologies. To clarify this point further, note that in order to fulfill the stringent area and power requirements, the CMOS ICs will have to operate at a very small supply voltage, possibly 0.5 V whereas the feature size of the transistors will have to be pushed to very small values, possibly below the 10 nm node. To the best of the author's knowledge, 28 nm is the smallest feature size currently used in analogue CMOS ICs. Regardless of the actual, smallest feature size used in the industry, it is definitely nowhere close to that needed for the Body Dust particle. In addition, analogue CMOS IC design at 0.5 V supply voltage is still an active research area with many unknowns. At 0.5 V supply voltage, it is very difficult to represent analogue information with high SNR in the voltage domain.

Nevertheless, it does not mean that the Body Dust project should be put on hold until the required technology has matured. In fact, it is an opportune moment to investigate the fundamental analogue circuit design trade-offs in the Body Dust Project as well as conduct research into a suitable framework to design such analogue circuits. In other words, it is necessary to go back to basics and view analogue circuit design using a fundamental framework that will prevail even as the feature size and supply voltage decrease.

1.2 Project Aim and Objectives

The overarching goal of this FYP is centered around the investigation of the fundamental trade-offs and a systematic framework to design analogue CMOS ICs for the Body Dust Project. The g_m/I_D methodology is selected as the design framework, and as will be explained throughout this report, the g_m/I_D methodology offers the designer a good grasp of the inherent trade-offs and leads to a more optimum design.

While working toward the overarching aim, the following objectives are undertaken in this FYP.

- Simulate existing potentiostat topologies in order to identify areas for improvement.

- Conduct research into the g_m/I_D methodology.

- Apply the g_m/I_D methodology to designed *optimised* potentiostats and develop a MATLAB script to aid the implementation process.

1.3 Project Contributions

To the best of the author's knowledge, this is the first instance of the g_m/I_D methodology that has been applied to the design of potentiostats. More importantly, this FYP has extended the

g_m/I_D methodology into the design of low power circuits. In most of the existing work, the g_m/I_D methodology is being used to design high performance analogue circuits that are not necessarily optimised for low power dissipation.

A new figure of merit (a weighted Euclidean distance) was proposed that compares op-amps based on their area, power and noise.

Based on the work in this FYP, a conference paper (first draft) was written and is included in the Appendix. This project is still ongoing and the author intends to continue working on this project over the summer holidays with a concrete goal of publishing this work.

1.4 Report Overview

This report is divided into 7 chapters. Chapter 1 outlines the project motivation as well as project objectives. Chapter 2 and Chapter 3 form the background material. Chapter 4 focuses on the circuit analysis, whereas Chapter 5 goes into the details of the implementation. Chapter 6 presents the results of this FYP and the evaluation of the results achieved. Finally, Chapter 7 draws the conclusion of this report and recommends future work.

2 Potentiostats
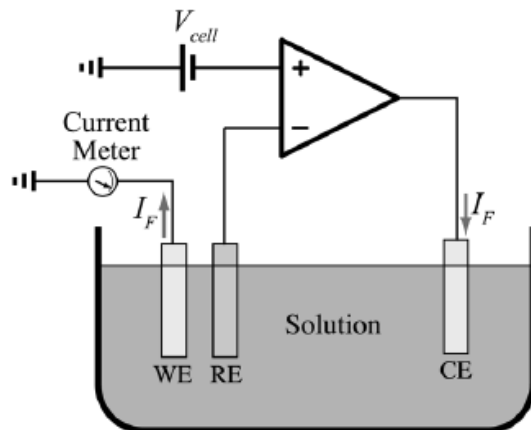
Figure 2: A three-electrode amperometric electrochemical sensing system (potentiostat is represented by an op-amp symbol) [2, Figure 1].

A potentiostat, along with the working electrode (WE), counter electrode (CE), and reference electrode (RE) form the basis of three-electrode amperometric electrochemical sensing systems as shown in Figure 2. In essence, a potentiostat is the electronic circuitry that fulfills two basic functions of maintaining a desired potential difference (V_{Cell}) between the WE and RE and measuring the amount of current directed through the CE [6].

2.1 Cell Potential Control Configurations

For the first function of controlling the cell potential $V_{Cell} = V_{WE} - V_{RE}$, there are three configurations possible, grounded RE, grounded WE and grounded CE as presented in [7]. However, as argued in [2], the grounded WE and grounded RE are electrically identical configurations. Thus, there are only two unique configurations from an electrical point of view.
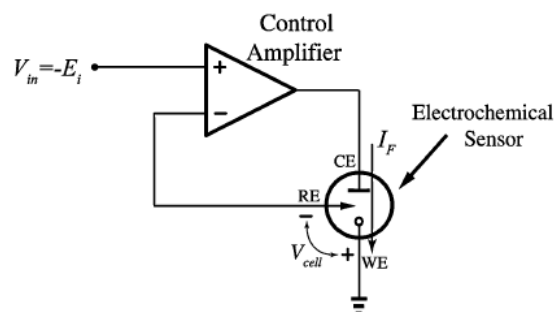
Figure 3: Grounded WE configuration [2, Figure 2].

Figure 3 depicts the grounded WE configuration. This configuration is simple and popular because the WE is kept at a fixed and known potential and the control amplifier forces V_{Cell} to be equal to a certain desired V_{in} via negative feedback. Note that V_{in} needs to be negative

in order for V_{Cell} to be positive, thus, a dual supply voltage system is necessary. The accuracy of this configuration depends on the input offset voltage and voltage gain of the control amplifier. The electrochemical sensor current is often denoted as I_F for Faradic current, since the electrochemical sensor current is generated from redox reactions that obey Faraday's law of electrolysis.
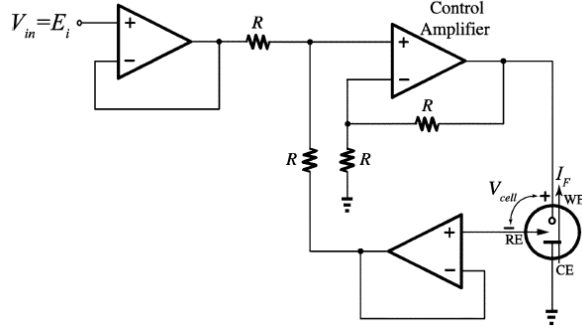
Figure 4: Grounded CE configuration [2, Figure 3].

On the other hand, the grounded CE configuration has also been investigated as an alternative to the grounded WE configuration. A basic implementation is shown in Figure 4. Unfortunately, the grounded CE configuration requires more active and passive components, thus, taking up more area and becoming more susceptible to thermal noise and component mismatch. In addition, the WE is no longer at a fixed and known potential, which makes it more vulnerable to noise and interference, affecting the accuracy of V_{Cell}. However, as argued in [7], the grounded CE configuration might prove more useful than the grounded WE configuration in situations where the WE suffers from electromagnetic interference (EMI) issues that are difficult to mitigate.

2.2 Current Measurement Configurations

The most straightforward method to measure current is to employ a transimpedance amplifer (TIA) which converts a current input to a voltage output. Figure 5 illustrates one possible realisation in which the TIA establishes a virtual ground at the WE and also produces a voltage output that is linearly proportional to the sensor current I_F.

This topology is valued for its relative simplicity and its scalability when measuring vast ranges of sensor currents since the feedback resistance in the TIA, R_M can be programmed appropriately to allow for tunable gain. Alternatively, the TIA can be implemented using switched capacitor circuits if it is not possible for large resistances to be implemented on chip [8]. As an added bonus, in this topology, the sensor current and V_{Cell} are measured with respect to ground.

However, its shortcomings are also rather obvious. Firstly, WE is connected to a virtual ground instead of the true ground. Therefore, the WE can be corrupted by environmental noise and high frequency EMI if not properly shielded [2]. Subsequently, the output level of the TIA
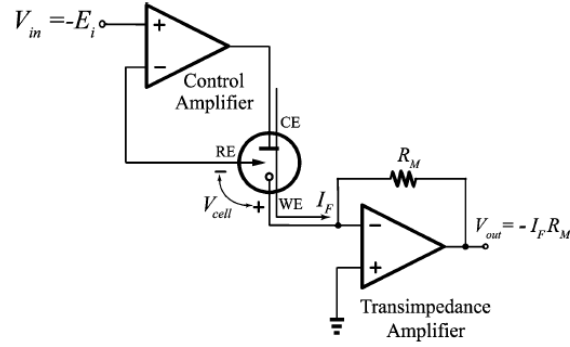
Figure 5: Transimpedance amplifier that measures current [2, Figure 4].

could be corrupted by the environmental noise that goes through the feedback resistance R_M. Secondly, this topology could have stability concerns since the input resistance of the TIA, which should ideally be infinite actually demonstrates an inductive behaviour as reported in [7], [9], and [10]. The input impedance of the TIA is in series with the electrochemical sensor, which is a large capacitive load as seen in Figure 6.
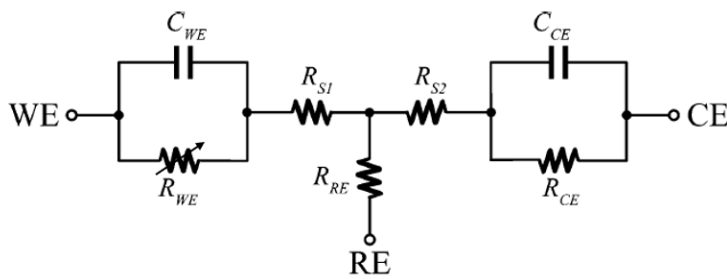
Typical values for a glucose sensor:
R_{WE} = variable
C_{WE} = 1 µF
R_{S1} = 10 Ω
R_{RE} = 6.3 kΩ
R_{S2} = 10 Ω
C_{CE} = 1 nF
R_{CE} = 1 kΩ

Figure 6: Equivalent circuit (only passive components) of an electrochemical cell for glucose sensing [2, Figure 13].

Figure 6 shows the equivalent circuit for a glucose sensor as proposed by Ahmadi and Jullien in [2]. Although this equivalent circuit is only an approximation and does not model any temporal behaviour of the analyte, it has been widely adopted by other works such as [3], [11] and will be used in this FYP. R_{xE} and C_{xE} model the charge-transfer resistance and the double-layer capacitance of the respective electrodes; R_{Sx} models the respective solution resistances.

The series connection between the capacitive glucose sensor load and inductive input resistance of the TIA could result in oscillations and destabilise the potential-control loop [2].

Thirdly, this topology will not be valid when working with single supply voltage systems. For instance, in oxygen-electrode-based glucose sensors, the cell potential is approximately -600 mV with respect to the standard Ag/AgCl RE [12]. The glucose sensor current occurs due to oxygen being reduced (loss of electrons) at the surface of the WE [12]. The potential at RE will be higher than that at the WE by 0.6 V and the sensor current will flow from CE to WE. As a result, the voltage output of the TIA will be negative and below the ground potential, which is

only feasible with dual supply voltage systems.

Two alternative current measurement topologies have been proposed to circumvent the issue with single supply voltage as encountered in the transimpedance amplifier topology. These two topologies measure sensor current by inserting a resistor in the current path at either the WE [9] or the CE [10].
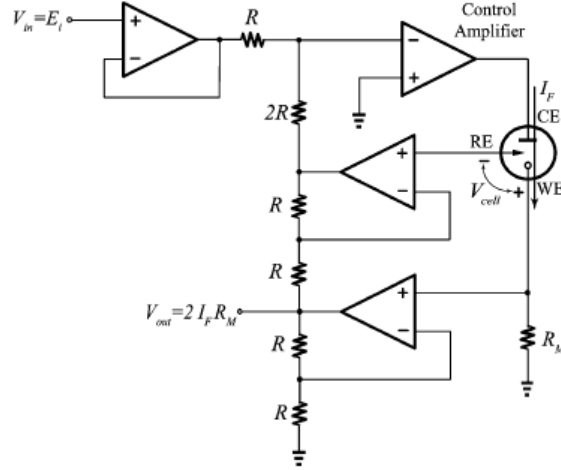
Figure 7: Resistor inserted at WE for current measurement [2, Figure 7].

Figure 7 shows a possible realisation of measuring current by inserting a resistor R_M in the current path at WE. This topology is capable of sensing small values of currents by tuning R_M. However, with a more complex design, this topology is also vulnerable to more component mismatch and thermal noise. In addition, a bigger sensor current would limit the voltage swing headroom at CE since there will be a bigger voltage drop across R_M.
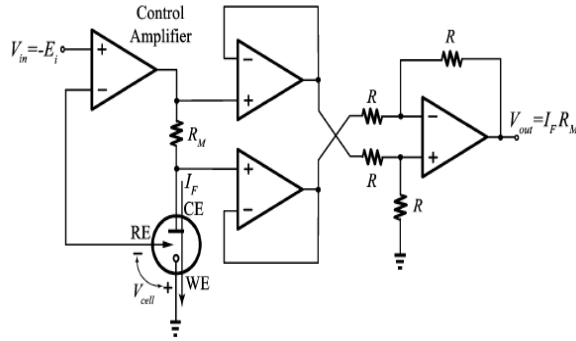
Figure 8: Resistor inserted at CE for current measurement [2, Figure 8].

Figure 8 shows a possible realisation of measuring current by inserting a resistance R_M in the current path at CE. The output voltage generated across R_M is measured using an instrumentation amplifier [2]. This topology shares the disadvantages of the topology in Figure 7 but has better stability performance because there is only one active component (control amplifier) in the control feedback loop [2].

2.3   Current Mirror-based Potentiostat Topology

In light of the strengths and limitations of the potentiostat topologies reviewed above, Ahmadi and Jullien proposed a new topology based on current mirrors in [2]. This new topology avoids many of the disadvantages explained above and remains one of the most influential potentiostat topologies. The novelty of this topology lies in using a current mirror to create a copy of the sensor current and measure the *mirrored current* instead of the original sensor current. Figure 9 depicts this new topology. Note that the sensor, transistor $M1$ and the amplifier A_1 form a feedback loop that fixes the cell potential to be at a desired level, such as -0.6 V for oxygen-based glucose sensors [2]. This topology presents several advantages. Firstly, this is a grounded WE configuration, which makes the WE less vulnerable to noise and EMI pickup. Secondly, this topology requires fewer components which is helpful to alleviate noise and component matching concerns.
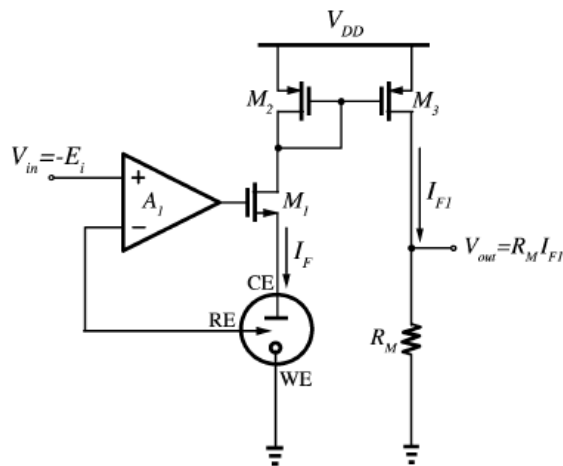
Figure 9: Ahmadi and Jullien's current mirror-based potentiostat [2, Figure 10].

However, the accuracy of this potentiostat will be limited by the accuracy of the current mirror. Current mirrors suffer from current mismatches due to channel length modulation. Standard techniques such as using cascode current mirrors and longer transistors can help to reduce the impact of channel length modulation.

The potentiostat proposed by Ahmadi and Jullien has been simulated in Cadence Spectre. Despite its advantages, it consumes too much area and dissipates too much power, rendering it to be unacceptable in the Body Dust Project context. A transistor-level schematic of the potentiostat proposed by Ahmadi and Jullien along with its Cadence Spectre simulation results (TSMC 180 nm) can be found in the Appendix.

2.4   1-V Low Power High Accuracy Potentiostat Topology

Trakoolwattana and Thanachayanont proposed a potentiostat topology (Figure 10) that boasts low power dissipation and high accuracy [3]. In this potentiostat topology, the op-amp A_1 is *bulk-driven* which allows for a low supply voltage operation (V_{DD} is 1 V) and a low input common-mode voltage level (0.4 V assuming 0.6 V cell potential for glucose). This topology also achieves high accuracy, i.e. good matching between the sensor current (I_F) and the mirrored sensor current (I_{F1}) through the use of a wide-swing cascode current mirror and the op-amp A_2.

A$_2$ serves two purposes. Firstly, it acts as a transimpedance amplifier to produce a voltage output. Secondly, it helps to keep the drain voltages of M_1 and M_2 identical via negative feedback, ensuring good accuracy between the sensor current and its copy.
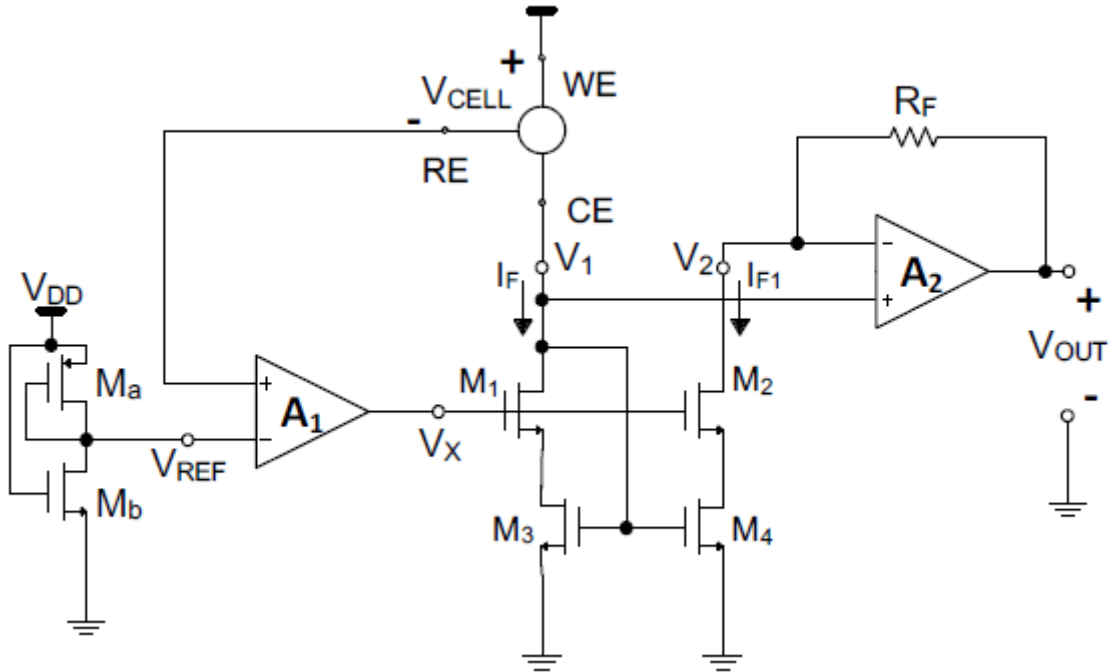
Figure 10: Trakoolwattana and Thanachayanont's low power, high accuracy potentiostat topology [3, Figure 2].

This is an elegant topology that is effective, robust and compact. With its low power and high accuracy, Trakoolwattana and Thanachayanont's topology is a strong candidate for the Body Dust Project. This topology forms the backbone of this FYP and its trade-offs have been studied in-depth. This topology has been optimised using the g_m/I_D methodology and has been simulated in Cadence Spectre (TSMC 180 nm and TSMC 65 nm). The three main building blocks of this topology, A_1, A_2 and the current mirror (M_1 to M_4) will be discussed in greater details in the subsequent material of this report.

2.5 Summary

In this chapter, the potentiostat has been discussed. Various potentiostat topologies have been analysed, with emphasis on their strengths and weaknesses. In addition, the potentiostat topologies proposed by Ahmadi and Trakoolwattana have been analysed in detail. Trakoolwattana's topology was selected as the basis for this FYP.

3 g_m/I_D Methodology

The central idea of the g_m/I_D methodology, as its name suggests is to analyse and design analogue ICs according to the transistor's inversion level, for which the g_m/I_D (transconductance efficiency) figure of merit is a good proxy. The inversion level is a *qualitative* concept that describes the degree to which the MOSFET channel is inverted. The terms "inversion level" and "inversion region" are used synonymously in the literature. In the strong inversion region, the square law holds true to a reasonable degree of accuracy and the MOSFET is biased with a high gate overdrive voltage ($V_{ov} \geq 200\,\text{mV}$). In the weak inversion region, also known as the sub-threshold region, the square law is not valid, and the MOSFET is biased with a gate voltage that is lower than its threshold voltage. The MOSFET exhibits a well-known exponential I-V relationship as seen in Equation 1 [13].

$$I_{DS\,(W.I)} = I_{D0}\frac{W}{L}\exp(\frac{V_{GS}}{nU_T})(1 - \exp\frac{-V_{DS}}{U_T}) \tag{1}$$

If $V_{DS} > 3\,U_T$, then I_{DS} in weak inversion is independent of V_{DS}, and Equation 1 is reduced to Equation 2.

$$I_D = I_0\exp(V_{GS}/nU_T) \tag{2}$$

where $I_0 = I_{D0}\frac{W}{L}$.

The transition from the weak inversion region to the strong inversion region is aptly described as the moderate inversion region. It is apparent that there is a discontinuity in the MOSFET I-V characteristic between the weak inversion region (exponential relationship) and the strong inversion region (quadratic relationship). There is no closed-form expression that captures all three inversion regions as well as "short channel effects" accurately. Unfortunately, this makes the task of quantifying the inversion level, or identifying the boundaries between inversion regions harder.

There have been numerous attempts to quantify the inversion level and there are two prominent examples, the inversion coefficient (IC) and the normalised charge density (q) as advocated by Enz at al [14] and Jespers [15] respectively.

$$IC = \frac{I_D}{I_S} = \frac{I_D}{2nU_T^2\mu C_{ox}\frac{W}{L}} = \left(\frac{L}{2nU_T^2\mu C_{ox}}\right)J_D \tag{3}$$

$$q = -\frac{Q_i}{2nU_T C_{ox}} \tag{4}$$

The inversion coefficient (Equation 3) is basically the drain current density that is normalised by the specific current I_S. The onset of the strong inversion region is at $IC > 10$ ($V_{ov} \approx$

220 mV); the middle of the moderate inversion region is at $IC = 1$ ($V_{ov} \approx 40\ mV$); the onset of the weak inversion is at $IC < 0.1$ ($V_{ov} \approx -72\ mV$) [16].

On the other hand, the normalised charge density (Equation 4) is typically used in the basic EKV model and is closely related to the inversion coefficient. In the same manner as the inversion coefficient, $q < 0.1$, $q = 1$, $q \gg 1$ correspond to weak, moderate and strong inversion levels respectively [15]. One might argue that they are similar concepts that merely represent different physical quantities.

Despite the fact that the inversion coefficient and the normalised charge density are continuous functions across all three inversion regions, they suffer from the drawback of dependency on μC_{ox}. μC_{ox} is dependent on bias conditions and is normally derived a posteriori from real measurements of many transistors. This means that there is a palpable degree of variability in the μC_{ox} value. In addition, when working with advanced CMOS technology, modern analogue designers almost never use the μC_{ox} parameter, which is essentially a relic of the antiquated square law model.

Therefore, we need a better quantifier for the inversion level and g_m/I_D proves to be the ideal candidate.

3.1 Motivation and Historical Background

The well-known MOSFET square law has been used to analyse and design analogue and digital integrated circuits extensively, ever since the MOSFET was invented by Atalla and Kahng in 1959 [17]. The square law is highly attractive because it provides a good set of algebraic equations that yields considerable insight into MOSFET behaviour without weighing designers down with too many complex, unwieldy theoretical solid state physics concepts. Consequently, the square law has formed the basis of practically all undergraduate education in analogue electronics as seen in popular textbooks such as [18] and [19].

However, the MOSFET square law is based on an ideal drift current model and does not take into account non-idealities that arise when the channel length decreases, also termed as "short channel effects". The predominant short channel effects are Drain-Induced Barrier Lowering (DIBL) and punch-through, surface scattering, velocity saturation, impact ionization, and hot electrons [20].

These short channel effects result in errors in the predicted circuit parameters using the square law. For instance, in the strong inversion region (overdrive voltage in the range of several hundred mV), the predicted transconductance, g_m can be off from the actual value by 20-60% [4]. In weak inversion, the square law is no longer valid and cannot be applied, since the current in the MOSFET is primarily diffusion current, instead of drift current.

In light of the modeling limitations of the square law, modern circuit simulation employs complicated device models such as PSP [21] or BSIM6 [22] that are derived from the underlying

solid state physics laws that govern MOSFET behaviour. These models are only suited for powerful computers and as a result, these models do not provide the same level of intuition to the designer compared to the square law. The wide chasm between simulation and hand-analysis results has led to a design flow that is heavily reliant on iterative and time-consuming SPICE-based tweaking.

This heavily iterative style of simulation and tweaking in SPICE causes the engineer to lose insight into analogue circuit design. This style of design is also incapable of providing insights into the fundamental limits of performance of a chosen architecture and would typically result in poor design choices and sub-optimal circuit designs. Furthermore, this design style is inherently time-consuming and is incompatible with the time-to-market pressures and deadlines in the industry. Should an analogue designer adopt this design style, he (she) would be reduced to nothing more than a "SPICE monkey"!

The g_m/I_D methodology was first proposed by Silveira et al in 1996 and was published in the IEEE Journal of Solid-State Circuits (JSSC) [23]. The g_m/I_D methodology was developed in part to address the growing discrepancy between hand-analysis and simulation results in analogue CMOS IC design, as well as to provide a novel, systematic way of designing circuits with a strong grasp of the inherent trade-offs. It was crucial to bridge this gap for the overall progress in IC design. Since the publication of the seminal paper by Silveira et al in 1996, the g_m/I_D methodology has been further developed and applied in a wide range of analogue circuit design contexts. The work presented in [4], [15], [24], [25] and [26] are prominent examples that utilise the g_m/I_D methodology and have been studied thoroughly during the literature review phase of this FYP.

1st Generation	2nd Generation
gm/ID design using the EKV model.	gm/ID design using pre-computed lookup tables.
[24], [25], [26]	[15]
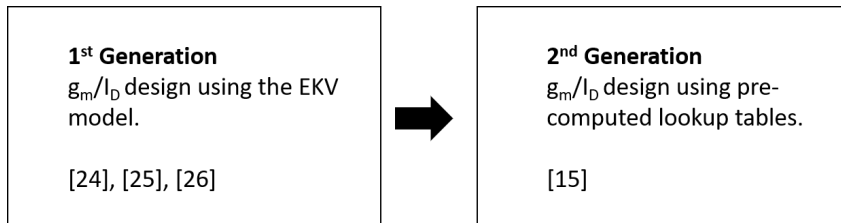
Figure 11: Evolution of the g_m/I_D methodology.

As shown in Figure 11, the work in [24], [25], and [26] can be classed as the first generation application of the g_m/I_D methodology. These works characterise g_m/I_D using either a semi-empirical or model-driven method. In the semi-empirical method, g_m/I_D and other design variables are derived from real measurements or from data generated from advanced MOSFET models like PSP and BSIM. The model-driven method utilises a simple and reliable analytical model such as the EKV model [27], an analytical model that is more advanced than the square law but is not so complicated as to warrant the need for computer simulations.

The work in [15] is the most recent advancement in the development of the g_m/I_D methodology and utilises a SPICE-generated lookup table. This lookup table contains a transistor's equivalent small signal parameters and this data is closely linked to the behaviour of the foundry-provided SPICE model. Therefore, this enables close alignment between the desired specifica-

tions and actual simulation results. The lookup table approach to the g_m/I_D methodology is adopted in this FYP and will be explained in greater detail in Section 3.4. A demonstration of this lookup table approach is given in Section 5.1.

3.2 Transistor Figures of Merit

3.2.1 g_m/I_D

g_m/I_D is a fundamental figure of merit that influences the performance of analogue circuits greatly. g_m/I_D bridges g_m, a small-signal quantity with I_D, a large-signal quantity. Qualitatively, g_m/I_D is a measure of how much g_m (gain) we can get in return for each unit of I_D (bias current) that we invest. I_D can also be viewed as a proxy for the power dissipation. Thus, g_m/I_D encapsulates the most important small-signal and large-signal parameters of an analogue circuit. The subsequent material in this report will emphasise the huge influence that g_m/I_D has in analogue circuit performance.

g_m/I_D also gives an indication of the inversion region [23]. Equation 5 clarifies this further [23].

$$\frac{g_m}{I_D} = \frac{1}{I_D}\frac{\partial I_D}{\partial V_G} = \frac{\partial \ln I_D}{\partial V_G} = \frac{\partial \ln[\frac{I_D}{W/L}]}{\partial V_G} \tag{5}$$

Therefore, g_m/I_D is simply the slope (first derivative) of the $\ln I_D/V_G$ characteristic. In the weak inversion region, g_m/I_D is given by Equation 6.

$$\frac{g_m}{I_D} = \frac{1}{I_D}\frac{\partial I_D}{\partial V_G} = \frac{1}{I_D}\frac{I_D}{nU_T} = \frac{1}{nU_T} \tag{6}$$

It is an established fact that g_m/I_D reaches a maximum in the weak inversion region, where the current-voltage relationship is exponential. The maximum possible g_m/I_D is given by $\frac{1}{U_T} = 38.46$ S/A, assuming $U_T \approx 26$ mV. The maximum g_m/I_D for bulk transistors is generally between 20 to 30 S/A [4]. As the transistor is biased toward the strong inversion region, g_m/I_D decreases in value. g_m/I_D greater than 20 S/A corresponds to weak inversion; g_m/I_D between 20 to 10 S/A corresponds to moderate inversion; g_m/I_D between 2 to 10 S/A corresponds to strong inversion [4]. It should also be noted that these ranges of g_m/I_D remain relatively constant across transistor technologies [4]. Therefore, g_m/I_D can be used as a good proxy for the inversion level.

g_m/I_D is also closely related to the drain saturation voltage (V_{Dsat}), a parameter that many analogue circuit designers use to estimate the voltage headroom.

$$\frac{2}{g_m/I_D} = V_{Dsat} \tag{7}$$

We now seek to verify that Equation 7 is valid in both the strong and weak inversion regions.

In the strong inversion region, where the square law holds to a reasonable degree,

$$I_D = \frac{1}{2}\mu C_{ox}\frac{W}{L}(V_{GS} - V_T)^2 \tag{8}$$

$$g_m = \mu C_{ox}\frac{W}{L}(V_{GS} - V_T) \tag{9}$$

$$\frac{2}{g_m/I_D} = (V_{GS} - V_T) = V_{Dsat} \tag{10}$$

In the weak inversion region, assuming the drain current of the transistor is saturated, that is $V_{DS} > 3\,U_T$, I_D is given by

$$I_D = I_0 \exp(V_{GS}/nU_T) \tag{11}$$

$$g_m = \frac{1}{nU_T}I_D \tag{12}$$

$$\frac{2}{g_m/I_D} = 2nU_T \approx 3U_T \tag{13}$$

We can infer from Equation 13 that $\frac{2}{g_m/I_D}$ corresponds well with the value that we know to be approximately true for the drain saturation voltage of a MOSFET in the weak inversion region.

In addition, g_m/I_D is (to the first-order) independent of the width as shown in Equation 5, which makes it useful for device sizing. g_m/I_D can be employed to find the transistor widths by following this procedure [4]:

1. Derive g_m from the design specifications.

2. Select the transistor lengths to satisfy g_m, speed, area, and matching requirements[2].

3. Decide on g_m/I_D, taking the relevant trade-offs into consideration.

4. Determine I_D/W from g_m/I_D (Figure 16).

5. Derive $I_D = \frac{g_m}{\frac{g_m}{I_D}}$.

6. Derive $W = \frac{I_D}{\frac{I_D}{W}}$.

---

[2]The designer needs to take a small leap of faith here since the transistor lengths may not be easily decided upon. The designer can return to this step after the sizing procedure is complete to improve the circuit further.

The g_m/I_D-centric sizing procedure applies to most high performance circuits, however, it must be slightly modified when designing low-power circuits biased in the weak inversion region. This is because g_m/I_D is approximately constant in the weak inversion region and as a result, many different designs will be mapped to nearly the same g_m/I_D. In other words, a small error in g_m/I_D could lead to a design that is completely off-target. In this case, the g_m/I_D-centric sizing procedure will have to be modified into a J_D-centric sizing procedure, in which the current density (J_D) is used as the design variable. The J_D-centric sizing procedure is a variant of the g_m/I_D-centric sizing procedure and will be demonstrated in greater detail in Section 5.2. Nevertheless, the central tenets of the g_m/I_D methodology still stand. g_m/I_D should be used as the tuning "knob" when designing analogue circuits, instead of the gate overdrive voltage (V_{ov}) commonly used in the square law.

In order to study the trade-offs between power dissipation, noise, distortion, and bandwidth, the g_m/I_D methodology also incorporates other fundamental transistor figures of merit such as $g_m/2\pi C_{gg}$ (transit frequency), g_m/g_{ds} (intrinsic gain), $f_T \times \frac{g_m}{I_D}$, and f_{co}.

3.2.2 $g_m/2\pi C_{gg}$

The transit frequency, f_T is formally defined as the frequency at which the magnitude of small-signal current gain of a common source amplifier reaches 0 dB [18].

$$f_T = g_m/2\pi C_{gg} \tag{14}$$

where C_{gg} is the total gate capacitance and is given by Equation 15.

$$C_{gg} = C_{gb} + C_{gd} + C_{gs} \tag{15}$$

f_T is viewed as the speed bottleneck of a circuit and acts as a warning to designers to steer clear from this frequency. Furthermore, f_T also provides information about the stability of operational amplifiers, since non-dominant poles are usually located at some fraction of f_T. On the other hand, in RF electronics, f_{max} is commonly used which indicates the maximum oscillation frequency at unity power gain [28].

Consequently, MOSFETs typically do not operate at or even near f_T, largely due to the fact that parasitic elements such as transcapacitances and gate resistance make it very challenging to model the MOSFET at high frequencies [4]. Therefore, a general rule of thumb is to assume that $f_T/10$ is the upper frequency limit for predictable circuit operation [28]. Note that the f_T, or rather the $f_T/10$ constraint is relatively relaxed in biomedical contexts, where signals do not vary at high frequencies, unlike the signals in communication systems. This is an important property that grants biomedical electronic circuits the freedom and possibility to operate with low power dissipation.

17

3.2.3 g_m/g_{ds}
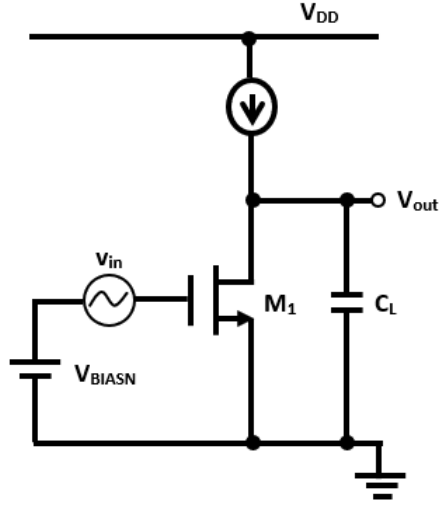
Figure 12: Intrinsic Gain Stage (IGS).

g_m/g_{ds} is the low frequency gain of an intrinsic gain stage. Ideally, g_m should be as large as possible, while g_{ds} should be as small as possible. g_m/g_{ds} is a fundamental figure of merit of a transistor, and some variant of g_m/g_{ds} can be found in practically all small-signal analysis of analogue circuits, ranging from single stage to differential amplifiers. For example, cascode amplifiers and op-amps usually have a voltage gain that is a function of g_m/g_{ds} raised to an integer power n [4].

3.2.4 $f_T \times \frac{g_m}{I_D}$

The thermal noise in a MOSFET can be represented by a voltage source referred to its input (gate) with power spectral density given by Equation 16 [29].

$$v_{i,t}^2 = 4kT\gamma_n \frac{1}{g_m} \, \Delta f \tag{16}$$

where γ_n is normally assumed to be 2/3 in long channel transistors. In order to reduce the thermal noise, g_m needs to increase. There are two methods to do so.

1. Fix device geometry, increase drain current. However, the power dissipation will rise.

2. Fix drain current and increase width. However, the transit frequency will be reduced since g_m/I_D has increased[3].

---

[3]See Figure 13.

For a fair comparison, it is necessary to craft a new figure of merit that takes into account thermal noise, DC bias current and gain-bandwidth. However, before doing so, it would be useful to elaborate on the relationship between the gain-bandwidth product (GBW) of the IGS and its f_T.

The cutoff (-3dB) frequency of the IGS is given by

$$f_c \approx \frac{g_{ds}}{2\pi C_L} \tag{17}$$

The low frequency (DC) gain of the IGS is given by

$$A_{v0} = -\frac{g_m}{g_{ds}} \tag{18}$$

The unity gain frequency of the IGS is given by

$$f_u \approx \frac{g_m}{2\pi C_L} \tag{19}$$

Since the IGS is a first-order system, its GBW is a constant and is given by the product of A_{v0} and f_c, which is f_u. The GBW can be expressed in a different form, by introducing a new variable, $\mathfrak{R}$.

$$\mathfrak{R} = \frac{C_L}{C_{gg}} \tag{20}$$

$$\frac{f_T}{f_u} = \frac{f_T}{GBW} = \frac{g_m/(2\pi C_{gg})}{g_m/(2\pi C_L)} = \mathfrak{R} \Rightarrow GBW = \frac{f_T}{\mathfrak{R}} \tag{21}$$

Returning to the goal of devising a new figure of merit, a simple candidate is proposed in Equation 22 [4].

$$\frac{GBW}{Thermal\ Noise \times I_D} = \frac{\frac{f_T}{\mathfrak{R}}}{4kT\gamma_n \frac{1}{g_m} \times I_D} \tag{22}$$

Assuming $\mathfrak{R}$ and γ_n are fixed and constant, Equation 22 reduces to $f_T \times \frac{g_m}{I_D}$, which should be as large as possible.

3.2.5 f_{co}

In a MOSFET, the flicker noise is generally modelled by Equation 23 [18].

$$v_{i,f}^2 = \frac{K_f}{C_{ox}WL} \cdot \frac{1}{f} \, \Delta f \tag{23}$$

With the exception of area (WL), all other parameters in Equation 23 are beyond the control of the designer. As such, increasing the area of the transistor is the only way to reduce flicker noise in the transistor. However, at the circuit level, techniques such as correlated double sampling and chopping can be employed to reduce the flicker noise [30].

Since flicker noise is only dominant in low frequency circuits, a more meaningful parameter to consider is the flicker noise corner frequency, f_{co}, which is the frequency at which the flicker noise and thermal noise power spectral densities are equal. As a general rule of thumb, flicker noise can be neglected at frequencies 1 to 2 orders of magnitude above f_{co}.

An expression for f_{co} can be found by equating Equation 16 and Equation 23.

$$4kT\gamma_n \frac{1}{g_m} = \frac{K_f}{WL} \cdot \frac{1}{f_{co}} \Rightarrow f_{co} = \frac{K_f}{4kT\gamma_n} \cdot \frac{g_m}{WL} \tag{24}$$

3.3 Inherent Trade-offs

Having established the fundamental figures of merit, it follows naturally to investigate the major trade-offs in analogue circuits from a g_m/I_D perspective. Note that Figure 13 to Figure 16 were plotted using the lookup table data generated from the TSMC 180 nm SPICE model for an N-channel IGS. In Figure 13 to Figure 16, the V_{DS} of the IGS was fixed to 0.9 V. The exact value of V_{DS} being used is not of great significance since the plots are known to scale approximately with V_{DS} [4].
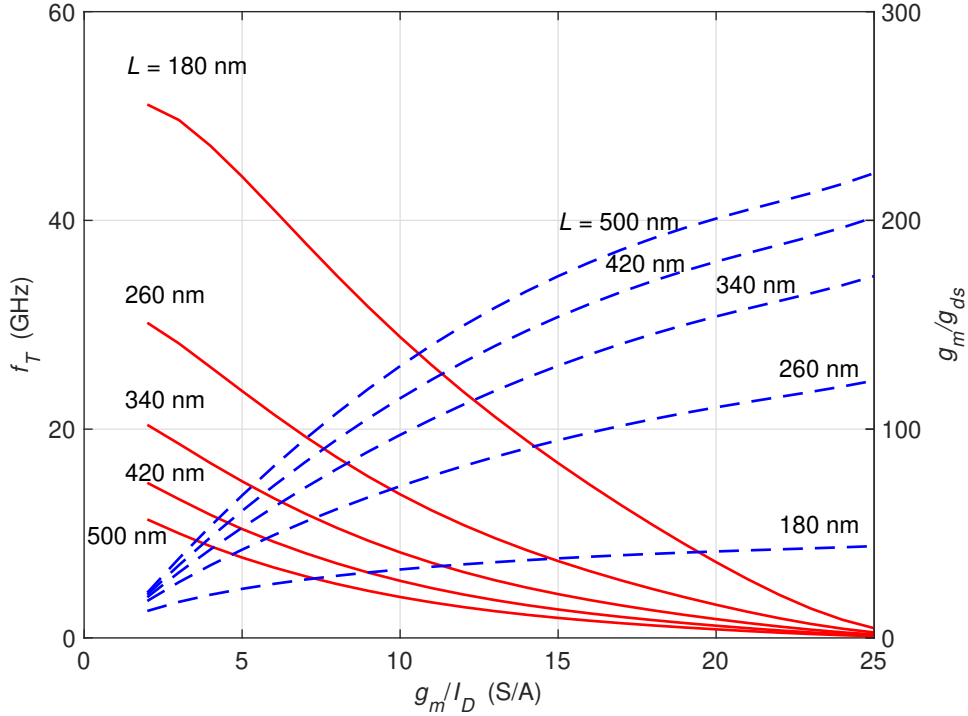
Figure 13: f_T and g_m/g_{ds} against g_m/I_D for a range of lengths; $V_{DS} = 0.9$ V; $V_{SB} = 0$ V; TSMC 180 nm.

Figure 13 shows that designing for a high transit frequency, i.e. bandwidth, and designing for a high gain are mutually exclusive objectives. There are distinct trade-offs that emerge depending on the transistor's g_m/I_D:

- Large $g_m/I_D \rightarrow$ weak inversion $\rightarrow$ large gain and transistor area but slow.

- Small $g_m/I_D \rightarrow$ strong inversion $\rightarrow$ small gain and transistor area but fast.

Figure 13 underscores the point that in biomedical applications, where bandwidth is generally not a primary concern, the transistors can be biased in the weak inversion region for greater gain. Interestingly, the trade-offs between speed, gain and a third variable, area can also be inferred from Figure 13. It is a well-known fact that a transistor with a larger size is slower since it has a greater total gate capacitance. Therefore, there is a direct relationship between the transit frequency of a transistor and its area.

21
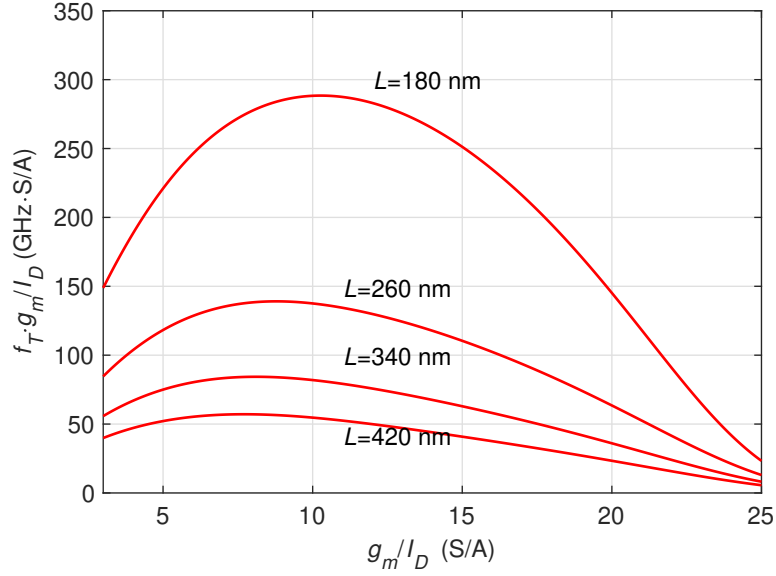
Figure 14: $f_T \times \frac{g_m}{I_D}$ against g_m/I_D for a range of lengths; $V_{DS} = 0.9$ V; $V_{SB} = 0$ V; TSMC 180 nm.

Figure 14 depicts the trade-off between thermal noise and GBW for a transistor. It can be seen that the maxima of the plots occur at $g_m/I_D \approx 7$ to 10 S/A, which corresponds to the moderate-strong inversion region. Shorter channel lengths have a higher maxima since f_T is larger with short channel lengths. It must be pointed out that $f_T \times \frac{g_m}{I_D}$ may not be an appropriate figure of merit in applications where the f_T requirement is relaxed. $f_T \times \frac{g_m}{I_D}$ can be modified by replacing the f_T variable by the pre-selected f_T value. Subsequently, the plots in Figure 14 will be constrained to satisfy this f_T value. Nevertheless, the argument that the moderate-strong inversion region offers the best compromise between thermal noise and GBW still stands.
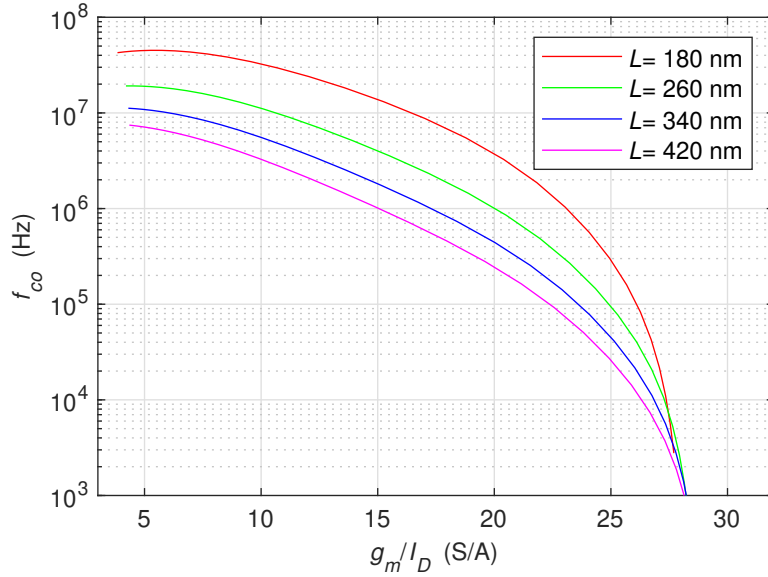
Figure 15: Flicker noise corner frequency against g_m/I_D for a range of lengths; $V_{DS} = 0.9$ V; $V_{SB} = 0$ V; TSMC 180 nm.

Figure 15 confirms Equation 24 in that f_{co} decreases for larger g_m/I_D and larger lengths. Biasing transistors in the weak inversion region and using long channels help to reduce f_{co}. The plots in Figure 15 are obtained using a more direct method that does not need to compute estimates for γ_n and K_f [4]. The thermal noise current density (STH) data can be extracted and stored via SPICE simulations. Likewise, the flicker noise power spectral density at 1 Hz can be extracted from SPICE simulations. Thermal noise is approximately white and is normally represented as a horizontal line on noise voltage against frequency graphs. Since the flicker noise power spectral density is known at 1 Hz and assuming a rolloff that is exactly equal to $\frac{1}{f}$, a straight line equation of the flicker noise can be derived. The intersection of this downward-sloping line with the horizontal thermal noise line will yield f_{co}.
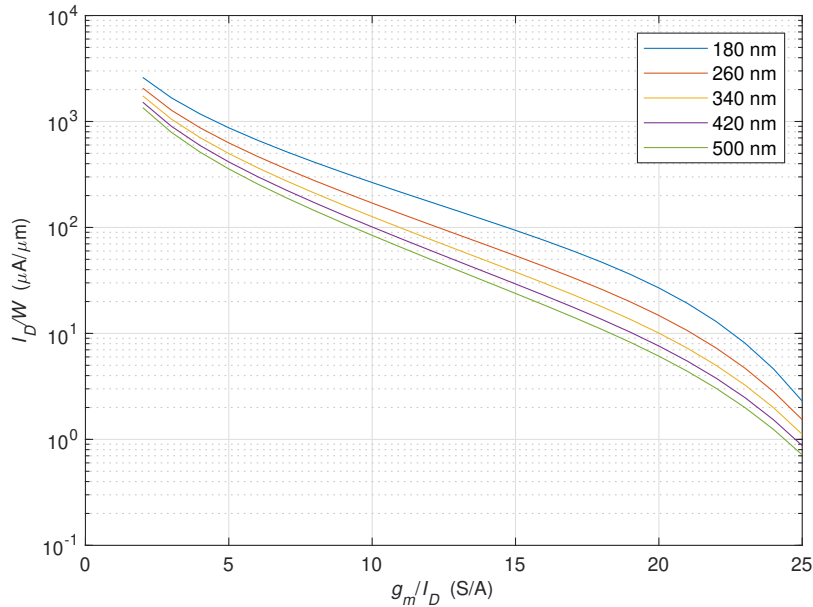
Figure 16: I_D/W against g_m/I_D for a range of lengths; $V_{DS} = 0.9$ V; $V_{SB} = 0$ V; TSMC 180 nm.

Figure 16 provides a link between g_m/I_D and I_D/W, which is used in calculating the required width toward the end of the sizing procedure mentioned in Section 3.2.1. It is also apparent from Figure 16 that a large (small) g_m/I_D implies large (small) transistor area.

The inherent trade-offs can be summarised as follow. The optimum intrinsic gain, input-referred flicker noise, and power dissipation can be achieved by biasing the transistor toward the weak inversion region and setting a long transistor length. However, the intrinsic transistor bandwidth (f_T) is optimised under the opposite conditions of biasing the transistor toward the strong inversion region and setting a short transistor length. Figure 17 depicts the trade-offs graphically. The moderate inversion proves to be the best compromise for high performance analogue circuits.
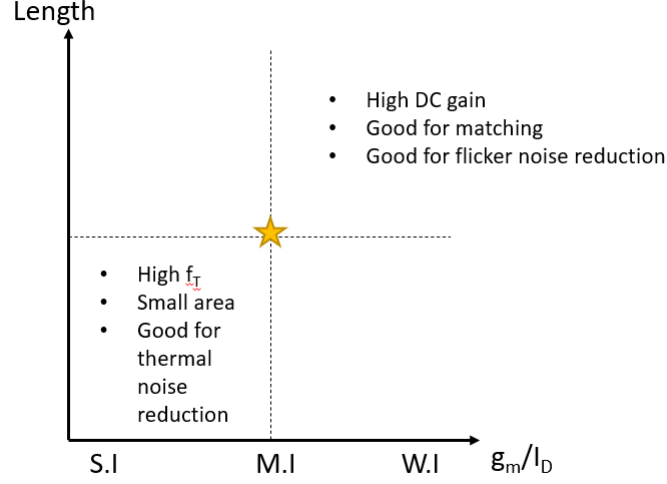
Figure 17: Trade-offs plane; the star denotes the best compromise spot.

Lastly, on a more relevant note to this FYP, the trade-offs between transistor area, noise and power can also be expressed in Equation 25, which relates the total noise in a transistor (thermal and flicker noise) to its area (A) and power dissipation (current I). K_w and K_f are the technology constants associated with the thermal and flicker noise respectively; p is the exponent in the g_m versus I_D relationship. The complete derivation of Equation 25 can be found in [31] and will not be repeated here.

$$\bar{v_n^2} = \left(\frac{K_w(p)}{I^p} \right) \Delta f + \left(\frac{K_f}{A} \right) \ln \left(\frac{f_h}{f_l} \right) \tag{25}$$

Equation 25 shows that the DC current (power dissipation) must be increased to reduce the input-referred thermal noise, whereas the area must be increased to reduce the input-referred flicker noise. Equation 25 can complement the g_m/I_D methodology in understanding the trade-offs in a circuit.

3.4 Pre-Computed Lookup Table Approach
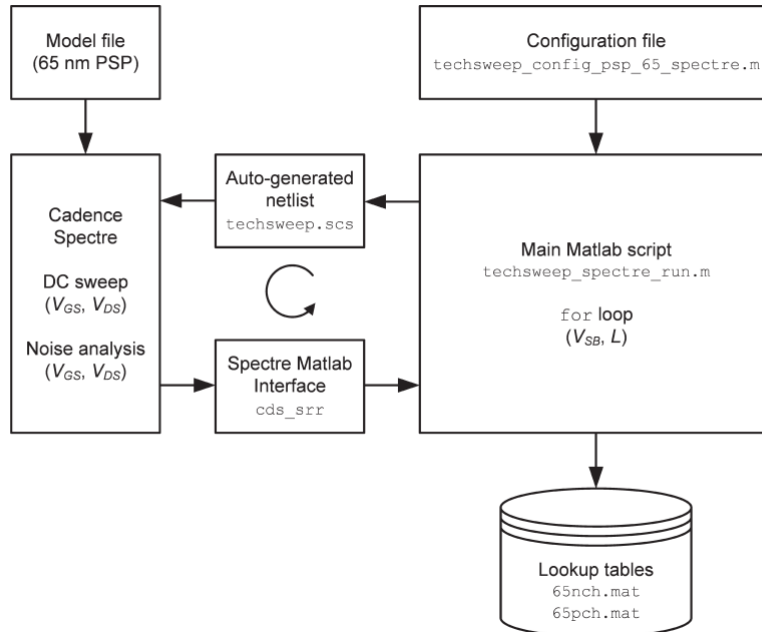
Figure 18: Example setup in to generate a lookup table [4, Figure A.2.1].

The work by Jespers and Murmann in [4] constitutes the second generation of the g_m/I_D methodology. Taking advantage of the exponential increase in computational power over the past few decades, Jespers and Murmann have championed the use of SPICE-generated lookup tables to complement the g_m/I_D methodology. Figure 18 depicts the example flow used in [4] to generate lookup tables. Essentially, there is a configuration file that enables the designer to key in the relevant Cadence path setup information. The main MATLAB script includes a netlist specification for MOSFET simulation that is directly fed[4] into Cadence Spectre. For a pre-selected width, the main MATLAB script runs DC sweeps (and noise simulations) along the four variables, V_{GS}, V_{SB}, V_{DS}, L and stores the data into an easily accessible data structure. It is not necessary to sweep the width since the parameters have been found to scale approximately with the device width [4].

Jespers and Murmann have also provided two simple companion MATLAB functions (`look_up` and `look_upVGS`) in [4] that basically return the desired information (g_m/I_D, g_m/g_{ds}, f_T etc) based on the provided input arguments. These functions perform interpolation where necessary. In summary, Jespers and Murmann provided i) a script that returns a SPICE-generated lookup table and ii) companion functions to retrieve the relevant parameters from the lookup table. The work in this FYP uses these MATLAB scripts extensively, although it must be reiterated that these MATLAB scripts are not "plug-and-play" programs that take over the intuition behind analogue circuit analysis and design. The analysis and design of circuits are guided by the g_m/I_D methodology. The provided MATLAB scripts merely perform the basic setup and

---

[4]Note that it is necessary to have both Cadence and MATLAB installed in the same workstation to enable direct data.

number crunching.

3.5   Summary

In this chapter, the g_m/I_D methodology was discussed in detail. Firstly, the motivation and historical background of the g_m/I_D methodology was introduced. Next, the intuition and reasoning behind the g_m/I_D methodology were elaborated. The fundamental transistor figures of merit as well as the inherent trade-offs in a circuit were explained with the aid of MATLAB plots generated using the lookup table data.

4   Analysis of Trakoolwattana & Thanachayanont's Potentiostat

In this chapter, the three main building blocks of Trakoolwattana and Thanachayanont's potentiostat will be analysed in detail with emphasis on the frequency response of the two op-amps. The inherent trade-offs in the circuits will also be investigated. The aim of this chapter is to establish a solid understanding of the circuits, which serves as the foundation for the design and implementation of the circuits.

4.1   Two-Stage Op-amp
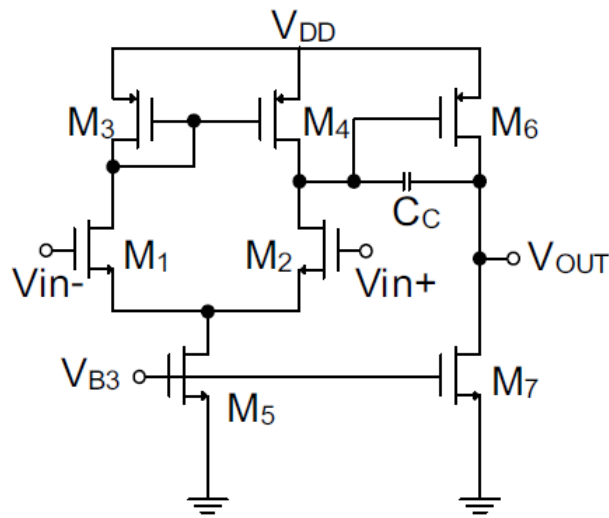
Figure 19: Two-stage op-amp used in this FYP [3, Figure 4b].

The two-stage op-amp in Figure 19 has an N-channel differential amplifier as its first stage, followed by a P-channel common source amplifier with an NMOS load. This popular op-amp topology offers many advantages such as large open loop voltage gain, good common mode rejection ratio (CMRR), and a small number of transistors etc. The main limitation of this topology is the non-dominant pole introduced by the load capacitance and the output impedance. A Miller compensation capacitance is used to improve the stability of the op-amp by transforming the overall system to a first-order system.

This op-amp has differential inputs and a single-ended output. It is assumed that this two-stage op-amp drives capacitive loads (a few pF), otherwise a third stage (buffer) will be required to drive resistive loads. There are two high impedance nodes in this op-amp and they influence the frequency response of the op-amp greatly.

The two-stage op-amp can be analysed from many aspects, and Gray and Meyer have published an excellent tutorial in [32]. For the sake of practical relevance and brevity, only the frequency response (small signal transfer function and stability) of the two-stage op-amp will be examined in detail in Section 4.1.1.

4.1.1 Small Signal Transfer Function

The small signal transfer function, $H(s)$ from the differential inputs to the single-ended output can be expressed in the form of a two-pole, one-zero system.

$$H(s) = \frac{A_{v0}(1 - \frac{s}{z_1})}{(1 + \frac{s}{p_1})(1 + \frac{s}{p_2})} \tag{26}$$

where A_{v0} is the DC (low frequency) gain, p_1 is the dominant pole, p_2 is the output pole, and z_1 is the zero introduced by the Miller compensation capacitor. The two-stage op-amp has been analysed extensively and a complete derivation can be found in [19].

A_{v0} is given by the multiplication of the DC gain of the 2 individual stages[5].

$$|A_{v0}| = |A_{v1}|\,|A_{v2}| = \frac{g_{m2}}{g_{o2} + g_{o4}}\frac{g_{m6}}{g_{o6} + g_{o7}} \tag{27}$$

The poles are influenced by the junction capacitances of the transistors situated along the gain path. The locations of the poles are given below.

- p_1 is the dominant pole (lowest frequency) and is located at

$$p_1 = -\frac{1}{g_{ds2} + g_{ds4}}\frac{1}{(1 + |A_{v2}|)C_C} \approx -\frac{1}{g_{ds2} + g_{ds4}}\frac{1}{|A_{v2}|\,C_C} \tag{28}$$

- p_2 is the output pole and is located at

$$p_2 = -\frac{g_{m6}C_C}{C_1 C_2 + C_C(C_1 + C_2)} \tag{29}$$

where C_1 accounts for the gate capacitance seen at M_6 and the junction capacitances at M_2 and M_4,

$$C_1 = C_{gs6} + C_{db2} + C_{db4} + C_{gd2} + C_{gd4} \tag{30}$$

C_2 accounts for the load capacitance, C_L and the junction capacitances of M_6 and M_7.

$$C_2 = C_L + C_{db6} + C_{db7} + C_{gd6} + C_{gd7} \tag{31}$$

Note that the Miller compensation capacitance has caused the dominant pole to be located at a very low frequency and the non-dominant output pole to be located at a much higher frequency. Therefore, this compensation technique is known as "pole splitting".

---

[5]Note that the transistor numbering in the design equations match the transistor numbering used in Figure 19.

The position of the zero, z_1 is given by

$$z_1 = +\frac{g_{m6}}{C_C} \qquad (32)$$

z_1 is introduced because C_C offers a feedforward path for the signal to travel from the input to the output. At high frequencies, C_C is essentially shorted out and M_6 becomes a diode-connected transistor. The impedance of M_6 (looking from the drain) is approximately $1/g_{m6}$, and thus a zero at Equation 32 is introduced. This zero is undesirable since it degrades the phase margin of the amplifier. There are numerous ways to mitigate the effect of this zero, such as inserting a resistor in series with C_C [33] or even a source follower in between C_C and the output so as to break the feedforward path [34]. However, these methods introduce other problems such as increased number of transistors and (possibly) passive components, an increase in thermal noise and even an increase in power dissipation. Undeniably, z_1 degrades the phase margin and may justify the need for techniques to cancel its effect. However, in practice, an adequate phase margin can be achieved without the use of a nulling resistor or source follower via a careful choice of transistor dimensions and C_C.

Strictly speaking, the current mirror (M_3 and M_4) introduces another pole-zero pair (doublet). However, this doublet is generally located at very high frequencies (at least three times the unity gain frequency of the amplifier) [15]. Consequently, this doublet has negligible impact on the phase margin and will not be discussed.

4.1.2 Unity Gain Frequency and Phase Margin

On its own with no compensation, the two-stage op-amp is generally unstable since it contains many poles that would degrade the phase margin. Therefore, the objective of compensation is to reduce the two-stage op-amp into a first-order system, which is unconditionally stable. In other words, it is desirable for the two-stage op-amp to exhibit only one *dominant* pole in which p_1 is much smaller than p_2.

The dominant pole condition can be summarised as follow

$$\frac{p_1}{p_2} \leq 0.1 \qquad (33)$$

The requirement in Equation 33 for the dominant and non-dominant poles to be spaced one decade apart is arbitrary. The one-decade-apart condition is imposed to justify the use of various approximations in estimating the pole and zero locations. In practice, the designer does not need to follow this condition exactly. In fact, should this condition be relaxed, the second stage of this op-amp can have a much lower power dissipation. This interesting observation will be explained further in the implementation of the two-stage op-amp (Section 5.3).

Under the dominant pole condition, the unity gain frequency (UGF) of the op-amp can be approximated using the equation for the UGF of the IGS (a first-order system).

$$f_u \approx \frac{g_{m2}}{2\pi C_C} \tag{34}$$

It must be reiterated that for a first-order system, the terms GBW and UGF are approximately equal and are treated as equivalent.

The phase margin (PM) is defined by the phase response of the closed-loop unity-feedback system when its gain is at 0 dB. The phase margin gives an indication of the closed-loop stability of a system. A general rule of thumb is for circuits to have at least 45° phase margin. The step response of a feedback system with a smaller value of phase margin is more likely to exhibit ringing. A phase margin of 60° is generally considered a optimal value, since the step response of the feedback system shows faint ringing and a rapid settling time [18].

The phase margin and gain bandwidth product of an amplifier are usually fixed specifications that must be met. The value of C_C plays a huge role in determining if these specifications can be achieved. Therefore, this warrants for a closer analysis of C_C.

4.1.3 Miller Compensation Capacitance

C_C is generally determined by applying some heuristic, followed by minor tweaking if necessary. This is because it is very difficult to obtain an exact, analytical expression for C_C since the positions of poles in the circuit are only reasonable and intuitive approximations. The high frequency small signal model of the MOSFET is also an approximation that assumes quasi-static capacitances even though real MOSFETs have been known to exhibit non-quasi-static capacitances (transcapacitances). By using these valid approximations, the designer has no choice but to adopt some heuristic to choose a suitable C_C and thereafter apply some fine-tuning. This section discusses two heuristics as proposed by Allen and Jespers in detail.

In [35], Allen proposed a heuristic that the zero be placed 10 times and the non-dominant pole be placed at least 2.2 times above the amplifier's GBW respectively to achieve a phase margin of 60°. This is derived below.

Using the angular frequency notation, the expression for 60° phase margin is given by

$$PM = \pi - \arctan \frac{\omega}{|p_1|} - \arctan \frac{\omega}{|p_2|} - \arctan \frac{\omega}{z_1} = \frac{\pi}{3} \tag{35}$$

Assuming $\omega = \omega_u$, Equation 35 can be re-written as

$$\frac{2\pi}{3} = \arctan \frac{\omega_u}{A_{v0}} + \arctan \frac{\omega_u}{|p_2|} + \arctan 0.1 \tag{36}$$

Assuming A_{v0} is large, $\arctan \frac{\omega_u}{A_{v0}}$ tends to zero, and $\arctan \frac{\omega_u}{|p_2|}$ tends to $\frac{2\pi}{3} - \arctan 0.1 \approx$ 1.9947, which results in $|p_2| \geq 2.2\,\omega_u$.

Allen simplified the expression for the location of the non-dominant pole (Equation 29) further.

$$p_2 \approx -\frac{g_{m6}C_C}{(C_1 + C_C)C_2} \approx -\frac{g_{m6}}{C_2} \tag{37}$$

This simplification is justified since C_C and C_2 are usually several orders of magnitudes larger than C_1, which is composed entirely of MOSFET parasitic capacitances.

The following conditions hold

$$\frac{g_{m6}}{C_C} > 10\,\frac{g_{m2}}{C_C} \Rightarrow g_{m6} > 10\,g_{m2} \tag{38}$$

$$\frac{g_{m6}}{C_2} > 2.2\,\frac{g_{m2}}{C_C} \tag{39}$$

Equation 38 and Equation 39 can be combined to give

$$C_C > 0.22\,C_2 \approx 0.22\,C_L \tag{40}$$

Since C_L is known a priori, Equation 40 offers a simple rule of thumb when selecting C_C.

In [15], Jespers proposed a similar rule of thumb that must be derived computationally. The main difference is that Jespers did not apply the simplification in Equation 37 but arrived at a value for C_C using Equation 29 and a `for` loop.

Let

$$\frac{f_{z1}}{f_u} = Z = 10 \tag{41}$$

$$\frac{f_{p2}}{f_u} = NDP = 4 \tag{42}$$

Rewriting Equation 34 and Equation 32 in a different form,

$$g_{m2} = 2\pi f_u C_C \tag{43}$$

$$g_{m6} = 2\pi f_{z1} C_C = \frac{f_{z1}}{f_u} g_{m2} = Z g_{m2} \tag{44}$$

Since f_u, Z, NDP are known a priori, C_C can be determined from Equation 44 by rearranging into a different form and applying the well-known quadratic formula.

Rearranging Equation 44 and replacing the frequency terms with angular frequencies,

$$\omega_{z1} = \frac{g_{m6}C_C}{C_1C_2 + C_C(C_1 + C_2)} = \frac{\omega_{z1}C_C^2}{C_1C_2 + C_C(C_1 + C_2)} \tag{45}$$

$$\frac{\omega_{z1}}{\omega_u} = \frac{\omega_{z1}C_C^2}{C_1C_2 + C_C(C_1 + C_2)}/\omega_u \Rightarrow \frac{NDP}{Z} = \frac{C_C^2}{C_1C_2 + C_C(C_1 + C_2)} \tag{46}$$

$$C_C^2 - \frac{NDP}{Z}(C_1 + C_2)C_C - \frac{NDP}{Z}C_1C_2 = 0 \tag{47}$$

Ignoring the negative root of Equation 47,

$$C_C = \frac{1}{2}\frac{NDP}{Z}\left(C_1 + C_2 + \sqrt{(C_1 + C_2)^2 + 4C_1C_2\frac{Z}{NDP}}\right) \tag{48}$$

Equation 48 requires the values of C_1 and C_2, none of which is known before the transistors are sized. However, it is possible to start with a reasonable estimate for C_C, derive the corresponding g_{m2} and g_{m6}, and size the transistors using the g_m/I_D methodology. The parasitic capacitances can be obtained once the transistors are sized and fed back into Equation 48 to obtain a more accurate guess for C_C. This procedure can be implemented with an iterative `for` loop in MATLAB that eventually converges at a final set of values. This `for` loop is explained in greater detail in Section 5.3.

In summary, Allen's heuristic is an useful starting point for choosing a value of C_C. It is a "coarse" knob that can be complemented by a "finer" knob (Jespers' heuristic).

4.2 Bulk-driven Folded Cascode Op-amp
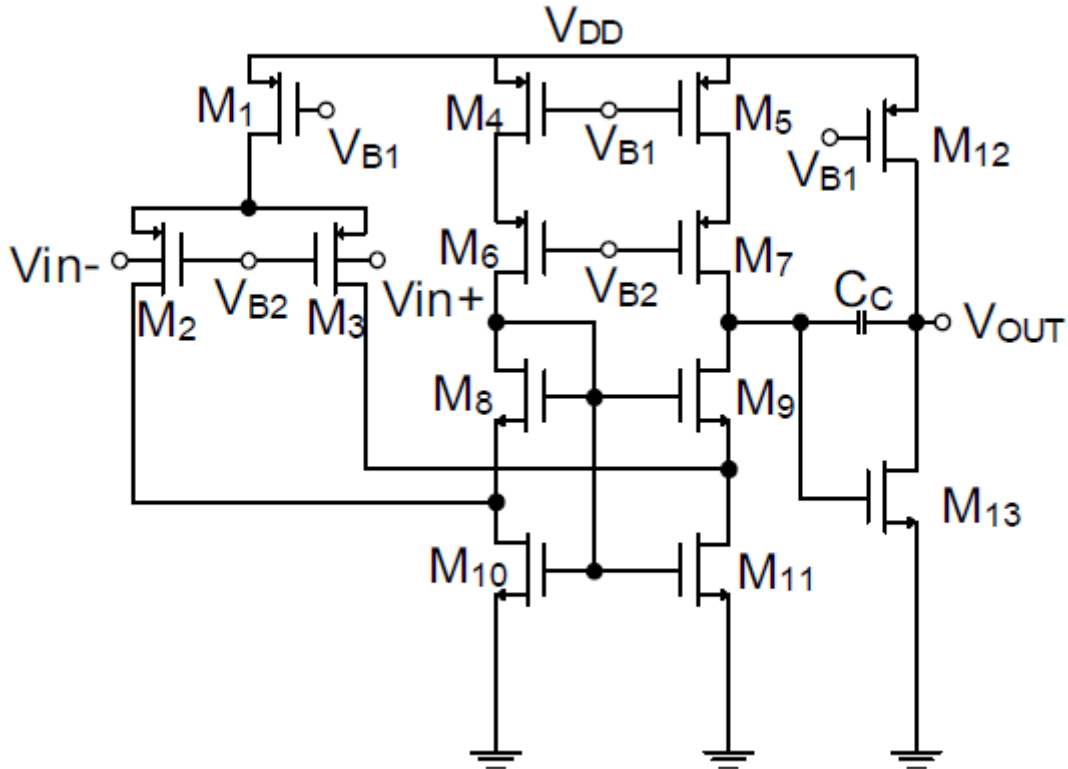
Figure 20: Bulk-driven Folded Cascode Op-amp [3, Figure 4a].

Figure 20 depicts a bulk-driven folded cascode stage followed by a conventional common source amplifier with an active load as the second stage. The input stage of this op-amp is unorthodox since most op-amps are gate-driven. The bulk-driven input stage deserves special attention and will be explained further below.

The bulk-driven technique was first introduced by Guzinski et al in 1987 [36]. Since then, there has been a growing number of published papers on this topic. The bulk-driven technique allows for low supply voltages and is especially attractive in low power applications such as implantable biosensors.

In a bulk-driven MOSFET, its gate terminal needs to be connected to a bias voltage to form a channel (inversion layer) between the drain and source terminals. Once this channel is established, the current flowing through this channel can be modulated by applying at signal at the bulk terminal, i.e. varying V_{BS}. The bulk-driven MOSFET operates in a similar manner as the junction gate field-effect transistor (JFET) [37]. Note that the term "bulk-driven" can be misleading as it seems to suggest that the MOSFET gate can be left floating, which is not true.

An important advantage of the bulk-driven technique is that it frees the designer from the threshold voltage requirement (to a certain extent). The threshold voltage of a MOSFET can be given by Equation 49.

$$V_{\text{th}} = V_{\text{th0}} + \gamma(\sqrt{|2\phi_F - V_{\text{BS}}|} - \sqrt{|2\phi_F|}) \tag{49}$$

where V_{th0} is the zero bias threshold voltage, γ is the body effect coefficient, ϕ_F is the Fermi potential [18].

Equation 49 shows that by varying $|V_{BS}|$, the threshold voltage can be reduced. However, there is an important caveat. $|V_{BS}|$ can be increased or decreased as long as the parasitic Bipolar Junction Transistors (BJTs) in the MOSFET do not turn on completely[6]. A reduction in the threshold voltage gives the designer the freedom to operate at much lower supply voltages.

Another advantage of the bulk-driven technique is the improvement in linearity. In the seminal paper by Guzinski et al, the bulk-driven technique was applied in the differential input stage of an operational transconductance amplifier. Since the bulk transconductance is smaller than the gate transconductance, a bulk-driven differential amplifier has a wider input linear range.

Nevertheless, the bulk-driven techniques has many other limitations. Firstly, the bulk-driven technique requires isolated bulk contacts. Therefore, in a typical p-substrate, n-well process, only P-type transistors can support the bulk-driven technique. Moreover, these bulk-driven P-type transistors need to be fabricated in different wells and this could make good matching between transistors difficult to achieve. It will be very challenging to apply good layout techniques such as inter-digitisation and common centroiding. Secondly, as mentioned previously, the bulk transconductance (g_{mb}) is typically smaller than the conventional gate transconductance (g_m) by a factor of 0.2 to 0.4 [37]. Therefore, by using the bulk-driven technique, the MOSFET transit frequency is lower and could lead to a poorer frequency response. In addition, with a smaller transconductance, the equivalent input-referred noise in a bulk-driven MOSFET could be more significant than that in a gate-driven MOSFET [37].

Apart from the input stage, the rest of the folded cascode op-amp in Figure 20 consists of standard circuit topologies and will be discussed by analysing their DC gain and frequency response.

---

[6]The parasitic BJTs can cause latch-up in a MOSFET and can lead to unwanted transistor behaviour. By inspection of the structure of the MOSFET, it can be seen that a parasitic BJT can be formed by the source of a NMOS, the p-substrate and an n-well.

4.2.1 Small Signal Transfer Function

The folded cascode op-amp is a large, complicated circuit and subsequently, it is not worthwhile performing a complete small signal analysis. It is more intuitive to analyse the folded cascode by taking advantage of its symmetrical structure and using half-circuits. To this end, we would need to leverage on Lemma 1.

Lemma 1 *"In a linear circuit, the voltage gain is equal to $-G_m R_{out}$, where G_m denotes the transconductance of the circuit when the output is shorted to ground and R_{out} represents the output resistance of the circuit when the input voltage is set to zero." [18, p. 66]*
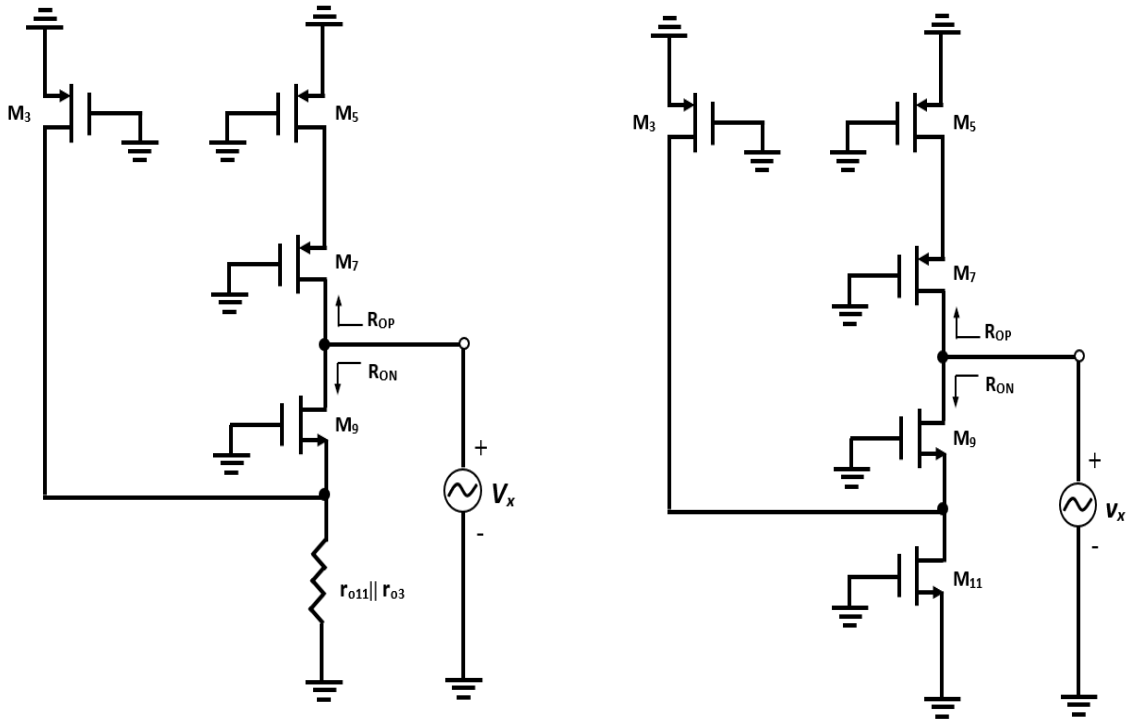
(a) Equivalent half-circuit.　　　　(b) Alternative representation of the circuit in (a).

Figure 21: Equivalent circuit to derive G_m and R_{out} of the folded cascode stage.

As shown in Figure 21a, we can make the observation that $r_{o3}||r_{o11}$ is generally several orders of magnitude larger than the impedance looking into the source of M9, which is $(g_{m9} + g_{mb9})^{-1}||r_{o9}$. Therefore, the drain current of M3 (I_{D3}) is approximately equal to the short circuit current and $G_m \approx g_{mb3}$[7].

By inspection of Figure 21b, $R_{out} = R_{OP}||R_{ON}$. Invoking a well-known approximation for the output resistance of a cascoded stage,

---

[7]The input transistors are bulk-driven. Therefore, G_m is approximately equal to the bulk transconductance, instead of the usual gate transconductance.

$$R_{OP} \approx (g_{m7} + g_{mb7})\, r_{o7} r_{o5} \tag{50}$$

$$R_{ON} \approx (g_{m9} + g_{mb9})\, r_{o9}\, (r_{o3}||r_{o11}) \tag{51}$$

Therefore, the gain of the folded cascode stage can be given by

$$|A_v| \approx g_{mb3} \left\{ [(g_{m7} + g_{mb7})\, r_{o7} r_{o5}] \;||\; [(g_{m9} + g_{mb9})\, r_{o9}\, (r_{o3}||r_{o11})] \right\} \tag{52}$$

Next, we proceed to discuss the frequency response of the folded cascode stage by considering the locations of its poles. By inspection, there is a dominant pole at the node connecting the drains of M_7 and M_9 (high impedance node), given by

$$p_1 \approx \frac{1}{2\pi R_{out} C_1} \tag{53}$$

where

$$C_1 = C_{gs13} + C_{db7} + C_{db9} + C_{gd7} + C_{gd9} \tag{54}$$

There is also a non-dominant pole introduced by the folded cascode, i.e. at the drains of M_{10} and M_{11}, given by

$$p_2 \approx \frac{g_{m9} + g_{mb9}}{2\pi C_2} \tag{55}$$

where

$$C_2 = C_{gs9} + C_{sb9} + C_{gd3} + C_{db3} + C_{gd11} + C_{db11} \tag{56}$$

Assuming that the folded cascode stage behaves like a first-order system, its GBW can be approximated by

$$GBW = \frac{g_{mb3}}{2\pi C_C} \tag{57}$$

The second stage (common source amplifier with active load) introduces a non-dominant pole and provides additional gain. Both the gain and non-dominant pole of the second stage has been discussed in Section 4.1.1. The Miller compensation capacitance has also been discussed in Section 4.1.3 and will not be repeated here.

4.3 Wide-swing Cascode Current Mirror
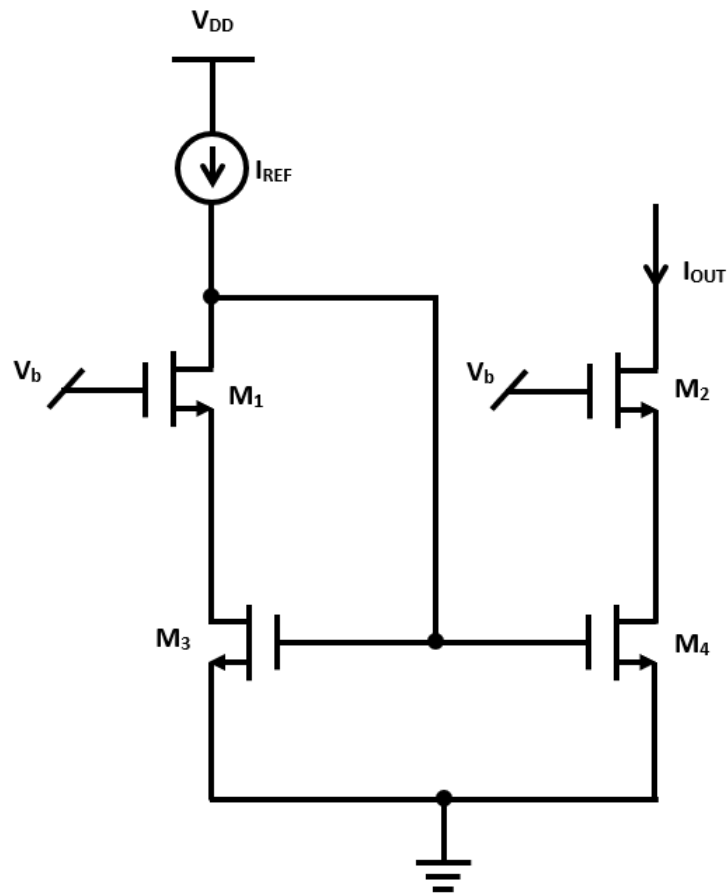
Figure 22: Wide-swing Cascode Current Mirror.

The wide-swing cascode current mirror in Figure 22 has two main advantages, wide output swing and a high output resistance. If properly biased, M_2 and M_4 consume minimal headroom while V_{DS1} and V_{DS3} are identical, allowing I_{out} to be an accurate copy of I_{REF}.

4.4 Summary

In this chapter, the two-stage op-amp, the folded cascode op-amp, and the wide-swing cascode current mirror were analysed, with emphasis on their small signal transfer function and stability. In particular, the influence of the Miller compensation capacitance on the op-amp stability and the various heuristics employed by designers to select a suitable value for the Miller compensation capacitance were discussed in detail. The operation of the unorthodox, bulk-driven technique and its implications were also introduced.

5 g_m/I_D-based Design and Implementation

Once the circuit topology has been analysed, the remaining design work to be done is to determine the transistor dimensions, bias conditions as well as various peripheral passive components like the Miller compensation capacitance. However, deciding on the transistor dimensions and bias currents is a multifaceted task with differing objectives. In addition, the design complexity grows rapidly with the number of transistors in the circuit. Therefore, the only way to proceed with determining the transistor dimensions and bias conditions is to arrange the objectives in a hierarchical order. To clarify this point further, note that some objectives are essential and must always be met. These objectives shape the specifications of the circuit. In comparison, other objectives are desirable but not essential. These non-essential objectives are regarded as the circuit attributes. For instance, the gain-bandwidth product and phase margin of an amplifier would be considered as specifications, whereas the power dissipation and area are considered attributes.

The specifications set the boundaries of a sizing space that represents satisfactory circuit designs. On the other hand, the attributes define the optimisation areas within the above sizing space. The crux of the design process lies in distinguishing the important transistors that influence the specifications from the secondary transistors. This will help to reduce the number of variables that the designer need to consider, making the design of complicated circuits possible.

In a sizing space, the g_m/I_D methodology can help the designer evaluate the trade-offs and guide him (her) to an optimal neighbourhood of designs[8]. The g_m/I_D methodology provides a systematic framework that significantly reduces the amount of time the designer has to spend on tweaking in Cadence Spectre.

This chapter explains the design principles and implementation of Trakoolwattana and Thanachayanont's potentiostat using the g_m/I_D methodology and pre-computed lookup tables. We begin by designing an intrinsic gain stage (IGS) to demonstrate the usage of the lookup tables as well as various auxiliary MATLAB functions. This is intended as an introductory example for those who are new to the g_m/I_D methodology to familiarise himself (herself) with the approach adopted in this FYP. Subsequently, the design and implementation of the op-amps and the wide-swing cascode current mirror will be explained in detail. Note that Trakoolwattana and Thanachayanont's potentiostat has been designed and simulated in both TSMC 180 nm and TSMC 65 nm technologies. In this chapter, the design and implementation process will be predominantly explained with circuits simulated in TSMC 180 nm technology. The general design process follows the same pattern for both technologies and will not be repeated for TSMC 65 nm, unless there is a distinct, worthwhile point to be made.

---

[8]Strictly speaking, to arrive at an optimal design point, one would need to define an objective function and employ optimisation techniques such as geometric programming. On its own, the g_m/I_D methodology does not include any optimisation techniques. Therefore, it cannot lead the designer to an optimal point, only to an optimal neighbourhood. Nevertheless, the the g_m/I_D methodology is amenable to optimisation techniques.

5.1 Sizing an Intrinsic Gain Stage (TSMC 180 nm)

Suppose we intend to size an IGS (find its width) with unity gain frequency (UGF) of 1 GHz when the load capacitance is 1 pF. Assume $L = 900$ nm, $g_m/I_D = 15$ S/A, $V_{DS} = 0.9$ V, and verify the results through Cadence Spectre simulations.

To do this, we need a SPICE-generated lookup table based on the TSMC 180 nm model file, as well as two helper functions `look_up` and `look_upVGS` provided by Boris Murmann on his website. `look_up` and `look_upVGS` are handy interpolation functions and their use will be demonstrated below. The internal details of `look_up` and `look_upVGS` can be found in Appendix 2 of [4].

From Equation 19,

$$g_m \approx 2\pi f_u C_L = 6.2832 \text{ mS} \tag{58}$$

$$I_D = \frac{g_m}{g_m/I_D} = \frac{0.0062832}{15} = 418.88 \ \mu\text{A} \tag{59}$$

To find the width, W, we need the current density $(I_D/W = J_D)$ which we can find using the `look_up` function.

```
JD = look_up(nch,'ID_W','GM_ID',15,'VDS',0.9,'L',0.9);
```

With the exception of the first 2 entries of `look_up`, every entry in single quotes is an input variable, and the consecutive entry (not in single quotes) is the actual input variable value being passed into the function. The first entry of `look_up` refers to the lookup table data structure. The second entry of `look_up` that is in single quotes is the output. If V_{GS} and V_{SB} are not specified, their default values are assumed.

- Default value of V_{DS} is $V_{DD}/2$ where V_{DD} is specified during the lookup table generation process.

- Default value of V_{SB} is 0.

- Default value of V_{GS} is the entire V_{GS} vector specified during the lookup table generation process.

We find $J_D = 1.2856 \times 10^{-6}$ A/μm, which yields $W = 325.82$ μm. The transistor is implemented with 20 fingers, each of width 16.291 μm. We can also use `look_upVGS` to derive the value of V_{GS} in our context.

```
VGS = look_upVGS(nch,'GM_ID',15,'VDS',0.9,'L',0.9);
```

`look_upVGS` follows similar syntax rules as `look_up`, except all entries are inputs. The only output is V_{GS}. V_{GS} is found to be 0.5233 V. Lastly, the DC gain can also be found using `look_up`.

```
DCgain = look_up(nch,'GM_GDS','GM_ID',15,'VDS',0.9,'L',0.9);
```

'GM_GDS' refers to g_m/g_{ds}, the equation for the DC gain of an intrinsic gain stage and is found to be 233.7026 = 47.373 dB. The design of this IGS is complete and we proceed to verify our design in Cadence Spectre with the test bench illustrated in Figure 23.
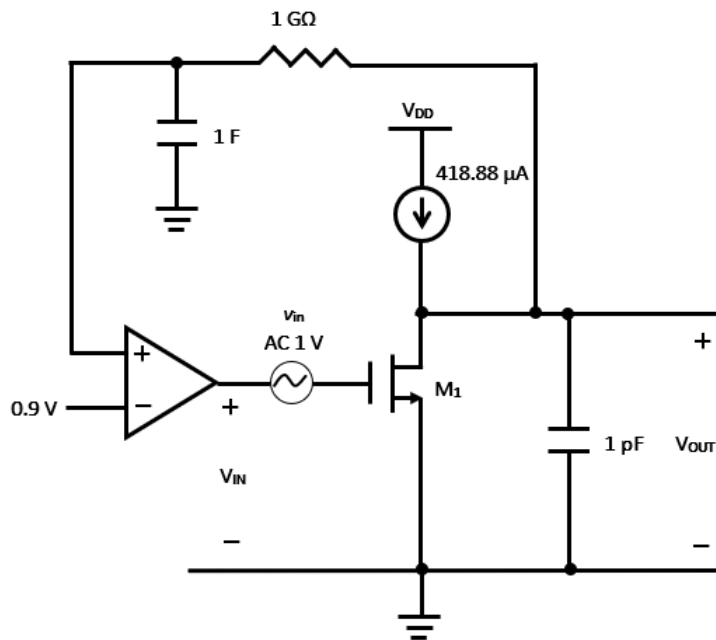
Figure 23: IGS test bench.

The test bench in Figure 23 works in the following manner. An ideal voltage op-amp[9] with 0.9 V tied to its inverting input terminal and its non-inverting terminal tied to a C-R network functions as a comparator. The C-R network was chosen such that it only passes DC signals. The overall test bench system employs negative feedback to produce a V_{GS1} value that ensures V_{DS1} is at 0.9 V, as intended. For instance, if V_{DS1} goes above 0.9 V, then the comparator output goes high, which leads to a larger V_{GS1}. Since I_{DS1} is fixed by the ideal current source, a rise in V_{GS1} will lead to a fall in V_{DS1}, countering the initial rise in V_{DS1}.

Table 1 shows that the g_m/I_D methodology can lead to accurate results without having to spend a lot time tweaking in Cadence Spectre. Our design is essentially right on-target.

	g_m/I_D-based Design	Cadence Spectre Simulation
g_m	6.2832 mS	6.3036 mS
g_m/I_D	15 S/A	15.0488 S/A
I_D	418.88 μA	418.88 μA
DC Gain	47.373 dB	47.337 dB
V_{GS}	0.5253 V	0.5234 V
V_{DS}	0.9 V	0.9 V

Table 1: IGS design results comparison

---

[9]Implemented in Cadence Spectre using a voltage-controlled voltage source.

5.2 g_m/I_D-based Design of the Two-Stage Op-amp (TSMC 180 nm)

We begin the design process for the two-stage op-amp by reminding ourselves of the desirable attributes that this op-amp should possess. The two-stage op-amp should have low power dissipation, small area and high SNR. As explained in Section 3.3, there is a trade-off between power and area. For low power dissipation, the transistor should be biased in the weak inversion region (large g_m/I_D). However, the transistors should be biased in the strong inversion region for a small total area. Clearly, the designer faces a dilemma.

In this design, the approach adopted was to bias transistors in the weak and moderate inversion regions. Placing transistors in the strong inversion region would lead to greater speed performance, which is not the crucial bottlenecks in our context. Furthermore, biasing transistors in the strong inversion would leave the designer with hardly any breathing room when optimising for other attributes like power dissipation and flicker noise. On the other hand, by placing transistors in a mix of weak and moderate inversion regions, the designer can optimise for gain, flicker noise and power dissipation. For the transistors that are biased in the moderate inversion region, the g_m/I_D sizing procedure mentioned in Section 3.2.1 is viable. However, for transistors that will be biased in the weak inversion region, a variant of the g_m/I_D-centric design procedure, i.e. the J_D-centric design procedure must be applied instead. In this design, a transistor with a g_m/I_D that is 19 S/A and above is considered to be in the weak inversion region. The J_D-centric design procedure is listed below.

1. Derive g_m from the design specifications.

2. Select the transistor lengths to satisfy g_m, speed, area, and matching requirements.

3. Decide on J_D for the transistors that need to be biased in the weak inversion region, taking the relevant trade-offs into consideration. For the transistors that are biased in the moderate or strong inversion regions, it is valid to decide on their g_m/I_D.

4. Determine g_m/I_D from J_D using the `look_up` function.

5. Derive $I_D = g_m/\left(g_m/I_D\right)$.

6. Derive $W = I_D/J_D$.

Firstly, we need to derive g_m from the design specifications. The transistor numbering in the two-stage op-amp is reproduced here for convenience.

- M_1, M_2 (NMOS) - Differential input transistor.

- M_3, M_4 (PMOS) - Current mirror transistor in the first stage.

- M_5 (NMOS) - Tail current source in the first stage.

- M_6 (PMOS) - Second stage common source amplifier.

- M_7 (NMOS) - Second stage current source.

Equation 34 shows that g_{m2} can be derived from the UGF specification, whereas g_{m6} is largely a function of g_{m2} and is set to be 10 times g_{m2} for stability reasons. In our context of pushing the circuit toward low power dissipation, the UGF of the two-stage op-amp is likely to be very low, e.g. kHz range. This UGF specification can be further tuned by the designer.

Subsequently, the design variables to decide upon are the lengths of the transistors. At first glance, there are seven pairs of widths and lengths On closer inspection, there are only five unique pairs of transistor widths and lengths to be determined since the current mirror and differential input stage are comprised of matched transistors. It is not always straightforward to decide on the lengths since there are many trade-offs. Therefore, the choice of lengths should be guided by the application context of the circuit. In this FYP, the lengths were decided based on the required gain and power dissipation constraints. This is only a rough estimate in the right direction and the designer may need to come back to this step and make modifications if necessary.

Next, the most important step of this design procedure is to determine the g_m/I_D or J_D of the transistors. It is undeniable that the two most important transistors with wide-reaching influence are M_2 and M_6, the differential input transistor and the second stage P-type common source amplifier respectively. M_2 and M_6 set the specifications of the circuit and will be biased in the weak inversion region. Their respective J_D will not be set to fixed values. Instead, a reasonable range of values will be considered, i.e. $(J_D)_2$ and $(J_D)_6$ will be vectors, not scalars. This allows the designer to evaluate the impact on the circuit performance as the two input variables ($(J_D)_2$ and $(J_D)_6$) are varied.

On the other hand, the current mirror transistors (M_3 and M_4), current source transistors (M_5 and M_7) are secondary and are biased in the moderate inversion. Therefore, their g_m/I_D will not be used as "tuning knobs" but will be fixed based on the circuit trade-offs. The choice of $\left(\frac{g_m}{I_D}\right)_3$, $\left(\frac{g_m}{I_D}\right)_5$, and $\left(\frac{g_m}{I_D}\right)_7$ is discussed below.

Consider the design of the second stage, a P-type common source amplifier with an active N-type load (Figure 24).

$$A_{v0} = -\frac{g_{m6}}{g_{ds6} + g_{ds7}} = -\frac{\left(\frac{g_m}{I_D}\right)_6}{\left(\frac{g_{ds}}{I_D}\right)_6 + \left(\frac{g_{ds}}{I_D}\right)_7} \tag{60}$$

From Equation 60, it can be seen that to increase the low frequency (DC) gain of this stage, $\left(\frac{g_{ds}}{I_D}\right)_7$ needs to be reduced.

The trade-offs between $\left(\frac{g_{ds}}{I_D}\right)_7$ and $\left(\frac{g_m}{I_D}\right)_7$ can be explored by plotting their relationship for a range of transistor lengths. From Figure 25, it can be seen that a small g_{ds}/I_D of the load transistor M_7 corresponds to it being biased in the strong inversion region (small value
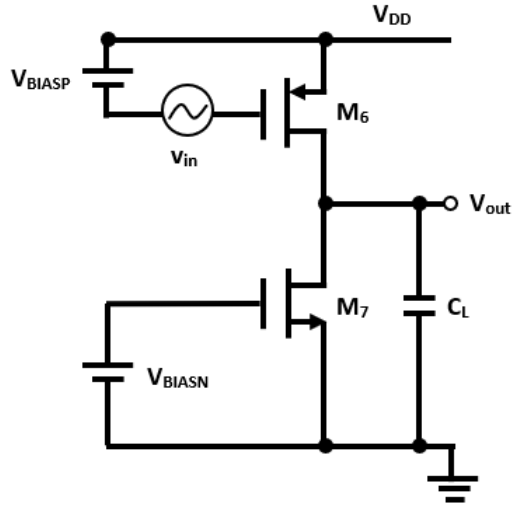
Figure 24: P-type common source amplifier with an active N-type load.

of g_m/I_D). However, with longer lengths, it is possible to achieve sufficiently small g_{ds}/I_D at moderate to weak inversion regions. Since it is desirable for the circuit to have less power dissipation, it would be better for M$_7$ to be biased in the moderate to weak inversion region. Thus, a compromise between gain and power dissipation was reached and M$_7$ should have a longer length with its $\left(\frac{g_m}{I_D}\right)_7$ chosen to be approximately 14-15 S/A.
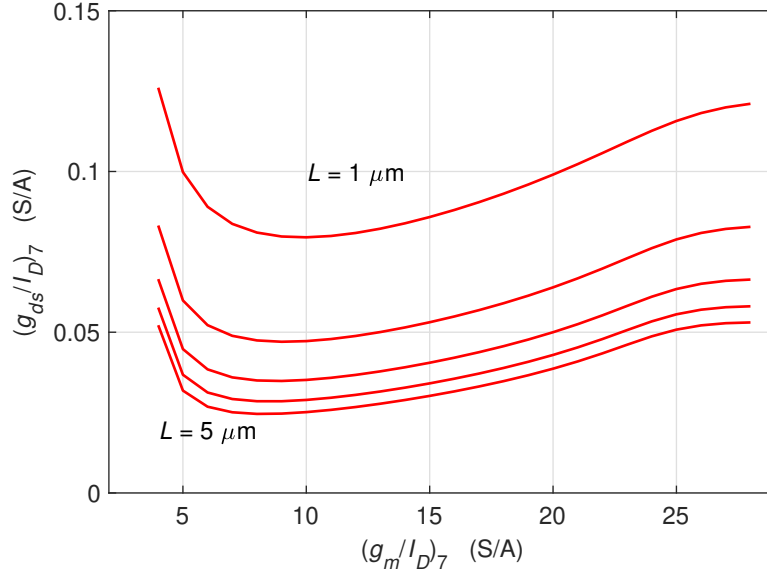
Figure 25: $\left(\frac{g_{ds}}{I_D}\right)_7$ against $\left(\frac{g_m}{I_D}\right)_7$ for a range of lengths; $V_{DS} = 0.5$ V; $V_{SB} = 0$ V; TSMC 180 nm.
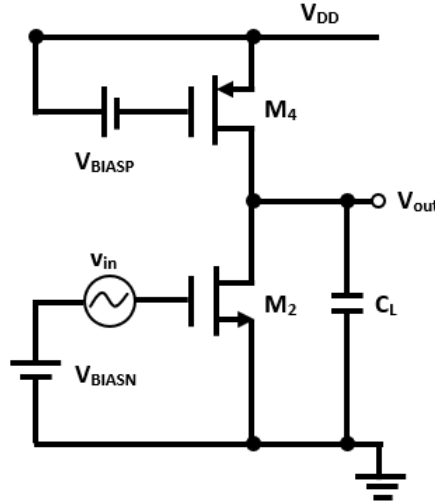
Figure 26: N-type common source amplifier with an active P-type load.

Moving on to the differential amplifier (N-type) with a current mirror load (P-type). We need to determine $\left(\frac{g_m}{I_D}\right)_4$. Observe that this half-circuit in Figure 26 is also a common source amplifier with an active load. Therefore, we can leverage on the previous analysis. In a similar fashion to M_7, we opt for M_4 to have a longer length and be biased in the moderate to weak inversion region. Figure 27 illustrates plots for a PMOS transistor in which these plots are slightly shifted downward compared to the NMOS plots in Figure 25. Therefore, for a comparable g_{ds}/I_D, we can afford to bias M_4 closer toward the weak inversion region, i.e. at a higher g_m/I_D of approximately 18-19 S/A.
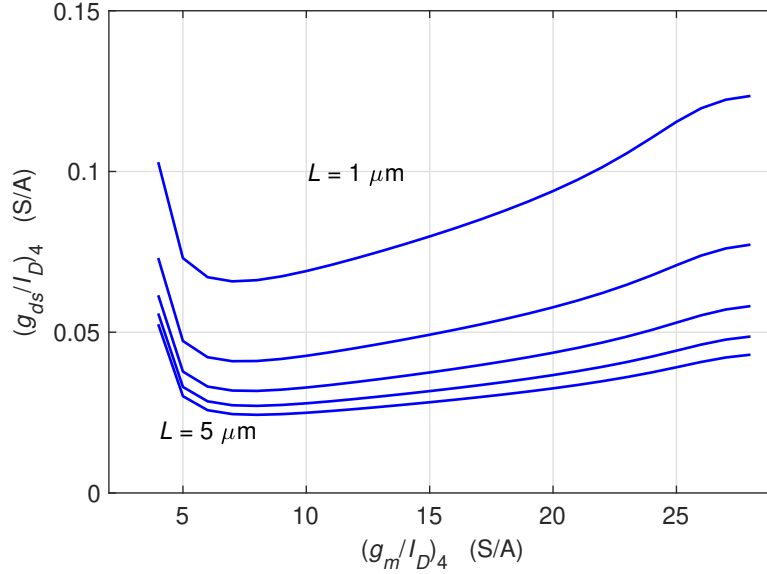
Figure 27: $\left(\frac{g_{ds}}{I_D}\right)_4$ against $\left(\frac{g_m}{I_D}\right)_4$ for a range of lengths; $V_{DS} = 0.5$ V; $V_{SB} = 0$ V; TSMC 180 nm.

The remaining transistor to consider is M_5, the tail current source of the differential amplifier. It is desirable for the transistor M_5 to have a long length. This is because M_5 should have a output resistance that is as large as possible to approximate an ideal current source. As a first-order approximation, the small signal output resistance of a transistor, r_O is proportional to the transistor length.

In addition, the transistor M_5 has to be able to support a much larger current, i.e. twice the amount of current flowing through M_4. With a much larger current (I_{D5}), $(g_m/I_D)_5$ will have to be pushed to a lower value compared to $(g_m/I_D)_4$. Taking the above into consideration, it was decided to bias M_5 near the boundary between the weak and strong inversion regions. $(g_m/I_D)_5$ was assigned a value of 10 S/A.

Having arrived at the values of g_m/I_D, the sizing of the two-stage op-amp is essentially finished. The Miller compensation capacitance is decided using the heuristics explained in Section 4.1.3. The remaining steps of the sizing procedure (steps 4-6) are straightforward and can be easily accomplished using the `look_up` and `look_upVGS` functions.

5.3 Implementation of the Two-Stage Op-amp (TSMC 180 nm)

In this section, the implementation of the two-stage op-amp is explained in detail. The implementation process is aided by a bespoke MATLAB script. This MATLAB script i) imports the SPICE-generated lookup table data (TSMC 180 nm), ii) plots a contour graph of constant gain, area and DC current curves, and iii) outputs design parameters such as transistor widths based on the point selected on the contour plot. The algorithm behind this MATLAB script will be examined. Finally, the design was simulated in Cadence Spectre (TSMC 180 nm) and the results were compared.

To begin the implementation, we must decide on the set of target specifications which is given in Table 2.

Unity Gain Frequency	80 kHz
Phase Margin	$\geq 45°$
Supply Voltage	1 V
Input Common Mode Voltage Level	0.4 V
Load Capacitance	5 pF

Table 2: Target specifications for two-stage op-amp.

In order to achieve a low power dissipation, the bandwidth must be sacrificed and the two-stage op-amp was designed for a low UGF of 80 kHz. The two-stage op-amp must also exhibit a phase margin of at least 45°.

Next, we fix the transistor lengths and the g_m/I_D ratios of the secondary transistors. The transistor lengths and $(g_m/I_D)_{4,5,7}$ are shown in Table 3. The lengths of the transistors were chosen to be quite long for sufficient gain and to keep the currents flowing through the transistors low, whereas the choice of $(g_m/I_D)_{4,5,7}$ has been explained in Section 5.2.

	M_2	M_4	M_5	M_6	M_7
Lengths (μm)	5	5	4	2	5
g_m/I_D (S/A)	-	18.5	10	-	14.5

Table 3: Transistor lengths and g_m/I_D for the two-stage op-amp.

The sizing procedure takes 2 input variables, which are J_{D2} and J_{D6}. To this end, we establish a 2-D sizing space and various constants in MATLAB as seen in Listing 1.

```
1   clc
2   close all
3   clear all
4
5   load 180v3nch % The lookup table data for TSMC 180 nm NMOS
6   load 180v3pch % The lookup table data for TSMC 180 nm PMOS
7
8   % M2 : Diff input pair (NMOS)
9   % M4 : Current mirror active load (PMOS)
10  % M5 : 1st stage bias tail current source (NMOS)
11  % M6 : 2nd stage common source amplifier (PMOS)
12  % M7 : 2nd stage active load transistor (NMOS)
13
14  % Constants ————————————————————————————————
15  U_T = 0.026;        % Thermal Voltage at 300 K
16  VDD = 1;            % Supply voltage V
17  fu = 80000;         % Unity gain frequency a.k.a GBW
18  C_L = 5e-12;        % Output (load) cap
19  Z = 10;             % Relative position of zero with respect to GBW.
20  NDP = 4;            % Relative position of non-dominant pole to GBW.
21
22  % Voltage is in V
23  VDS6 = 0.452;         % Voltage across 2nd stage common source
        amplifier
24  VG2 = 0.4;
25
26  % gm/ID is in S/A
27  gm_ID4 = 18.5;
28  gm_ID7 = 14.5;
29  gm_ID5 = 10;
30
31  % All lengths are in um units
32  L2 = 5;
33  L4 = 2;
34  L5 = 5;
35  L6 = 5;
36  L7 = 4;
37
38  X = logspace(-9,-5.5,100);        % (A/um)
39  Y = logspace(-9,-5.5,100);        % (A/um)
40  [JD_2, JD_6] = meshgrid(X,Y);
```

Listing 1: Generate 2-D sizing space and establish constants.

Moving on, we need to determine g_{m2} by using Equation 34. To do so, we have to decide on a suitable value for the Miller compensation capacitance. Recall that there are two heuristics that can be applied, Allen's heuristic and Jespers' heuristic.

Jespers' heuristic was implemented separately in a `for` loop. The MATLAB implementation of Jespers' heuristic is given in Listing 2.

```matlab
Cc = 0.5 * C_L ; % Initial guess for Cc is 2.5 pF
for k = 1:10
    gm2  = 2*pi*fu*Cc;
    gmoverID2 = transpose(look_up(nch,'GM_ID','ID_W',X,'VDS',VDS2,'
        VSB',VDS5,'L',L2)) ;

    ID2 = gm2./gmoverID2 ;
    W2 = ID2./X ;
    CDD2 = W1.*transpose(look_up(nch,'CDD_W','ID_W',X,'VDS',VDS2,'
        VSB',VDS5,'L',L2)) ; % CDD2 = Cgd2 + Cdb2

    ID4 = ID2 ;
    JD4 = look_up(pch,'ID_W','GM_ID',gm_ID4,'VDS',VGS4,'L',L4) ;
    W4 = ID4./JD4 ;
    CDD4 = W4.*look_up(pch,'CDD_W','GM_ID',gm_ID4,'VDS',VGS4,'L',L4)
         ; % CDD4 = Cgd4 + Cdb4

    gm6  = transpose(Z*gm2);
    gmoverID6 = look_up(pch,'GM_ID','ID_W',Y,'VDS',VDS6,'VSB',0,'L',
        L6) ;
    ID6 = gm6./gmoverID6 ;
    W6 = ID6./transpose(Y) ;
    CDD6 = W6.*look_up(pch,'CDD_W','ID_W',Y,'VDS',VDS2,'L',L2) ; %
        CDD6 = Cgd6 + Cdb6

    ID7 = ID6 ;
    JD7 = look_up(nch,'ID_W','GM_ID',gm_ID7,'VDS',VDD-VDS6,'L',L7) ;
    W7 = ID7./JD7 ;
    CDD7 = W7.*look_up(nch,'CDD_W','GM_ID',gm_ID7,'VDS',VDD-VDS6,'L
        ',L7) ; % CDD7 = Cgd7 + Cdb7

    C1 = CDD2 + CDD4 + transpose(W2.*look_up(pch,'CGS_W','ID_W',Y,'
        VDS',VDS2,'L',L2)) ;
    C2 = C_L + CDD6 + CDD7;
    Cc = 0.5*NDP/Z*(C1 + transpose(C2) + sqrt((C1+transpose(C2)).^2
        + C1.*transpose(C2)*4*Z/NDP));
```

```
29  end
```

Listing 2: For loop that implements Jespers' heuristic.

The `for` loop in Listing 2 is self-explanatory. A noteworthy point to mention is that in line 28 of Listing 2, the value of the Miller compensation capacitance is calculated using Equation 48.

In contrast, Allen's heuristic is much easier to implement. Allen's heuristic ($g_{m6} > 10g_{m2}$ and $C_C > 0.22C_L$), which was intended for a phase margin of 60° is able to provide some buffer since a phase margin of 45° is required. In fact, a more optimal circuit can be achieved in our context by relaxing the dominant pole condition in Allen's heuristic. This point will be explained subsequently.

With a load capacitance of 5 pF, we set the Miller compensation capacitance to be 1.2 pF ($0.24C_C$). However, in TSMC 180 nm technology, the closest capacitance to 1.2 pF that can be realised is 1.19288 pF, implemented with a unit capacitance of 119.288 fF and a multiplier of 10.

In this MATLAB script, the default option is to use Allen's heuristic, which is the simpler of the two. Jespers' heuristic acts as a backup that would only be invoked should Allen's heuristic fail to achieve the necessary phase margin.

As a result, g_{m2} is given by

$$g_{m2} = 2\pi C_C f_u = 2\pi \times 1.19288 \times 10^{-12} \times 80 \times 10^3 = 599.607 \times 10^{-9} \text{ S/A} \tag{61}$$

Having established g_{m2}, we proceed to use the `look_up` function to determine $(g_m/I_D)_2$. To do so, we need to determine V_{DS2} as given by Equation 62.

$$V_{DS2} = V_{DD} - V_{DS4} - V_{DS5} \tag{62}$$

Since $V_{DS4} = V_{DS3} = V_{GS3} = V_{GS4}$, we can find V_{GS4} using `look_upVGS` in an iterative `for` loop and giving V_{DS4} an initial guess value of 0.6 V (Listing 3). The value that V_{GS4} converges to will be used in all subsequent steps.

```
1  VGS4 = look_upVGS ( pch , 'GM_ID' , gm_ID4 , 'VDS' , 0.6 , 'L' , L4 ) ; % this is
       an initial guess
2
3  for i = 1:5
4          VGS4 = look_upVGS ( pch , 'GM_ID' , gm_ID4 , 'VDS' , VGS4 , 'L' , L4 ) ;
5  end
```

Listing 3: Converge to a guess for V_{GS4}.

On the other hand, V_{DS5} was chosen to be very low, in order to free up more of the supply voltage for the transistors M$_2$ and M$_4$. As a first guess, V_{DS5} was set to 100 mV, which can be easily tuned after performing some simulations in Cadence Spectre.

After defining g_{m2}, the `look_up` function can be used to find $(g_m/I_D)_2$ (see Listing 4). I_{D2} and W_2 follow naturally. Since $I_{D4} = I_{D2}$ and $I_{D5} = 2 \times I_{D2}$, W_4 and W_5 can also be found easily.

```matlab
gm2  = 599.607*10^(-9) ;  % S
gmoverID2 = transpose(look_up(nch,'GM_ID','ID_W',X,'VDS',VDS2,'VSB',
    VDS5,'L',L2)) ; % row vector
ID2 = gm2./gmoverID2 ; % row vector
ID2_m = repmat(ID2,length(X),1) ; % matrix
W2 = ID2./X ; % row vector
W2_m = repmat(W2,length(X),1) ; % matrix

ID4 = ID2 ; % row vector
ID4_m = repmat(ID4,length(X),1) ; % matrix
JD4 = look_up(pch,'ID_W','GM_ID',gm_ID4,'VDS',VGS4,'L',L4) ; %
    scalar
W4 = ID4./JD4 ; % row vector
W4_m = repmat(W4,length(X),1) ; % matrix

ID5 = 2*ID1 ;  % a row vector
ID5_m = repmat(ID5,length(X),1) ; %rows are copies of each other
JD5 = look_up(nch,'ID_W','GM_ID',gm_ID5,'VDS',VDS5,'L',L5) ; % a
    scalar
W5 = ID5./JD5 ; % a row vector
W5_m = repmat(W5,length(X),1) ; %rows are copies of each other
```

Listing 4: Current and width matrices for M$_2$, M$_4$, and M$_5$.

It must be emphasised that J_{D2} is intended to be the input variable on the horizontal axis of the 2-D sizing space. Thus, I_{D2}, W_2 ,I_{D4}, W_4, I_{D5} and W_5 which are largely functions of J_{D2} are defined as row vectors. The current and width row vectors must be extended into matrices. This is because in order to plot the transistor current and width contour curves, the MATLAB command `contour` must be used. The MATLAB command `contour` requires as an input argument, a 2-D array that represents the grid over which the contour curves are plotted. I_{D2}, W_2 ,I_{D4}, and W_4 can be easily extended into matrices by using the `repmat` command in MATLAB. `repmat(A,c,1)` returns a matrix that is formed by stacking the row vector `A` on top of each other for `c` repetitions.

The transistor currents and widths in the second stage of the amplifier can be defined in a similar fashion (see Listing 5). As a first estimate, V_{DS6} is defined at $\frac{V_{DD}}{2}$ for maximum voltage headroom at the output, which can be adjusted after performing some simulations in Cadence

Spectre.

Using the dominant pole condition in Equation 33, we define $g_{m6} = 10 \times g_{m2}$. The main difference is that J_{D6} is intended to be the input variable on the vertical axis of the 2-D sizing space. Therefore, the current and width vectors of M_6 and M_7 are column vectors and the corresponding matrices are formed by column repetition.

```
gm6 = 10*gm2 ;
gmoverID6 = look_up(pch, 'GM_ID', 'ID_W',Y, 'VDS',VDS6, 'VSB',0 , 'L',L6)
    ; % col vector
ID6 = gm6./gmoverID6 ; % col vector
ID6_m = repmat(ID6,1,length(Y)) ; % matrix
W6 = ID6./transpose(Y) ; % col vector
W6_m = repmat(W6,1,length(Y)) ; % matrix

ID7 = ID6 ; % col vector
ID7_m = repmat(ID7,1,length(Y)) ; % matrix
JD7 = look_up(nch, 'ID_W', 'GM_ID',gm_ID7, 'VDS',VDD-VDS6, 'L',L7) ; %
    scalar
W7 = ID7./JD7 ; % col vector
W7_m = repmat(W7,1,length(Y)) ; % matrix
```

Listing 5: Current and width matrices for M_6 and M_7.

Next, as shown in Listing 6, the low frequency gain curves of the first stage and second stage can be derived by using the equation for the low frequency gain of the two-stage op-amp (see Equation 27).

```
gds_ID2 = transpose(look_up(nch, 'GDS_ID', 'ID_W',X, 'VDS',VDS2_test, '
    VSB',VS2_test, 'L',L2));
gds_ID2_m = repmat(gds_ID2,length(X),1) ;
gds_ID4 = look_up(pch, 'GDS_ID', 'GM_ID',gm_ID4, 'VDS',VGS4, 'L',L4);
AdB1 = 20*log10(gmoverID2./(gds_ID2 + gds_ID4)) ; % First stage gain
AdB1_m = repmat(AdB1,length(gmoverID1),1) ;

gds_ID6 = look_up(pch, 'GDS_ID', 'ID_W',Y, 'VDS',VDS6, 'L',L6);
gds_ID7 = look_up(nch, 'GDS_ID', 'GM_ID',gm_ID7, 'VDS',VDD-VDS6, 'L',L7)
    ;
AdB2 = 20*log10(gmoverID6./(gds_ID6 + gds_ID7)) ; % Second stage
    gain
AdB2_m = repmat(AdB2,1,length(gmoverID2)) ;
```

Listing 6: Low frequency gain of the first and second stages.

Finally, having defined or derived all the design variables, we proceed with plotting the contour curves of the total DC current (sum of all drain currents), total gain, and the total active area[10] (see Listing 7).

The MATLAB command `ginput` provides a handy way for the designer to select a point on the contour plot by using his (her) mouse. The coordinates of this selected point is designated as `[P1,P2]`. `[P1,P2]` is used as the query points that is fed into the standard MATLAB 2-D interpolation function (`interp2`) for interpolating the total current and gain values of the selected point. `[P1,P2]` is also used to calculate the widths, from which the total active area of this selected design is calculated.

```matlab
1  Isupply_m = ID5_m + ID7_m;
2  H = contour(X,Y,Isupply_m,'r');
3  clabel(H); xlabel('JD2'); ylabel('JD6');
4  hold
5
6  Sactive = 2*W2_m*L2 + W4_m*L4 +2*W5_m*L5 +W6_m*L6 + W7_m*L7 ;
7  S = contour(X,Y,Sactive,'b'); clabel(S);
8
9  total_gain_dB = AdB1_m+AdB2_m ;
10 G = contour(X,Y,total_gain_dB,'m—') ;
11 clabel(G);
12
13 % % % SELECT a point ————————————————————————————
14 % % % The design parameters with no semicolon suppression will have
       their values printed out in the MATLAB Command Window
15 [P1,P2] = ginput(1); plot(P1,P2,'kx') ; hold
16 [P1,P2]
17 % The suffix x indicates that the transistor parameter, be it width
       or current is that of the selected point.
18 ID2x = 0.5*interp2(X,Y,ID5_m,P1,P2,'makima') ;
19 % Note that P1 and P2 are current densities (ID/W)!
20 W2x = ID2x/P1 ;
21 ID4x = ID2x ;
22 W4x = ID4x/JD4 ;
23 ID6x = interp2(X,Y,ID6_m,P1,P2,'makima') ;
24 W6x = ID6x/P2 ;
25 ID7x = ID6x ;
26 W7x = ID7x/JD7 ;
27 ID5x = 2*ID2x ;
28 W5x = ID5x/JD5 ;
29 W = [W1x W2x W3x W4x W5x]
30 L = [L1 L2 L3 L4 L5]
```

---
[10]The active area is defined by summing the product of the length and width of each transistor in the circuit.

```
31  total_area = 2*L1*W1x + L2*W2x + 2*L3*W3x + L4*W4x + L5*W5x
32  Isupplyo = 2*ID1x + ID2x;
33  I = [ID1x 2*ID1x ID2x Isupplyx]
34  power = VDD*I
35  A1dBx   = interp2(X,AdB1_m,P1,P2,'makima') ;
36  A2dBx   = interp2(X,Y,AdB2_m,P1,P2,'makima') ;
37  gaindB = [A1dBx A2dBx A1dBx+A2dBx]
```

Listing 7: Plotting the contour curves and selecting a point.

Running the MATLAB sizing procedure yields the contour plot shown in Figure 28 and the results are summarised in Table 4. The contour plots can help the designer to visualise the pros and cons of his (her) design choices. From Figure 28, it can be seen that selecting a point with smaller J_{D2} and J_{D6} values, which means placing M_2 and M_6 in the weak inversion region leads to lower power dissipation. The gain also improves, however, the total transistor area increases.
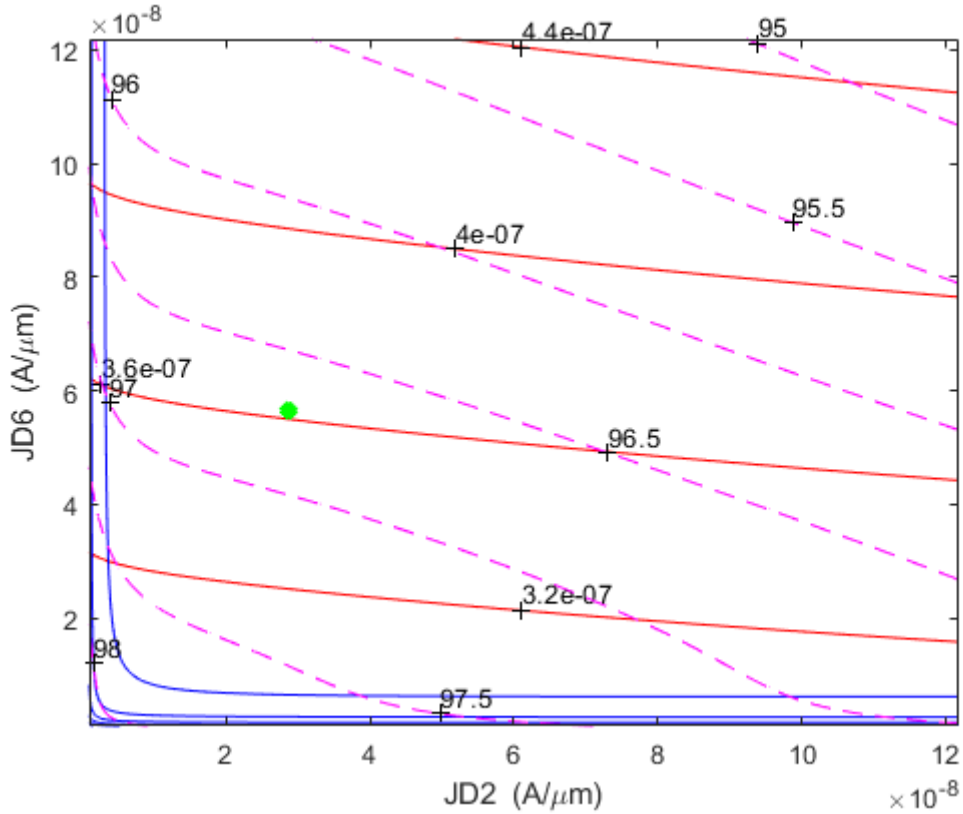

Figure 28: Total DC current (nW, red), total active area (μm^2, blue), and total gain (dB, dashed magenta) contour curves. The selected point is marked by the green dot.

	M$_2$ - 0.9500
	M$_4$ - 0.9482
Width (μm)	M$_5$ - 0.8788
	M$_6$ - 5.4322
	M$_7$ - 1.2064
Stage 1 Current in Both Branches (nA)	54.374
Stage 2 Current (nA)	307.710
Total Power Dissipation (nW)	362.084
Stage 1 Gain (dB)	49.503
Stage 2 Gain (dB)	47.203
Total Gain (dB)	96.706

Table 4: Two-stage op-amp design results for $g_{m6} = 10 \times g_{m2}$.

The set of transistor widths in Table 4 was simulated in Cadence Spectre. The phase margin achieved is 59.16°, which is essentially right on target as predicted by Allen's heuristic. Despite only dissipating a power of several hundred nW, this design can be further optimised for power. Note that the second stage is dissipating a disproportionately large amount of power compared to the first stage. In an attempt to reduce the power dissipation of the second stage, the condition, $g_{m6} > 10 \times g_{m2}$ was relaxed. This is justified since the phase margin has been over-designed.

The dominant pole condition was relaxed in the following way.

$$g_{m6} = M \times g_{m2}, \ M \in \{7, 7.5, 8, 8.5, 9\} \tag{63}$$

where M is a multiplicative constant. It has been found that for $M < 7$, there was a large deviation between the simulated UGF and the desired value (80 kHz). In addition, the minimum phase margin requirement (45°) would not be met. On the other hand, for $M > 9$, the power dissipation in the second stage only changed marginally, which defeated the purpose of relaxing the dominant pole condition.

The relaxed dominant pole condition in Equation 63 can be easily implemented in MATLAB by computing the 5 possible values of g_{m6} and using them in the subsequent calculations to obtain 5 possible designs. For a fair comparison, the design point that is marked by the green dot in Figure 28 was used in the 5 possible designs.

We find that $g_{m6} = 7.5 \times g_{m2}$ to be a good compromise that yields a reduced power dissipation in the second stage, an acceptable phase margin and unity gain frequency. Using this value of g_{m6} and running the MATLAB script once more, we arrive at the updated contour plot (Figure 29) and the results are collated in Table 5. The plots in Figure 29 have largely the same form as that in Figure 28, however, on closer inspection, it can be seen that the power dissipation has reduced. The values of the DC current (red) curves have reduced.
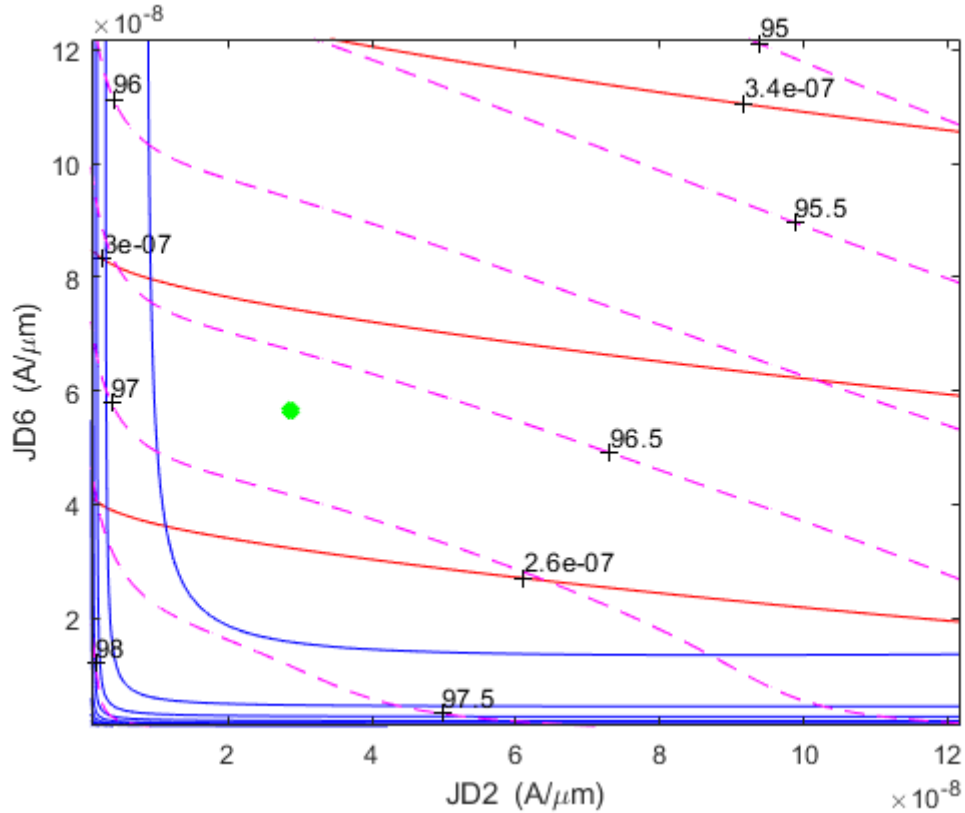
Figure 29: Optimised design contour curves. Total DC current (nW, red), total active area (μm^2, blue), and total gain (dB, dashed magenta) contour curves. The selected point is marked by the green dot.

As seen in Table 5, the predicted widths were given up to 4 decimal places. In practice, the analogue designer has to bear in mind transistor matching concerns and non-idealities in layout. From a practical point of view, it is better to work with transistors with dimensions that are rounded to the closest integer or at most to one decimal place. As such, the transistor widths were rounded up. In any case, the rounding up of transistor width values is only marginal.

In addition, it can be seen that the predicted results for the current and gain values are very close to the Cadence Spectre simulation results. It is worthwhile pointing out that the predicted unity gain frequency will always be slightly bigger than the actual unity gain frequency. This is due to the straight-line approximation that is inherent to Bode plots. This discrepancy is illustrated in Figure 30.

	Predicted Results	Cadence Spectre Simulation Results
Width (μm)	M_2 - 0.9500	M_2 - 1
	M_4 - 0.9482	M_4 - 1
	M_5 - 0.8788	M_5 - 1
	M_6 - 4.0512	M_6 - 4
	M_7 - 0.8997	M_7 - 1
Stage 1 Current in Both Branches (nA)	54.374	57.236
Stage 2 Current (nA)	229.490	226.584
Total Power Dissipation (nW)	283.864	283.820
Stage 1 Gain (dB)	49.503	49.759
Stage 2 Gain (dB)	47.203	47.252
Total Gain (dB)	96.706	97.011
Phase Margin	$\geq 45°$	53.990
Unity Gain Frequency (kHz)	80	70.918

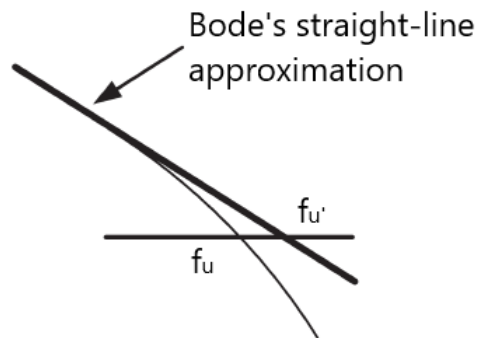Table 5: Optimised design ($g_{m6} = 7.5 \times g_{m2}$) simulation results.

Figure 30: Discrepancy between predicted and actual unity gain frequencies.

5.4 g_m/I_D-based Design of the Folded Cascode Op-amp (TSMC 180 nm)

In this section, the design process for the folded cascode op-amp is explained in detail. The folded cascode op-amp used by Trakoolwattana and Thanachayanont actually consists of two stages. The first stage is the folded cascode stage and the second stage is a standard common source amplifier with an active load. Nevertheless, this op-amp will be referred to as the folded cascode op-amp in this report to distinguish it from the standard two-stage op-amp. The design process for this folded cascode op-amp is largely similar to that of the standard two-stage op-amp. As such, we can leverage on some of the previous analysis. The folded cascode op-amp transistor numbering is reproduced here for convenience.

- M_1 (PMOS) - Folded cascode stage tail current source.

- M_2, M_3 (PMOS) - Bulk-driven input differential transistors.

- M_4 - M_7 (PMOS) - PMOS cascode current source in the first stage.

- M_8, M_9 (NMOS) - Cascode transistors to the input differential transistors.

- M_{10}, M_{11} (NMOS) - Tail current source to the cascode transistors M_8, M_9.

- M_{12} (PMOS) - Second stage active load.

- M_{13} (NMOS) - Second stage common source amplifier.

The design equations (see Section 4.2.1) show that the folded cascode stage can be divided into 3 groups of transistors. The first group comprises of the input transistors M_2 and M_3. The second group consists of M_8 and M_9, which form the cascode transistors to the input transistors. The third group of transistors consist of M_4-M_7 and M_{10}-M_{11}. These transistors act as current sources and active loads. Out of the 3 groups of transistors, the input transistors (M_2, M_3) and the cascode transistors (M_8, M_9) are more important.

The transconductance of the input transistors (M_2, M_3) and the second stage common source amplifier (M_{13}) can be derived by the UGF specification. In order to achieve low power dissipation, the bandwidth of the circuit must be sacrificed and we designed the folded cascode to have a UGF in the low kHz range. In a similar fashion to the previous op-amp, the transistor lengths were chosen to be rather long to provide sufficient gain and dissipate low power. Next, we decide on the various g_m/I_D. The input transistors (M_2, M_3), cascode transistors (M_8, M_9), current source transistors (M_{10}, M_{11}) and the second stage common source amplifier (M_{13}) were biased in the weak inversion region for greater gain and lower power dissipation. The transistors M_1, M_4 - M_7, M_{12} act as current sources. By leveraging on the analysis in Section 5.2, we bias these current source transistors to be in the moderate inversion region. The Miller compensation capacitance was chosen to be 1.2 pF, which is equivalent to 0.24 times of the load capacitance.

5.5   Implementation of the Folded Cascode Op-amp (TSMC 180 nm)

In this section, the implementation of the folded cascode op-amp is discussed. In principle, it is possible to write another MATLAB script that plots contour curves and have the designer select a design point. However, in this folded cascode op-amp, there are 3 transistors that constitute the "tuning" knobs. The 3 most important transistors are M_3, M_9 and M_{13}. In this case, it would be very difficult to plot a function of three independent variables. A possible solution is to define the 3 independent variables as 3 mutually perpendicular axes and encode the output variable using a color map. Nevertheless, the relevant trade-off information would not be easily seen from the plots, defeating the purpose of plotting the curves in the first place.

The implementation process of this folded cascode op-amp is rather different from that of the standard two-stage op-amp. A large portion of the implementation process of the standard two-stage op-amp was dedicated to preserving the relevant degrees of freedom in the circuit design by setting the current densities of the input and common source transistor as independent axes. The advantage of this approach is that useful plots to aid the designer can be obtained. On the other hand, the folded cascode op-amp is much more complex. Therefore, the focus of the implementation process of the folded cascode op-amp is to reduce the number of design variables by setting some design variables to known values. This makes the design of this op-amp more manageable. The choice of which design variables to fix and which to leave "free" requires intuition and sound judgement and will be explained in this section.

To proceed with the implementation of the folded cascode op-amp, a more detailed list of specifications (Table 6) was decided upon.

Stage 1 Current in Both Branches	≤ 100 nA
Stage 2 Current	≤ 100 nA
Total Gain	≥ 80 dB
Phase Margin	$\geq 45°$
Unity Gain Frequency	≥ 15 kHz
Supply Voltage	1 V
Input Common Mode Voltage Level	0.4 V
Load Capacitance	5 pF

Table 6: Target specifications for the folded cascode op-amp.

We aim to design the folded cascode op-amp for a very ambitious power dissipation target (less than 200 nW) as seen in Table 6. Unlike the standard two-stage op-amp, the UGF specification is not fixed at a single value but is given as a lower bound to grant the designer more freedom in pushing the power dissipation down. However, the phase margin requirement remains the same for stability reasons.

Next, the lengths and g_m/I_D ratios of the transistors were fixed, as seen in Table 7.

	M$_1$	M$_4$, M$_5$	M$_6$, M$_7$	M$_{12}$
Lengths (μm)	5	15	10	4
g_m/I_D S/A	15	14.5	17	15

Table 7: Transistor lengths and g_m/I_D for the folded cascode op-amp

We proceed to allocate the current flowing through the folded cascode stage. Observe that the current flowing through a single branch in the folded cascode stage, i.e. the current flowing through transistor M$_{11}$ is actually the sum of the currents flowing through M$_3$ and M$_9$. We allocate 28 nA to flow through M$_3$ and 17 nA to flow through M$_9$ to result in a sum of 45 nA flowing through the single branch. We would expect the Cadence-simulated currents to be very close but exactly equal to the allocated values. These allocated currents are meant as "anchors" to help the designer proceed.

Having decided on I_{D3}, I_{D9}, and I_{D11}, it is possible to use `look_up` to derive their widths computationally. However, in the interest of time, these widths were fixed a priori to rather large values. Recall that, M$_3$, M$_9$ and M$_{11}$ are intended to be biased in the weak inversion region for low power dissipation. From Figure 16, we can infer that a small value of J_D corresponds to a large value of g_m/I_D and width. Thus, it is reasonable and intuitive to set M$_3$, M$_9$ and M$_{11}$ to have rather large widths. For instance, the widths of M$_3$ and M$_9$ were set to be rather large, at 20 μm and 15 μm respectively. The width of M$_{11}$ was set to be 3 μm. The bias voltage V_{B2} and V_{B1} were adjusted such that I_{D3}, I_{D9}, I_{D11} are very close to 28 nA, 17 nA and 45 nA respectively.

After determining the dimensions for the transistors that are biased in weak inversion, the remaining transistors in the folded cascode stage can be determined using `look_up`.

For transistor M$_7$, its g_m/I_D was set to 17 S/A, and the current flowing through it should be very close to 17 nA.

The MATLAB function

```
JD7 = look_up(pch, 'ID_W', 'GM_ID', 17, 'VDS', VDS7, 'VSB', 0, 'L', 10)
```

yields J_{D7} to be 1.8597×10^{-8}A/μm.

As such, W_7 can be given by Equation 64.

$$W_7 = (I_D/J_D)_7 = \frac{17 \times 10^{-9}\ \text{A}}{1.8597 \times 10^{-8}\text{A}/\mu\text{m}} = 0.9141\mu\text{m} \tag{64}$$

In a similar fashion, for transistor M_5, its g_m/I_D was set to 14.5 S/A, and the current flowing through it should be very close to 17 nA.

The MATLAB function

```
JD5 = look_up(pch, 'ID_W', 'GM_ID', 14.5, 'VDS', VDS5, 'VSB', 0, 'L', 15)
```

yields J_{D5} to be $1.8484 \times 10^{-8} \text{A}/\mu\text{m}$.

As such, W_5 can be given by Equation 65.

$$W_5 = (I_D/J_D)_5 = \frac{17 \times 10^{-9} \text{ A}}{1.8484 \times 10^{-8} \text{A}/\mu\text{m}} = 0.9197\mu\text{m} \tag{65}$$

For transistor M_1, its g_m/I_D was set to 15 S/A, and the current flowing through it should be very close to 56 nA.

The MATLAB function

```
JD1 = look_up(pch, 'ID_W', 'GM_ID', 15, 'VDS', VDS1, 'VSB', 0,'L', 5)
```

yields J_{D1} to be $5.5098 \times 10^{-8} \text{A}/\mu\text{m}$.

As such, W_1 can be given by Equation 66.

$$W_1 = (I_D/J_D)_1 = \frac{56 \times 10^{-9} \text{ A}}{5.5098 \times 10^{-8} \text{A}/\mu\text{m}} = 1.016\mu\text{m} \tag{66}$$

Having obtained the widths of the transistors in the folded cascode stage (Table 8), we simulate the design in Cadence Spectre.

	M_1	M_2, M_3	M_4, M_5	M_6, M_7	M_8, M_9	M_{10}, M_{11}
Widths (μm)	$1.0160 \approx 1$	20	$0.9197 \approx 1$	$0.9141 \approx 1$	15	3

Table 8: Transistor widths for the folded cascode stage.

After some tuning, it was found that $V_{B1} = 0.51$V and $V_{B2} = 0.4$V yielded satisfactory currents.

	Desired values	Actual simulated values
I_D (nA)	M_1 - 56 M_2, M_3 - 28 M_4 - M_9 - 17 M_{10}, M_{11} - 45	M_1 - 55.714 M_2, M_3 - 27.857 M_4 - M_9 - 18.449 M_{10}, M_{11} - 46.307
g_m/I_D (S/A)	M_1 - 15 M_4, M_5 - 14.5 M_6, M_7 - 17 M_{12} - 15	M_1 - 15.114 M_2, M_3 - 26.415 M_4, M_5 - 14.508 M_7, M_7 - 17.080 M_8, M_9 - 26.438 M_{10}, M_{11} - 23.231

Table 9: Desired values versus actual simulated values for the folded cascode stage.

Moving on, we begin the design of the second stage. The active load M_{12} should be biased in the moderate inversion region whereas the common source amplifier should be biased in the weak inversion region. The choice of g_{m13} is largely influenced by the value of g_{mb3} and the UGF specification. From the simulation of the folded cascode stage, g_{mb3} was found to be 281.127 nS. As an rough approximation, the UGF of the entire op-amp can be calculated using Equation 57 to be 37.273 kHz, which is well beyond our UGF lower bound. Having obtained the value of g_{m3}, the value of g_{m13} is determined based on the dominant pole condition.

From Section 5.3, we observed that the more relaxed the dominant pole condition is, the less power the second stage dissipates. In Section 5.3, we set the second stage g_m to be 7.5 times of the g_m of the first stage and reached a decent compromise between the second stage power dissipation and the degree of agreement between the desired and simulated UGF values. However, in this design, our UGF specification is not so strict and thus, we aim to reduce the second stage power dissipation drastically.

The dominant pole condition is relaxed in the following manner.

$$g_{m13} = M \times g_{m3}, \ M \in \{6, 6.5, 7, 7.5, 8\} \tag{67}$$

It was found that $g_{m13} \approx 6 \times g_{m3}$ yielded a very low power dissipation in the second stage and fulfilling the phase margin requirement at the same time.

The performance results of the optimised design can be seen in Table 10.

	Desired values	Actual simulated values
Width (μm)	M_1 - 1.0160 M_2, M_3 - 20 M_4, M_5 - 0.9427 M_6, M_7 - 0.9141 M_8, M_9 - 15 M_{10}, M_{11} - 3 M_{12} - 1.0870 M_{13} - 2.1250	M_1 - 1 M_2, M_3 - 20 M_4, M_5 - 1 M_6, M_7 - 1 M_8, M_9 - 15 M_{10}, M_{11} - 3 M_{12} - 1 M_{13} - 2
Stage 1 Current in Both Branches (nA)	≤ 100	92.614
Stage 2 Current (nA)	≤ 100	69.068
Total Power Dissipation (nW)	≤ 200	161.682
Stage 1 Gain (dB)	-	70.277
Stage 2 Gain (dB)	-	44.890
Total Gain (dB)	≥ 80	115.167
Phase Margin	$\geq 45°$	47.680°
Unity Gain Frequency (kHz)	≥ 15	31.108

Table 10: Optimised folded cascode op-amp design results.

5.6 Summary

In this chapter, the design and implementation processes of the two-stage op-amp and the folded cascode op-amp were discussed. The use of the SPICE-generated lookup table data as well as two auxiliary functions `look_up` and `look_upVGS` was introduced with the IGS example. Next, the implementation of the two-stage op-amp was discussed, with emphasis on the MATLAB script that offers user interaction and plots contour curves. The implementation of the folded cascode op-amp followed a different, more direct approach and was elaborated. In implementing both op-amps, the use of the g_m/I_D methodology as the guiding principle yielded satisfactory circuit designs that showed a close agreement between the simulated and intended results.

6   Results and Evaluation

In this chapter, the performance of the original op-amps as designed by Trakoolwattana and Thanachayanont and that of the op-amps designed using the g_m/I_D methodology (TSMC 180 nm and TSMC 65 nm) are compared. A Figure of Merit (FoM) is proposed to quantify the performance of the op-amps based on their power dissipation, area and equivalent input noise.

6.1   Figure of Merit

Power dissipation, area and equivalent input noise can be considered as three mutually orthogonal vectors in a 3-D space. Ideally, the op-amp should have very small power dissipation, equivalent input noise and area. In other words, if we represent the op-amp's power dissipation, equivalent input noise and area as a point in the 3-D space, this point should be as close to the origin $\begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$ as possible.

Therefore, we can leverage on this observation and use a form of distance as a metric. The proposed FoM is a weighted Euclidean distance in the 3-D space and should be as small as possible. A standard Euclidean distance would not be appropriate because the power dissipation, area and equivalent input noise have vastly different orders of magnitude. For instance, the power dissipation achieved is typically in the nW range, whereas the equivalent input noise is in the μV range. As such, the equivalent input noise term will dominate the standard Euclidean distance, and diminish the influence of power dissipation and area. The proposed FoM is given in Equation 68.

$$FoM_i = \sqrt{\left(\frac{P_i}{\sum_i^n P_i} \times 100 \right)^2 + \left(\frac{N_i}{\sum_i^n N_i} \times 100 \right)^2 + \left(\frac{A_i}{\sum_i^n A_i} \times 100 \right)^2} \qquad (68)$$

where P, N, and A represent power dissipation, equivalent input noise and area respectively; n is the number of op-amps that are being compared. It must be emphasised that this FoM should be used to compare op-amps of the same topology and comparable performance.

To demonstrate the use of this FoM, consider the following hypothetical example. There are 2 op-amps that have the same topology but different transistor dimensions. The power dissipation, area and equivalent input noise of op-amp 1 are 100 nW, 200 μm^2, and $50 \times 10^{-6} \mu$V respectively. The power dissipation, area and equivalent input noise of op-amp 2 are 400 nW, 250 μm^2, and 35 μm^2 respectively.

FoM_1 and FoM_2 are calculated as shown below.

$$FoM_1 = \left(\left(\frac{100 \times 10^{-9}}{(100 + 400) \times 10^{-9}} \right)^2 + \left(\frac{50 \times 10^{-6}}{(50 + 35) \times 10^{-6}} \right)^2 + \left(\frac{200 \times 10^{-12}}{(200 + 250) \times 10^{-12}} \right)^2 \right)^{\frac{1}{2}}$$

$$\tag{69}$$

$$= 0.764 \tag{70}$$

$$FoM_2 = \left(\left(\frac{400 \times 10^{-9}}{(100 + 400) \times 10^{-9}} \right)^2 + \left(\frac{35 \times 10^{-6}}{(50 + 35) \times 10^{-6}} \right)^2 + \left(\frac{250 \times 10^{-12}}{(200 + 250) \times 10^{-12}} \right)^2 \right)^{\frac{1}{2}}$$

$$\tag{71}$$

$$= 1.057 \tag{72}$$

Since FoM_1 is smaller than FoM_2, we can conclude that op-amp 1 is better designed for our application. Note that this does not mean that op-amp 1 is superior to op-amp 2 in every aspect. Many other performance parameters such as gain, settling time, slew rate are not considered in this FoM. This FoM has been tailored according to the requirements of this FYP and it is assumed implicitly that the other relevant performance parameters of the op-amps are comparable. Clearly, an op-amp with 2 dB gain cannot be compared with another op-amp that has 80 dB gain using the proposed FoM. A universal figure of merit that captures every aspect of an op-amp's performance does not exist. Therefore, it is justified that the op-amp figure of merit is application-specific.

6.2 Two-Stage Op-amp Results (TSMC 180 nm)

The simulated results of the original two-stage op-amp designed by Trakoolwattana and Thanachayanont are collated in Table 11.

Stage 1 Current in Both Branches (nA)	64.159
Stage 2 Current (nA)	261.749
Total Power Dissipation (nW)	325.908
Stage 1 Gain (dB)	49.847
Stage 2 Gain (dB)	46.503
Total Gain (dB)	96.350
Phase Margin	64.007°
Unity Gain Frequency (kHz)	52.409
Total Active Area Without Miller Capacitance (μm^2)	34
2 pF Miller Capacitance Area (μm^2)	968.256
Total Input Referred Noise (μV, 0.001 Hz - 10 kHz)	20.910

Table 11: Original two-stage op-amp (TSMC 180 nm) results.

The distribution of power dissipation and active area (excluding the Miller compensation capacitance area)[11] between the two stages can be seen from Figure 31 and Figure 32.

Power Dissipation (nW)
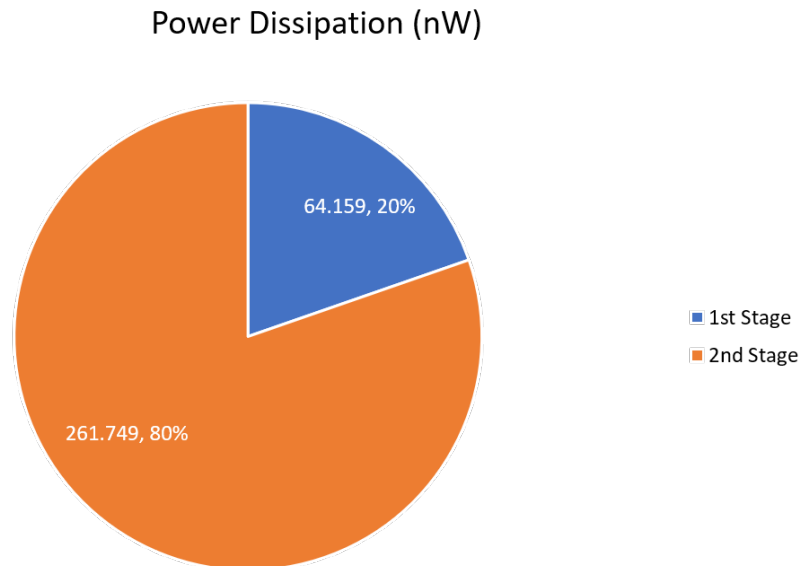
Figure 31: Power dissipation distribution in the original two-stage op-amp.

The simulated results of the g_m/I_D-based two-stage op-amp designed in Section 5.3 are collated in Table 11.

---

[11]For a fair comparison, the Miller compensation capacitance area has been excluded since the Miller compensation capacitance area will dominate the entire chip area and obscure any differences in the total transistor active area.
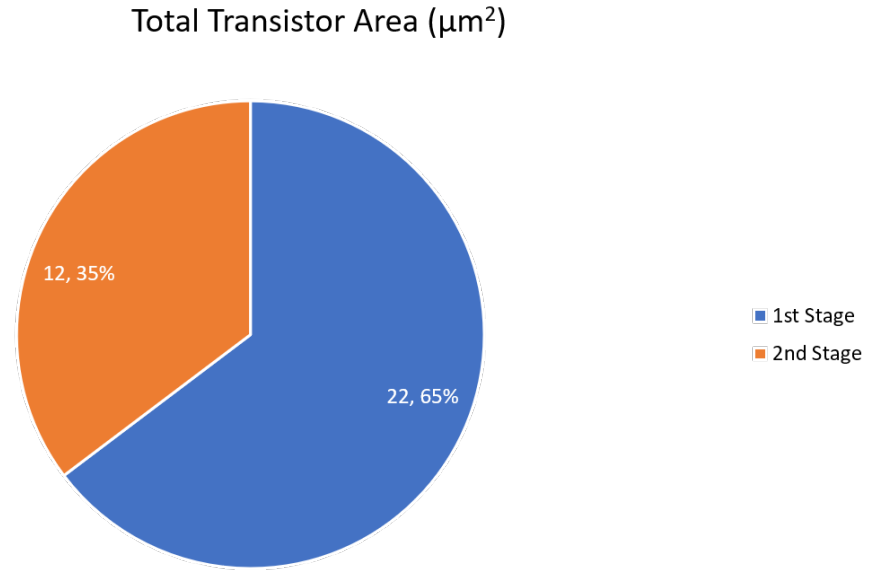
Figure 32: Total transistor area distribution in the original two-stage op-amp.

The distribution of power dissipation and active area (excluding the Miller compensation capacitance area) between the two stages of the g_m/I_D-based op-amp can be seen from Figure 33 and Figure 34.

Stage 1 Current in Both Branches (nA)	57.236
Stage 2 Current (nA)	226.584
Total Power Dissipation (nW)	283.820
Stage 1 Gain (dB)	49.794
Stage 2 Gain (dB)	47.252
Total Gain (dB)	97.046
Phase Margin	53.990°
Unity Gain Frequency (kHz)	79.476
Total Active Area Without Miller Capacitance (μm^2)	37
1.2 pF Miller Capacitance Area (μm^2)	568.516
Total Input Referred Noise (μV, 0.001 Hz - 10 kHz)	21.695

Table 12: g_m/I_D-based two-stage op-amp (TSMC 180 nm) results.

Power Dissipation (nW)

57.236, 20%

226.584, 80%

■ 1st Stage
■ 2nd Stage
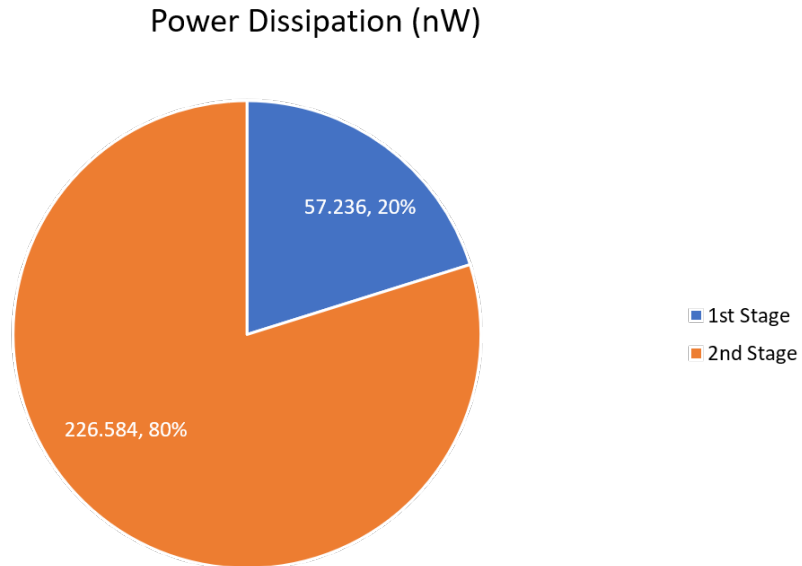
Figure 33: Power dissipation distribution in the g_m/I_D-based two-stage op-amp.

Total Transistor Area (μm^2)

13, 35%

24, 65%

■ 1st Stage
■ 2nd Stage
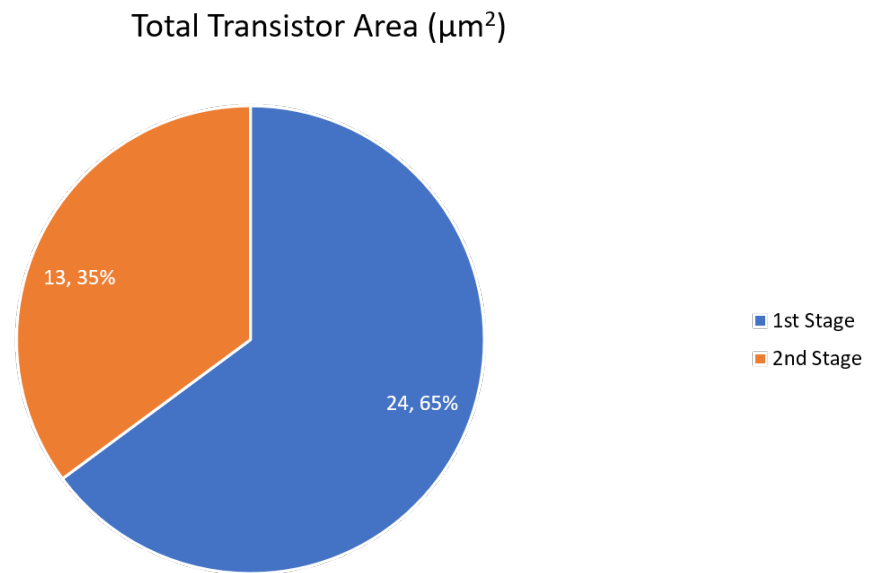
Figure 34: Total transistor area distribution in the g_m/I_D-based two-stage op-amp.

The original two-stage op-amp has a higher power dissipation but smaller equivalent input noise and area compared to the op-amp designed using the g_m/I_D methodology. The trade-off between power dissipation and equivalent input noise can be easily seen from the results. A higher power dissipation almost always leads to a better noise performance. Note that although the original two-stage op-amp has a higher power dissipation, its UGF is lower than that of the g_m/I_D-based op-amp. The reason for this is the larger Miller compensation capacitance used in the original op-amp.

Next, we compare the two op-amps by computing their FoM. The FoM of Trakoolwattana and Thanachayanont's design is denoted by one asterisk symbol, whereas the FoM of the g_m/I_D-based op-amp is denoted by two asterisk symbols.

$$FoM^* = \left(\left(\frac{325.908 \times 10^{-9}}{609.728 \times 10^{-9}} \right)^2 + \left(\frac{20.910 \times 10^{-6}}{42.605 \times 10^{-6}} \right)^2 + \left(\frac{34 \times 10^{-12}}{71 \times 10^{-12}} \right)^2 \right)^{\frac{1}{2}} \tag{73}$$

$$= 0.86942 \tag{74}$$

$$FoM^{**} = \left(\left(\frac{283.820 \times 10^{-9}}{609.728 \times 10^{-9}} \right)^2 + \left(\frac{21.695 \times 10^{-6}}{42.605 \times 10^{-6}} \right)^2 + \left(\frac{37 \times 10^{-12}}{71 \times 10^{-12}} \right)^2 \right)^{\frac{1}{2}} \tag{75}$$

$$= 0.86461 \tag{76}$$

Since $FoM^{**} < FoM^*$, the op-amp designed in Section 5.3 is a more optimal design. However, the two FoMs are very close to each other. This indicates that both designs are in the optimal neighbourhood of designs. Nevertheless, this does not distract from the argument that the g_m/I_D methodology provides a systematic framework that can guide the designer to an optimal region of designs and save him (her) a lot of time that would otherwise be spent on tweaking in Cadence Spectre.

6.3 Folded Cascode Op-amp Results (TSMC 180 nm)

The simulated results of the original folded cascode op-amp designed by Trakoolwattana and Thanachayanont are collated in Table 13.

Stage 1 Current in Both Branches (nA)	283.595
Stage 2 Current (nA)	126.046
Total Power Dissipation (nW)	409.641
Stage 1 Gain (dB)	78.0276
Stage 2 Gain (dB)	-1.638
Total Gain (dB)	76.3896
Phase Margin	49.137°
Unity Gain Frequency (kHz)	62.2024
Total Active Area Without Miller Capacitance (μm^2)	307
Miller Capacitance Area (μm^2)	1217.712
Total Input Referred Noise (μV, 0.001 Hz - 10 kHz)	31.267

Table 13: Original folded cascode op-amp results.

The distribution of power dissipation and active area (excluding the Miller compensation capacitance) between the two stages can be seen from Figure 35 and Figure 36.
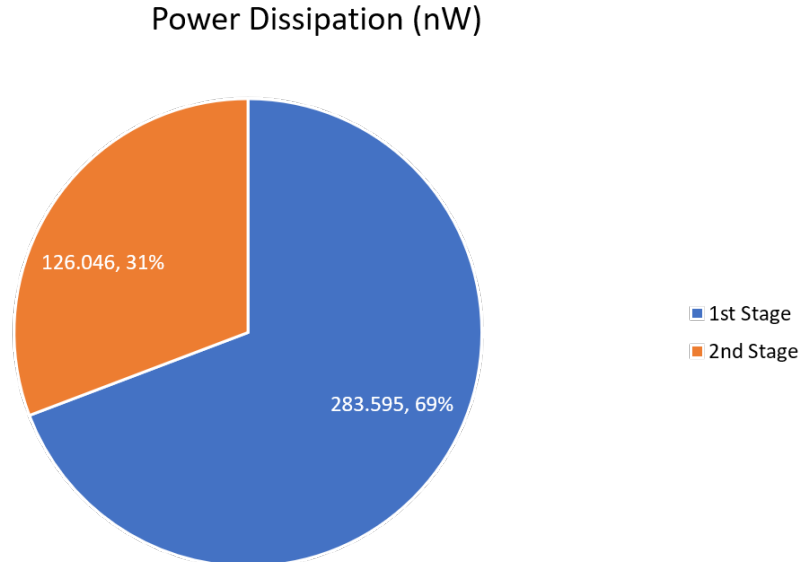
Figure 35: Power dissipation distribution in the original folded cascode op-amp.

The simulated results of the g_m/I_D-based folded cascode op-amp are collated in Table 14.

The distribution of power dissipation and active area (excluding the Miller compensation capacitance) between the two stages of the g_m/I_D-based folded-cascode op-amp can be seen from Figure 37 and Figure 38.
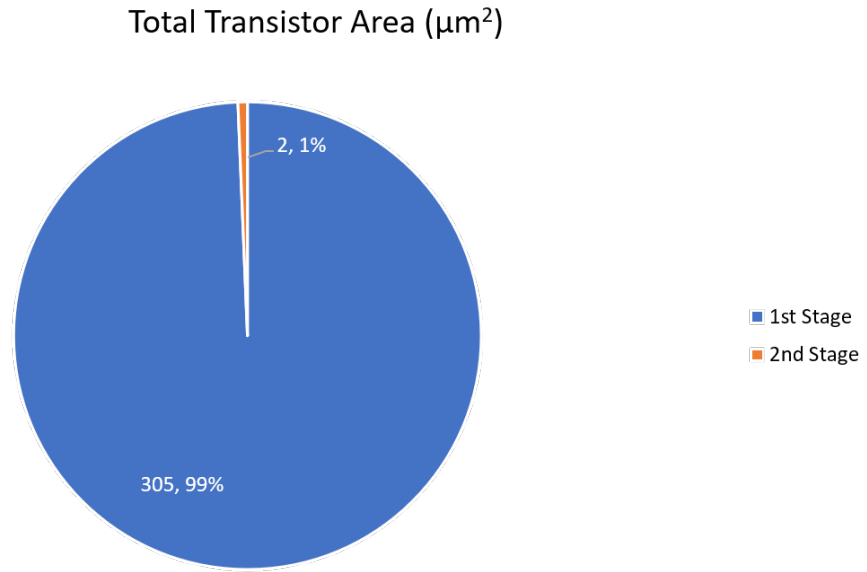
Figure 36: Total transistor area distribution in the original folded cascode op-amp.

Stage 1 Current in Both Branches (nA)	92.614
Stage 2 Current (nA)	69.060
Total Power Dissipation (nW)	161.674
Stage 1 Gain (dB)	70.277
Stage 2 Gain (dB)	44.890
Total Gain (dB)	115.167
Phase Margin	47.680°
Unity Gain Frequency (kHz)	30.327
Total Active Area Without Miller Capacitance (μm^2)	275
Miller Capacitance Area (μm^2)	572.292
Total Input Referred Noise (μV, 0.001 Hz - 10 kHz)	59.555

Table 14: g_m/I_D-based folded cascode op-amp results.

Power Dissipation (nW)
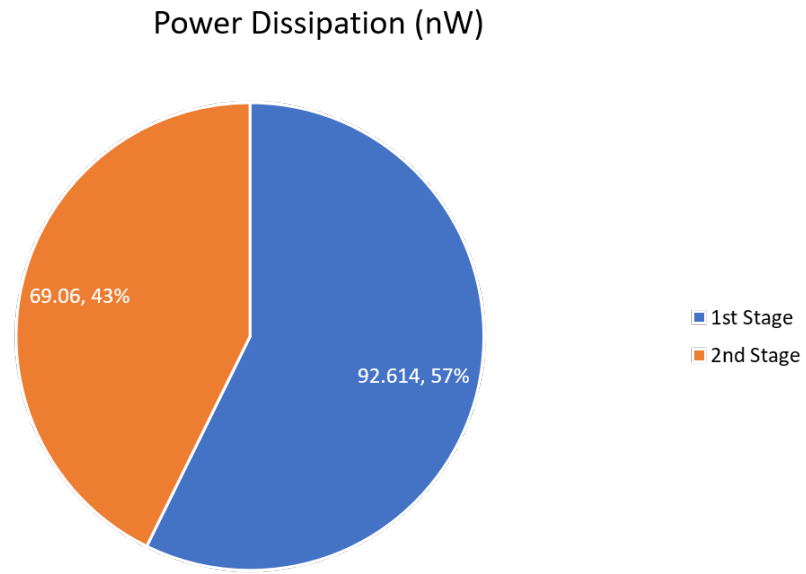
Figure 37: Power dissipation distribution in the g_m/I_D-based folded cascode op-amp.

Total Transistor Area (μm^2)
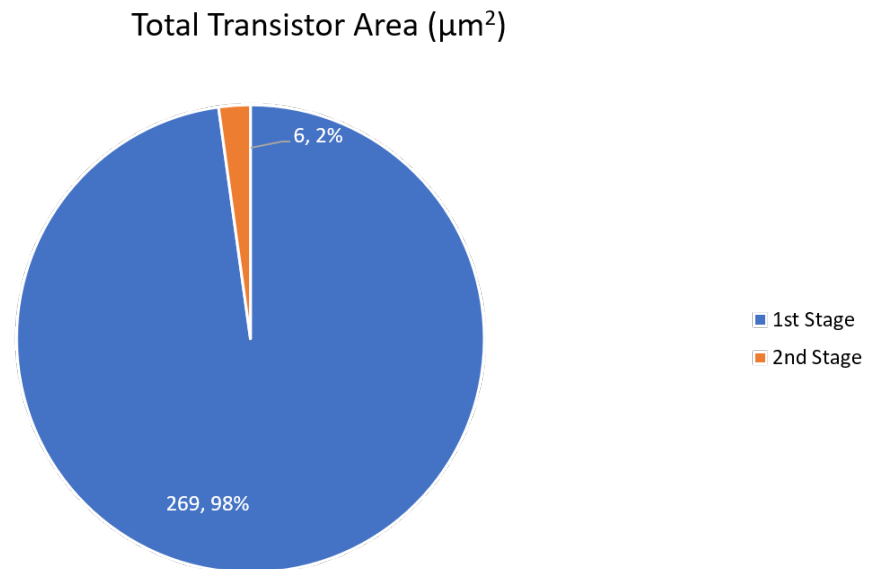
Figure 38: Total transistor area distribution in the g_m/I_D-based folded cascode op-amp.

The op-amp designed using the g_m/I_D methodology has a much lower power dissipation and area than the original op-amp. This is made possible by relaxing the dominant pole condition.

Next, we compare the folded cascode op-amp designed by Trakoolwattana and Thanachayanont against the op-amp designed in Section 5.5 by computing their FoM. The FoM of Trakoolwattana and Thanachayanont's design is denoted by one + symbol, whereas the FoM of the g_m/I_D op-amp is denoted by two ++ symbols.

$$FoM^+ = \left(\left(\frac{409.641 \times 10^{-9}}{571.315 \times 10^{-9}} \right)^2 + \left(\frac{31.267 \times 10^{-6}}{90.822 \times 10^{-6}} \right)^2 + \left(\frac{307 \times 10^{-12}}{582 \times 10^{-12}} \right)^2 \right)^{\frac{1}{2}} \tag{77}$$

$$= 0.95440 \tag{78}$$

$$FoM^{++} = \left(\left(\frac{161.674 \times 10^{-9}}{571.315 \times 10^{-9}} \right)^2 + \left(\frac{59.555 \times 10^{-6}}{90.822 \times 10^{-6}} \right)^2 + \left(\frac{275 \times 10^{-12}}{582 \times 10^{-12}} \right)^2 \right)^{\frac{1}{2}} \tag{79}$$

$$= 0.85635 \tag{80}$$

It is clear that the op-amp design in Section 5.5 is a much better op-amp. This is largely due to the fact that the power dissipation has been drastically reduced from 409.641 nW to 161.674 nW, a 60.5% improvement.

6.4 Top Level Results (TSMC 180 nm)

The optimised op-amps in TSMC 180 nm were connected to a wide-swing cascode current mirror to form the entire potentiostat topology given in Figure 10.
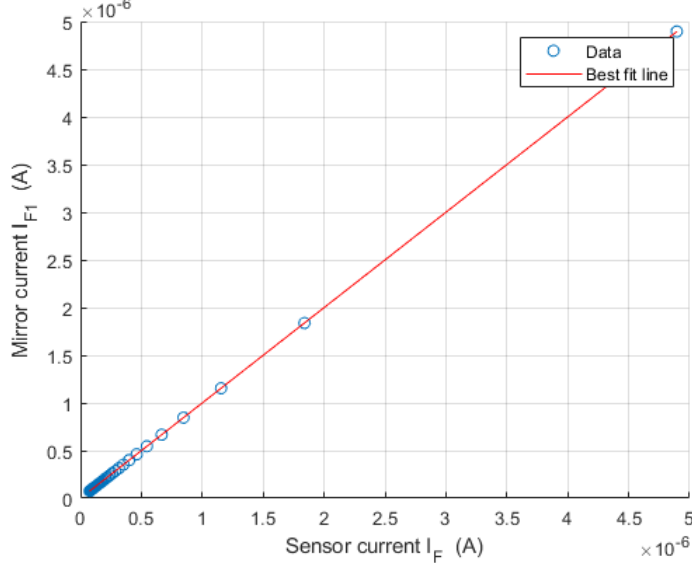
Figure 39: Linearity of the mirror (measured) current vs the sensor current.

In order to assess the accuracy and linearity of the potentiostat, R_{WE} (see Figure 6) was varied from 100 kΩ to 10 MΩ by performing a parametric sweep in Cadence Spectre to yield 50 data points. This was also equivalent to changing the sensor current from 4.895 μA to 60.615 nA. The corresponding mirror current data points were collated and plotted against the sensor data points as seen in Figure 39. A linear regression model was implemented in MATLAB. The results show that the coefficient of determination R^2 is 1.0000[12]. The potentiostat circuit based on optimised op-amps shows excellent linearity.

	This work	Trakoolwattana and Thanachayanont
V_{DD} (V)	1	1
Technology (nm)	180	180
R^2	1	0.99996
Maximum Power Dissipation $I_F = 4.9\mu$A (μW)	10.295	12.3

Table 15: Performance comparison between this work and the work by Trakoolwattana and Thanachayanont.

It can be seen from Table 15 that the potentiostat designed using the g_m/I_D methodology has a lower power dissipation than the original potentiostat.

---

[12]As a result of the limited machine precision, MATLAB gives the value of R^2 to be 1.0000. Obviously, a perfect coefficient of determination is impossible to achieve in practice. The take-home message is that the sensor current and the mirror current are extremely close to each other.

6.5 Op-amp Performance in TSMC 180 nm vs. TSMC 65 nm

The performance of the op-amps designed using the g_m/I_D methodology in TSMC 65 nm is summarised in Table 16 and Table 17.

Stage 1 Current in Both Branches (nA)	129.7
Stage 2 Current (nA)	143
Total Power Dissipation (nW)	272.7
Stage 1 Gain (dB)	28.086
Stage 2 Gain (dB)	22.516
Total Gain (dB)	50.602
Phase Margin	12.99°
Unity Gain Frequency (kHz)	103.451
Total Active Area Without Miller Capacitance (μm^2)	34
Miller Capacitance Area (μm^2)	569.270
Total Input Referred Noise (μV, 0.001 Hz - 10 kHz)	11.992

Table 16: g_m/I_D-based two-stage op-amp (TSMC 65 nm) results.

Stage 1 Current in Both Branches (nA)	80.92
Stage 2 Current (nA)	52.68
Total Power Dissipation (nW)	133.60
Stage 1 Gain (dB)	7.169
Stage 2 Gain (dB)	16.701
Total Gain (dB)	23.870
Phase Margin	61.339°
Unity Gain Frequency (kHz)	11.162
Total Active Area Without Miller Capacitance (μm^2)	331
Miller Capacitance Area (μm^2)	569.270
Total Input Referred Noise (μV, 0.001 Hz - 10 kHz)	100.053

Table 17: g_m/I_D-based folded cascode op-amp (TSMC 65 nm) results.

For a comparable power dissipation and area, the op-amps designed in TSMC 65 nm exhibit many shortcomings, and the most obvious of which is a reduced gain. Although the gain of an op-amp is not an optimisation objective in this FYP, a high gain is a necessary condition for all high performing analogue circuits. Therefore, even though the two-stage op-amp designed in TSMC 65 nm has a lower power dissipation and input referred noise, it cannot be argued that the op-amp in TSMC 65 nm is superior. There are many factors to consider when assessing an op-amp's performance.

In addition, the gate leakage current becomes a very significant problem that cannot be overlooked when designing in TSMC 65 nm. Typically, the analogue designer assumes that the gate of a MOSFET is an extremely high resistance that does not pass any current. However,

this assumption is no longer valid in shorter technology nodes and is especially pronounced in post-180 nm technology nodes. From an electrical current perspective, the two stages in the op-amp cannot be considered in isolation.

At this point, it must be explained that the relatively poor performance of the g_m/I_D designed op-amps is not the fault of the g_m/I_D methodology per se, but is due to the selected topology and CMOS technology. Fundamentally, the g_m/I_D methodology is a useful sizing tool that can help the designer to evaluate the trade-offs in a selected topology. The scope of the g_m/I_D methodology is bounded by the circuit topology and the chosen CMOS technology. The g_m/I_D methodology cannot help to solve performance problems that arise from an inappropriate circuit topology. To design in the 65 nm technology node, one would need to redesign the potentiostat topology completely. The single-ended output op-amps must be changed to fully differential op-amps to reap the benefits of differential operation. In addition, the transimpedance amplifier that generates the output voltage measurement should be changed to a switched-capacitor implementation.

6.6   Evaluation

The g_m/I_D methodology has been successfully extended to the design for low power dissipation. This can be seen in the performance of the optimised op-amps. The complete potentiostat which is formed by connecting the optimised op-amps showed better high level results than the original design by Trakoolwattana and Thanachayanont. This served to validate a central thesis of this FYP, in that the g_m/I_D methodology can lead the designer to an optimised designer. The simulated results showed great promise for the circuits designed in TSMC 180 nm. The trade-offs in this circuit were discussed extensively throughout this report and influenced the design and implementation processes greatly.

However, the performance of the op-amps in TSMC 65 nm were less than satisfactory. This is in part because analogue design does not scale with the technology nodes and the ever shrinking technology nodes present a mixed-blessing to analogue designers. Analogue designers have to work with devices that are increasingly difficult to model, making the design task more challenging. Note that, the g_m/I_D methodology works best on transistor-level designs. It cannot help to solve problems that are caused by an inappropriate circuit topology. Trakoolwattana and Thanachayanont topology is not suited for design in 65 nm technology which requires more advanced circuit techniques. Nevertheless, the immense practical value of the g_m/I_D methodology to the analogue designer, especially a novice designer has been explained and demonstrate throughout this report. The application of the g_m/I_D methodology to a potentiostat topology that is suitable for the 65 nm technology node is an exciting area to investigate, should time permits.

7 Conclusion

This project aims to investigate the fundamental trade-offs and a systematic framework to design potentiostats for the Body Dust Project. The g_m/I_D methodology was selected and it was demonstrated that the g_m/I_D methodology can lead to better designs.

Firstly, the operation of the potentiostat was discussed. A potentiostat simply needs to maintain a fixed potential difference between the working electrode and the reference electrode and measure the amount of current flowing through the counter electrode. There are many different potentiostat topologies, each with its advantages and disadvantages. Ahmadi and Jullien's work presents a pioneering potentiostat topology that is based on current mirrors. However, Ahmadi and Jullien's topology is dissipates too much power, rendering it unsuitable for the Body Dust Project. Trakoolwattana and Thanachayanont's topology is also current mirror-based but it has a very low power dissipation and was selected as the basis for this FYP.

Secondly, the g_m/I_D methodology was explained. The central tenet of the g_m/I_D methodology is to design circuits using g_m/I_D as the "tuning" knob. g_m/I_D is a fundamental transistor figure of merit that can be used to explore the inherent trade-offs in a circuit.

Next, Trakoolwattana and Thanachayanont's topology was analysed block by block. The frequency response of the op-amps were explained in detail. In addition, the impact of the Miller compensation capacitance was analysed rigorously, with emphasis on Allen's and Jespers' heuristics. The operation of the unorthodox, bulk-driven technique and its implications were also introduced.

The use of the SPICE-generated lookup table data as well as two auxiliary functions `look_up` and `look_upVGS` was introduced with the IGS example. Next, the implementation of the two-stage op-amp was discussed, with emphasis on the MATLAB script that offers user interaction and plots contour curves. The implementation of the folded cascode op-amp followed a different, more direct approach and was elaborated. In implementing both op-amps, the use of the g_m/I_D methodology as the guiding principle yielded satisfactory circuit designs that showed a close agreement between the simulated and intended results.

Finally, the implication of scaling the design of the potentiostat topology to the 65 nm technology node was discussed. In order to design in the 65 nm technology node, more advanced circuit techniques must be used.

7.1 Future Work

This FYP can be taken further by investigating new topologies that are better suited for TSMC 65 nm technology. In particular, the op-amps must be converted to fully differential op-amps. A switched-capacitor implementation of the transimpedance amplifier is also worth looking into.

Note that the equivalent circuit for the electrochemical cell (Figure 6) was adopted in this

FYP. This equivalent circuit is no longer valid for more advanced electrochemical cell. In [38], Beltrandi et al proposed a new VHDL-AMS model for the electrochemical cell that is more accurate. In future works, this new model for the electrochemical cell should be used instead.

Lastly, as mentioned before, the g_m/I_D methodology can only lead the designer to an optimal neighbourhood of designs. In future works, the g_m/I_D methodology could be complemented by optimisation techniques to lead the designer to a truly optimal design. In [39], the use of geometric programming techniques to design optimal op-amps has been successfully demonstrated. It is certainly worthwhile integrating geometric programming techniques into the g_m/I_D methodology.

References

[1] Sandro Carrara and Pantelis Georgiou. Body dust: Miniaturized highly-integrated low power sensing for remotely powered drinkable cmos bioelectronics. *arXiv preprint arXiv:1805.05840*, 2018.

[2] Mohammad Mahdi Ahmadi and Graham A Jullien. Current-mirror-based potentiostats for three-electrode amperometric electrochemical sensors. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 56(7):1339–1348, 2008.

[3] Rungtrakool Trakoolwattana and Apinunt Thanachayanont. 1-v low-power high-accuracy cmos potentiostat circuit for amperometric sensor. In *2014 11th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 1–4. IEEE, 2014.

[4] Paul GA Jespers and Boris Murmann. *Systematic Design of Analog CMOS Circuits*. Cambridge University Press, 2017.

[5] Guanyu Hao, Zhi Ping Xu, and Li Li. Manipulating extracellular tumour ph: an effective target for cancer therapy. *RSC advances*, 8(39):22182–22192, 2018.

[6] Allen J Bard, Larry R Faulkner, et al. Fundamentals and applications. *Electrochemical Methods*, 2(482):580–632, 2001.

[7] L Busoni, Marcello Carla, and Leonardo Lanzi. A comparison between potentiostatic circuits with grounded work or auxiliary electrode. *Review of scientific instruments*, 73(4):1921–1923, 2002.

[8] Ralf G Kakerow, Holger Kappert, Egbert Spiegel, and Yiannos Manoli. Low-power single-chip cmos potentiostat. In *Proceedings of the International Solid-State Sensors and Actuators Conference-TRANSDUCERS'95*, volume 1, pages 142–145. IEEE, 1995.

[9] R. Doelling. Potentiostats. *Bank Elektronik Application Note, 2nd ed. Clausthal-Zellerfeld, Germany: Bank Elektroni–Intelligent Controls GmbH*, mar 2000.

[10] Robert Greef. Instruments for use in electrode process research. *Journal of Physics E: Scientific Instruments*, 11(1):1, 1978.

[11] Milad Razzaghpour, Saul Rodriguez, Eduard Alarcon, and Ana Rusu. A highly-accurate low-power cmos potentiostat for implantable biosensors. In *2011 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 5–8. IEEE, 2011.

[12] E Wilkins and P Atanasov. Glucose monitoring: state of the art and future possibilities. *Medical engineering & physics*, 18(4):273–288, 1996.

[13] E. Vittoz and J. Fellrath. Cmos analog integrated circuits based on weak inversion operations. *IEEE Journal of Solid-State Circuits*, 12(3):224–231, 1977.

[14] Christian Enz, Maria-Anna Chalkiadaki, and Anurag Mangla. Low-power analog/rf circuit design based on the inversion coefficient. In *ESSCIRC Conference 2015-41st European Solid-State Circuits Conference (ESSCIRC)*, pages 202–208. IEEE, 2015.

[15] Paul Jespers. *The gm/ID Methodology, a sizing tool for low-voltage analog CMOS Circuits: The semi-empirical and compact model approaches.* Springer Science & Business Media, 2009.

[16] David M Binkley, CE Hopper, Steve D Tucker, Brian C Moss, James M Rochelle, and DP Foty. A cad methodology for optimizing transistor current and sizing in analog cmos design. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 22(2):225–237, 2003.

[17] Bo Lojek. *History of semiconductor engineering.* Springer, 2007.

[18] Behzad Razavi. *Design of analog CMOS integrated circuits.* Tata McGraw-Hill Education, 2002.

[19] Paul R Gray, Paul J Hurst, Stephen H Lewis, and Robert G Meyer. *Analysis and design of analog integrated circuits.* John Wiley & Sons, 2009.

[20] Bart Van Zeghbroeck. Principles of semiconductor devices. *Colarado University*, 34, 2004.

[21] Gennady Gildenblat, Xin Li, Weimin Wu, Hailing Wang, Amit Jha, Ronald Van Langevelde, Geert DJ Smit, Andries J Scholten, and Dirk BM Klaassen. Psp: An advanced surface-potential-based mosfet model for circuit simulation. *IEEE Transactions on Electron Devices*, 53(9):1979–1993, 2006.

[22] Yogesh Singh Chauhan, Sriramkumar Venugopalan, Maria-Anna Chalkiadaki, Muhammed Ahosan Ul Karim, Harshit Agarwal, Sourabh Khandelwal, Navid Paydavosi, Juan Pablo Duarte, Christian C Enz, Ali M Niknejad, et al. Bsim6: Analog and rf compact model for bulk mosfet. *IEEE Transactions on Electron Devices*, 61(2):234–244, 2013.

[23] Flandre Silveira, Denis Flandre, and Paul GA Jespers. A g/sub m//i/sub 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*, 31(9):1314–1319, 1996.

[24] Fernando Silveira and Denis Flandre. Operational amplifier power optimization for a given total (slewing plus linear) settling time. In *Proceedings. 15th Symposium on Integrated Circuits and Systems Design*, pages 247–253. IEEE, 2002.

[25] Takayuki Konishi, Kenji Inazu, Jun Gyu Lee, Masanori Natsui, Shoichi Masui, and Boris Murmann. Design optimization of high-speed and low-power operational transconductance amplifier using g m/i d lookup table methodology. *IEICE transactions on electronics*, 94(3):334–345, 2011.

[26] Denis Flandre, Alberto Viviani, J-P Eggermont, Bernard Gentinne, and PGA Jespers. Improved synthesis of gain-boosted regulated-cascode cmos stages using symbolic analysis and gm/id methodology. *IEEE Journal of Solid-State Circuits*, 32(7):1006–1012, 1997.

[27] Christian C Enz and Eric A Vittoz. *Charge-based MOS transistor modeling: the EKV model for low-power and RF IC design.* John Wiley & Sons, 2006.

[28] Thomas H. Lee. *The Design of CMOS Radio-Frequency Integrated Circuits.* Cambridge University Press, 2 edition, 2003. doi: 10.1017/CBO9780511817281.

[29] Asad A Abidi. High-frequency noise measurements on fet's with small dimensions. *IEEE transactions on electron devices*, 33(11):1801–1805, 1986.

[30] C. C. Enz and G. C. Temes. Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization. *Proceedings of the IEEE*, 84(11):1584–1614, 1996.

[31] R Sharpeshkar. Ultra low power bioelectronics: Fundamentals, biomedical applications, and bio-inspired system, 2010.

[32] Paul R Gray and Robert G Meyer. Mos operational amplifier design-a tutorial overview. *Ieee journal of solid-state circuits*, 17(6):969–982, 1982.

[33] Daniel Senderowicz, David A Hodges, and Paul R Gray. High-performance nmos operational amplifier. *IEEE Journal of Solid-State Circuits*, 13(6):760–766, 1978.

[34] Yannis P Tsividis and Paul R Gray. An integrated nmos operational amplifier with internal compensation. *IEEE Journal of Solid-State Circuits*, 11(6):748–753, 1976.

[35] Phillip E Allen and Douglas R Holberg. *CMOS analog circuit design.* Elsevier, 2011.

[36] A Guzinski. Body-driven differential amplifier for application in continuous-time active-c filer. In *Proc. European Conf. Circuit Theory and Design (ECCTD'87)*, pages 315–320, 1987.

[37] Shouli Yan and Edgar Sanchez-Sinencio. Low voltage analog circuit design techniques: A tutorial. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 83(2):179–196, 2000.

[38] M. Beltrandi, A. Vachoux, S. Carrara, Y. Leblebici, and G. De Micheli. Vhdl-ams model of an electrochemical cell to design vlsi bio-chips. In *2011 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 321–324, 2011.

[39] M. d. Hershenson, S. P. Boyd, and T. H. Lee. Optimal design of a cmos op-amp via geometric programming. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 20(1):1–21, 2001.

Appendices

A MATLAB Code

The MATLAB code used in this FYP can be found in the following *private* GitHub repository.

  https://github.com/giraffe-zhang/FYP_gm-ID_Code

  Please contact yz12316@ic.ac.uk for access to view the code in the GitHub repository.

B Conference Paper Draft

See below.

Design of Low-Power Highly-Accurate CMOS Potentiostat Using the g_m/I_D Methodology

Yaohua Zhang, Daryl Mah, Pantelis Georgiou

Imperial College London, UK

Abstract— A potentiostat is a key integrated circuit for the proposed Body Dust Particles for future healthcare applications. This paper first evaluates the existing CMOS potentiostat topologies and then adopts the most suitable topology to be designed and optimised with the g_m/I_D methodology. The g_m/I_D methodology has been extended to low-power circuit design. Simulated results show that the power dissipation can be drastically reduced from 409.641 nW to 161.674 nW, indicating a 60.5% improvement for the folded cascode op-amp in TSMC 180 nm CMOS technology. A new figure of merit (FoM) based on the weighted Euclidean distance in a 3-D space is proposed to quantify the performance in terms of low power dissipation, small area and high signal-to-noise ratio. The smaller the FoM is, the better the performance is. The FoM of 0.85635 was achieved in our design as compared with that of 0.95440 in the referenced design from simulations.

Index Terms— Body dust project, potentiostat, g_m/I_D methodology, CMOS

I. INTRODUCTION

The Body Dust Project was proposed by Dr S. Carrara and Dr P. Georgiou. In their Body Dust paper, a vision for wirelessly powered, drinkable CMOS integrated circuits (ICs) for the next generation of disease/tumour detection technologies was presented [1]. The idea is to develop tiny CMOS ICs, coated in a bio-compatible packaging with specific bio-molecules that are attracted to the source of a disease in organs and tissues like cancerous tumours. These CMOS ICs (dubbed Body Dust Particles) should be wirelessly powered and should also send back relevant data about the source of disease to medical professionals. For instance, transmitting the pH and glucose levels could indicate the nature of the tumour as well as any growth in size [2]. This method of detecting the source of a disease has the advantages of being minimally invasive compared to measures like biopsy, relatively small side-effects compared to X-ray and other radiation-based methods, and could be potentially cheap and affordable to the masses.

This paper describes the design of the potentiostat in TSMC's 180-nm CMOS technology with power dissipation, area, and signal-to-noise ratio (SNR) as the optimisation objectives. In Section II, the existing potentiostat topologies are briefly reviewed, and the most suitable topology is selected as a basis. In Section III, the g_m/I_D methodology is introduced. In Section IV, simulated results are compared to validate our design. Finally, the conclusion is drawn in Section V.

II. BRIEF REVIEW OF THE POTENTIOSTAT TOPOLOGIES

In essence, a potentiostat is an electronic circuitry that fulfils two basic functions of maintaining a desired potential difference between the working electrode (WE) and reference electrode (RE) and measuring the amount of current directed through the counter electrode (CE) for a three-electrode amperometric electrochemical sensor [3].

Ahmadi and Jullien provided a detailed review of the then existing potentiostat topologies. Considering the strengths and limitations of the potentiostat topologies reviewed, they proposed a new topology based on current mirrors in [3]. The novelty of this topology lies in using a current mirror to create a copy of the sensor current and measure the mirrored current instead of the original sensor current. It presents several advantages. Firstly, this is a grounded WE configuration, which makes the WE less vulnerable to noise and EMI pickup [3]. Secondly, this topology requires fewer components which is helpful to alleviate noise and component matching concerns. Despite its advantages, this topology consumes too much area and dissipates too much power, rendering it to be unacceptable in the Body Dust Project context.

Trakoolwattana and Thanachayanont proposed a alternative potentiostat topology, as shown in Fig. 1, which boasts low power dissipation and high accuracy [4]. In this potentiostat topology, the op-amp A_1 is bulk-driven that allows for a low supply voltage operation (V_{DD} is 1 V) and a low input common-mode voltage level (0.4 V assuming 0.6 V cell potential for glucose). This topology also achieves high accuracy, i.e. good matching between the sensor current (I_F) and the mirrored sensor current (I_{F1}) using a wide-swing cascode current mirror and the op-amp A_2. A_2 serves two purposes. Firstly, it acts as a transimpedance amplifier to produce a voltage output. Secondly, it helps to keep the drain voltages of M_1 and M_2 identical via negative feedback, ensuring good accuracy between the sensor current and its copy. This is an elegant topology that is effective, robust and compact. With its low power and high accuracy, Trakoolwattana and Thanachayanont's topology is a strong candidate for the Body Dust Project and is therefore chosen for further optimisation using the g_m/I_D methodology.
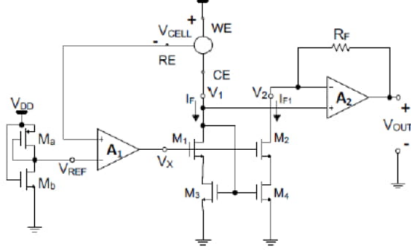
Fig. 1. Trakoolwattana and Thanachayanont's low power, high accuracy potentiostat topology [4, Figure 2].

III. DESCRIPTION OF THE DESIGN METHODOLOGY

The g_m/I_D methodology was first proposed by Silveira et al in 1996 and was published in the IEEE Journal of Solid-State Circuits (JSSC) [5]. The g_m/I_D methodology was developed in part to address the growing discrepancy between hand-analysis and simulation results in analogue CMOS IC design, as well as to provide a novel, systematic way of designing circuits with a strong grasp of the inherent trade-offs. It was crucial to bridge this gap for the overall progress in IC design. Since the publication of the seminal paper by Silveira et al in 1996, the g_m/I_D methodology has been further developed and applied in a wide range of analogue circuit design contexts [6].

g_m/I_D is a fundamental figure of merit that influences the performance of analogue circuits greatly. g_m/I_D bridges g_m, a small-signal quantity with I_D, a large-signal quantity. Qualitatively, g_m/I_D is a measure of how much gm (gain) we can get in return for each unit of I_D (bias current) that we invest. I_D can also be viewed as a proxy for the power dissipation. Thus, g_m/I_D encapsulates the most important small-signal and large-signal parameters of an analogue circuit.

$$\frac{g_m}{I_D} = \frac{1}{I_D}\frac{\partial I_D}{\partial V_G} = \frac{\partial \ln I_D}{\partial V_G} = \frac{\partial \ln[\frac{I_D}{W/L}]}{\partial V_G}$$

(1)

In the weak inversion region, g_m/I_D is given by

$$\frac{g_m}{I_D} = \frac{1}{I_D}\frac{\partial I_D}{\partial V_G} = \frac{1}{I_D}\frac{I_D}{nU_T} = \frac{1}{nU_T}$$

(2)

It is an established fact that g_m/I_D reaches a maximum in the weak inversion region, where the current-voltage relationship is exponential. The maximum possible g_m/I_D is given by $1/U_T$= 38.46 S/A, assuming $U_T \approx 26$ mV. The maximum g_m/I_D for bulk transistors is generally U_T between 20 to 30 S/A [6]. As the transistor is biased toward the strong inversion region, g_m/I_D decreases in value. g_m/I_D greater than 20 S/A corresponds to weak inversion; g_m/I_D between 20 to 10 S/A corresponds to moderate inversion; gm/I_D between 2 to 10 S/A corresponds to strong inversion

[6]. It should also be noted that these ranges of g_m/I_D remain relatively constant across transistor technologies [6]. Therefore, g_m/I_D can be used as a good proxy for the inversion level. g_m/I_D is also closely related to the drain saturation voltage V_{Dsat} by

$$\frac{2}{g_m/I_D} = V_{Dsat}$$

(3)

In addition, g_m/I_D is (to the first-order) independent of the width as given in (1), which makes it useful for device sizing. g_m/I_D can be employed to find the transistor widths by following this procedure [6]: 1) Derive g_m from the design specifications, 2) select the transistor lengths to satisfy g_m, speed, area, and matching requirements, 3) decide on g_m/I_D, taking the relevant trade-offs into consideration, 4) Determine I_D/W from g_m/I_D, 5) Derive $I_D = g_m/g_m/I_D$, 6) derive $W = I_D/I_D/W$.

The g_m/I_D-centric sizing procedure applies to most high performance circuits, however, it must be slightly modified when designing low-power circuits biased in the weak inversion region. This is because g_m/I_D is approximately constant in the weak inversion region and as a result, many different designs will be mapped to nearly the same g_m/I_D. In other words, a small error in g_m/I_D could lead to a design that is completely off-target. In this case, the g_m/I_D-centric sizing procedure will have to be modified into a J_D-centric sizing procedure, in which the current density (J_D) is used as the design variable. The J_D-centric sizing procedure is a variant of the g_m/I_D-centric sizing procedure and interested readers can refer to [7] for more details.

IV. RESULTS AND DISCUSSION

As shown in Fig. 1, the potentiostat basically consists of a current mirror, op-amps A_1 and A_2. Fig. 2 shows a bulk-driven folded-cascode two-stage op-amp for A_1. Fig. 3 shows a conventional two-stage op-amp for A_2. The task at hand is to optimise A_1 and A_2. A new FoM based on the weighted Euclidean distance in a 3-D space is proposed and given in (4) to quantify the performance in terms of low power dissipation, small area and low noise power.

$$FoM_i = \sqrt{\left(\frac{P_i}{\sum_i^n P_i} \times 100\right)^2 + \left(\frac{N_i}{\sum_i^n N_i} \times 100\right)^2 + \left(\frac{A_i}{\sum_i^n A_i} \times 100\right)^2}$$

(4)

where P, N, and A represent the power dissipation, equivalent input noise and area respectively; n is the number of op-amps that are being compared. It must be emphasised that this FoM should be used to compare op-amps of the same topology and comparable performance.
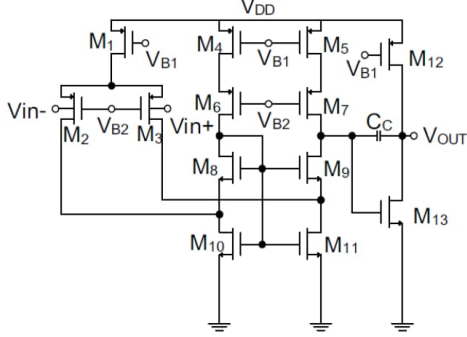
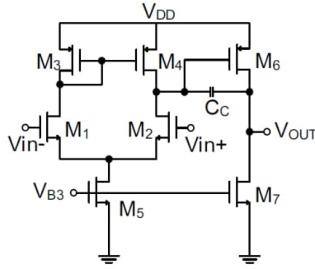Fig. 2. Op-amp schematic for A_1 [4, Figure 4a].

Fig. 3. Op-amp schematic for A_2 [4, Figure 4b].

A. Op-amp A_1

Table 1 lists the simulated results of A_1 using the design data by Trakoolwattana and Thanachayanont in [4].

TABLE 1

Stage 1 Current in Both Branches (nA)	283.595
Stage 2 Current (nA)	126.046
Total Power Dissipation (nW)	409.641
Stage 1 Gain (dB)	78.0276
Stage 2 Gain (dB)	-1.638
Total Gain (dB)	76.3896
Phase Margin	49.137°
Unity Gain Frequency (kHz)	62.2024
Total Active Area Without Miller Capacitance (μm^2)	307
Miller Capacitance Area (μm^2)	1217.712
Total Input Referred Noise (μV, 0.001 Hz - 10 kHz)	31.267

The FoM for the A_1 design can be calculated as 0.95440 with the data in Table 1.

$$FoM^+ = \left(\left(\frac{409.641 \times 10^{-9}}{571.315 \times 10^{-9}} \right)^2 + \left(\frac{31.267 \times 10^{-6}}{90.822 \times 10^{-6}} \right)^2 + \left(\frac{307 \times 10^{-12}}{582 \times 10^{-12}} \right)^2 \right)^{\frac{1}{2}}$$
$$= 0.95440$$

(5)

Table 2 gives the simulated results of A_1 using the improved g_m/I_D methodology in this work.

TABLE 2

Stage 1 Current in Both Branches (nA)	92.614
Stage 2 Current (nA)	69.060
Total Power Dissipation (nW)	161.674
Stage 1 Gain (dB)	70.277
Stage 2 Gain (dB)	44.890
Total Gain (dB)	115.167
Phase Margin	47.680°
Unity Gain Frequency (kHz)	30.327
Total Active Area Without Miller Capacitance (μm^2)	275
Miller Capacitance Area (μm^2)	572.292
Total Input Referred Noise (μV, 0.001 Hz - 10 kHz)	59.555

Similarly, the FoM for our A_1 design can be calculated with the data in Table 2 as 0.85635.

$$FoM^{++} = \left(\left(\frac{161.674 \times 10^{-9}}{571.315 \times 10^{-9}} \right)^2 + \left(\frac{59.555 \times 10^{-6}}{90.822 \times 10^{-6}} \right)^2 + \left(\frac{275 \times 10^{-12}}{582 \times 10^{-12}} \right)^2 \right)^{\frac{1}{2}}$$
$$= 0.85635$$

(6)

It is clear that our op-amp design for A_1 using the g_m/I_D methodology is a much better op-amp. This is largely due to the fact that the power dissipation has been drastically reduced from 409.641 nW to 161.674 nW, showing a 60.5% improvement.

B. Op-amp A_2

Running a MATLAB sizing procedure yields the contour plot shown in Fig. 4. The contour plots can help the designer to visualise the pros and cons of his (her) design choices. From Fig. 4, it can be seen that selecting a point with smaller J_{D2} and J_{D6} values, which means placing M_2 and M_6 in the weak inversion region leads to lower power dissipation. The gain also improves; however, the total transistor area increases. The plots for an optimised design are shown in Fig. 5. On closer inspection, it can be seen that the power dissipation has reduced. The values of the DC current (red) curves have reduced.

Table 3 lists the simulated results of A_2 using the design data by Trakoolwattana and Thanachayanont in [4]. The calculated FoM for the A_2 design is 0.86942 with the data in Table 3. Table 4 gives the simulated results of A_2 using the improved g_m/I_D methodology in this work. The calculated FoM for our A_2 design is 0.86461 with the data in Table 4. Our design is a more optimal design. However, the two FoMs are very close to each other. This indicates that both designs are in the optimal neighbourhood of designs. Nevertheless, this does not distract from the argument that the g_m/I_D methodology provides a systematic framework that can guide the designer to an optimal region of designs and save him (her) a lot of time that would otherwise be

spent on tweaking in Cadence Spectre.
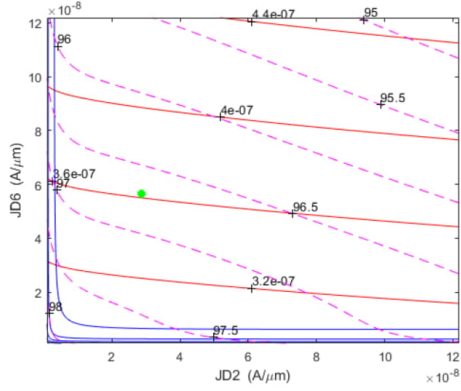
Fig. 4. Total DC current (nW, red), total active area (μm², blue), and total gain (dB, dashed magenta) contour curves. The selected point is marked by the green dot.
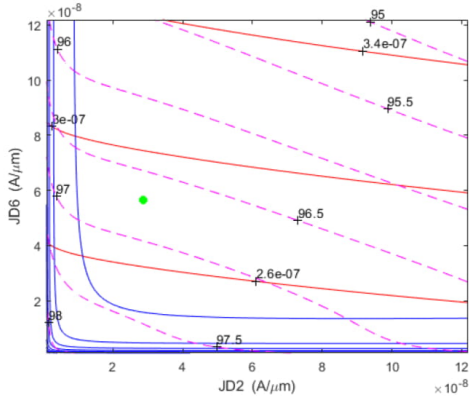
Fig. 5. Optimised design contour curves. Total DC current (nW, red), total active area (μm², blue), and total gain (dB, dashed magenta) contour curves. The selected point is marked by the green dot.

TABLE 3

Stage 1 Current in Both Branches (nA)	64.159
Stage 2 Current (nA)	261.749
Total Power Dissipation (nW)	325.908
Stage 1 Gain (dB)	49.847
Stage 2 Gain (dB)	46.503
Total Gain (dB)	96.350
Phase Margin	64.007°
Unity Gain Frequency (kHz)	52.409
Total Active Area Without Miller Capacitance (μm^2)	34
2 pF Miller Capacitance Area (μm^2)	968.256
Total Input Referred Noise (μV, 0.001 Hz - 10 kHz)	20.910

TABLE 4

Stage 1 Current in Both Branches (nA)	57.236
Stage 2 Current (nA)	226.584
Total Power Dissipation (nW)	283.820
Stage 1 Gain (dB)	49.794
Stage 2 Gain (dB)	47.252
Total Gain (dB)	97.046
Phase Margin	53.990°
Unity Gain Frequency (kHz)	79.476
Total Active Area Without Miller Capacitance (μm^2)	37
1.2 pF Miller Capacitance Area (μm^2)	568.516
Total Input Referred Noise (μV, 0.001 Hz - 10 kHz)	21.695

C. The Optimised Potentiostat

A linear regression analysis is made to the simulated data. Fig. 6 shows the simulated I_F and I_{Fl} in the range of 60.615 nA to 4.895 μA. As shown, a linear line is obtained, which demonstrates that the circuit exhibits excellent linearity with $R^2 = 1.0000$.
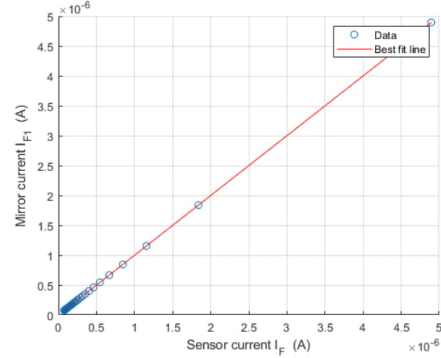
Fig. 6. Simulated linearity of the entire potentiostat optimised with g_m/I_D methodology.

V. CONCLUSION

A potentiostat based on the novel topology by Trakoolwattana and Thanachayanont was optimised in TSMC 180 nm CMOS technology using the g_m/I_D methodology for the proposed Body Dust particles for future healthcare applications. The existing g_m/I_D methodology was also extended to low-power circuit design. A new FoM based on the weighted Euclidean distance in a 3-D space was proposed to quantify the performance in terms of low power dissipation, small area and high signal-to-noise ratio. The smaller the FoM is, the better the performance is. The FoM of 0.85635 was achieved in our design as compared with that of 0.95440 in the referenced design from simulations.

REFERENCES

[1] Sandro Carrara and Pantelis Georgiou. Body dust: Miniaturized highly-integrated low power sensing for remotely powered drinkable cmos bioelectronics. arXiv preprint arXiv:1805.05840, 2018.

[2] Guanyu Hao, Zhi Ping Xu, and Li Li. Manipulating extracellular tumour pH: an effective target for cancer therapy. RSC advances, 8(39):22182–22192, 2018

[3] Mohammad Mahdi Ahmadi and Graham A Jullien., "Current-mirror-based potentiostats for three-electrode amperometric electrochemical sensors," *IEEE Transactions on Circuits and Systems I*: Regular Papers, vol. 56, no. 7, pp. 1339–1348, 2008.

[4] Rungtrakool Trakoolwattana and Apinunt Thanachayanont. "1-V low-power high-accuracy cmos potentiostat circuit for amperometric sensor," In *Proc. 2014 11th IEEE International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Informa-tion Technology (ECTI-CON)*, pp. 1–4, 2014.

[5] Flandre Silveira, Denis Flandre, and Paul GA 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 micro power OTA," *IEEE journal of solid-state circuits*, vol. 31, no. 9, pp. 1314–1319, 1996.

[6] Paul GA Jespers and Boris Murmann. Systematic Design of Analog CMOS Circuits. Cambridge University Press, 2017.

[7] Yaohua Zhang, Designing CMOS Integrated Circuits for Body Dust Project, Final Year Project Report, Imperial College London, 2020.

C g_m/I_D-based Two-Stage Op-amp Cadence Screenshots (TSMC 180 nm)

See below.
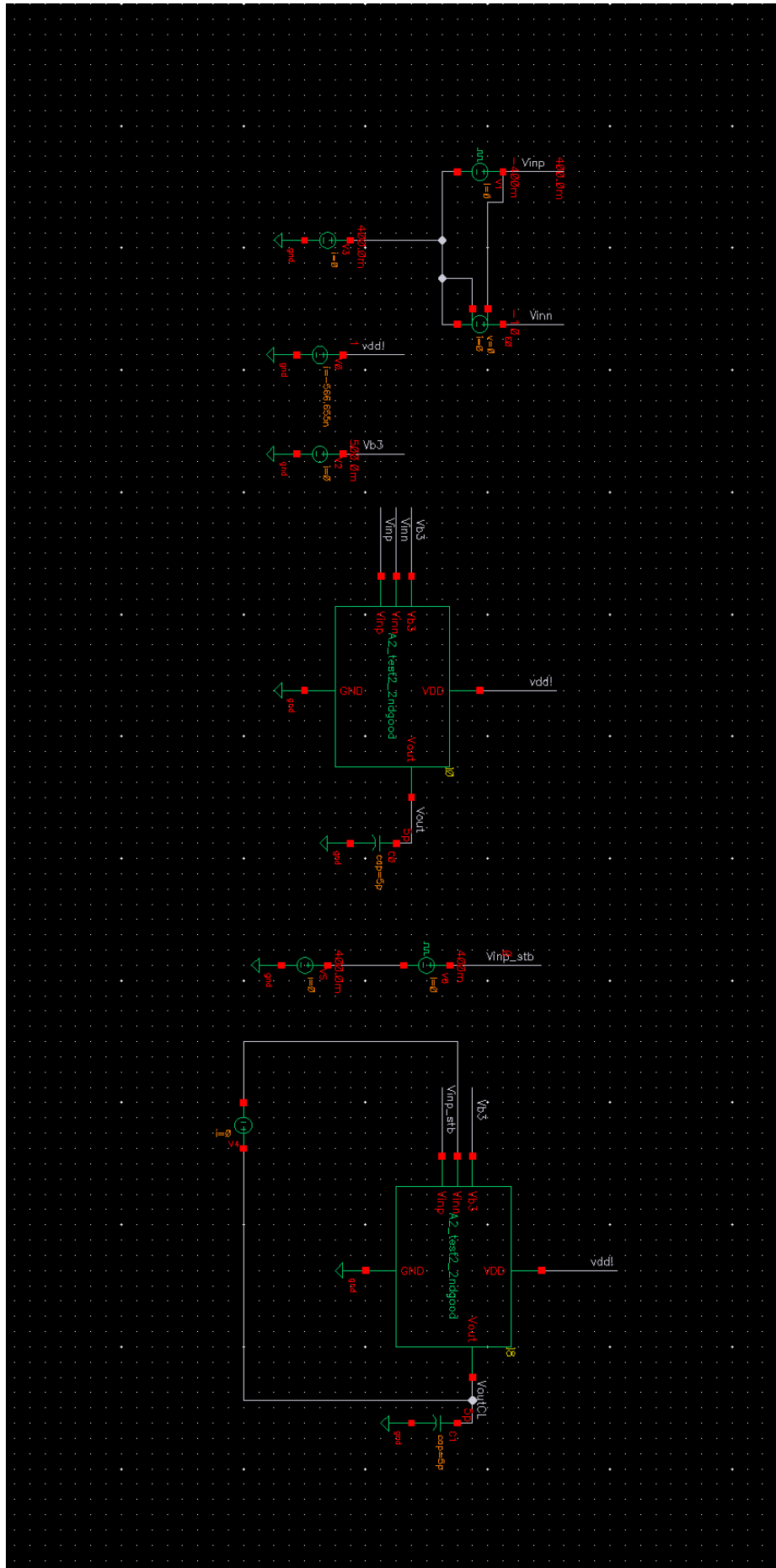
Figure 40: g_m/I_D-based two-stage op-amp test bench schematic; TSMC 180 nm.
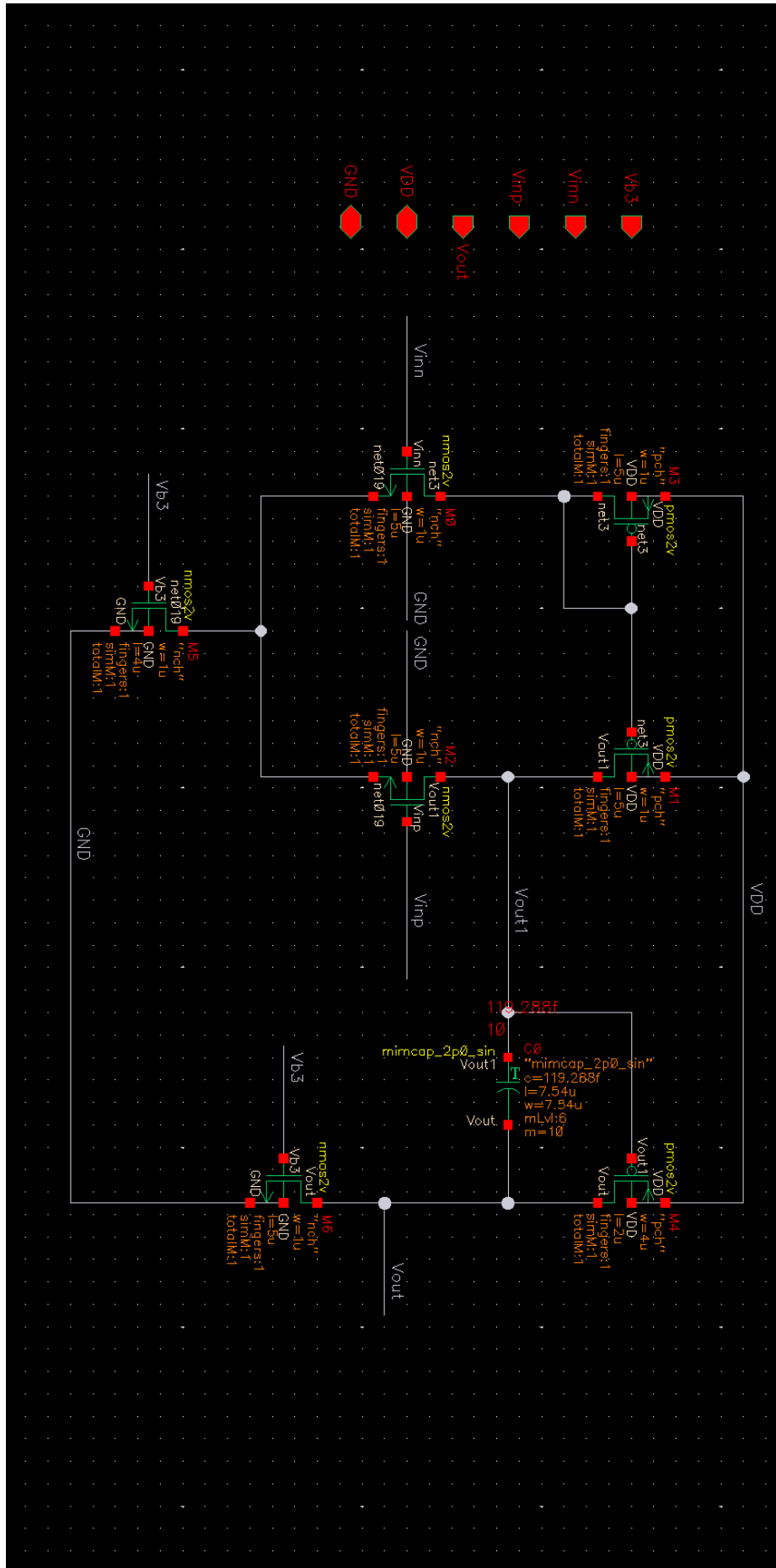
Figure 41: g_m/I_D-based two-stage op-amp transistor dimensions; TSMC 180 nm.
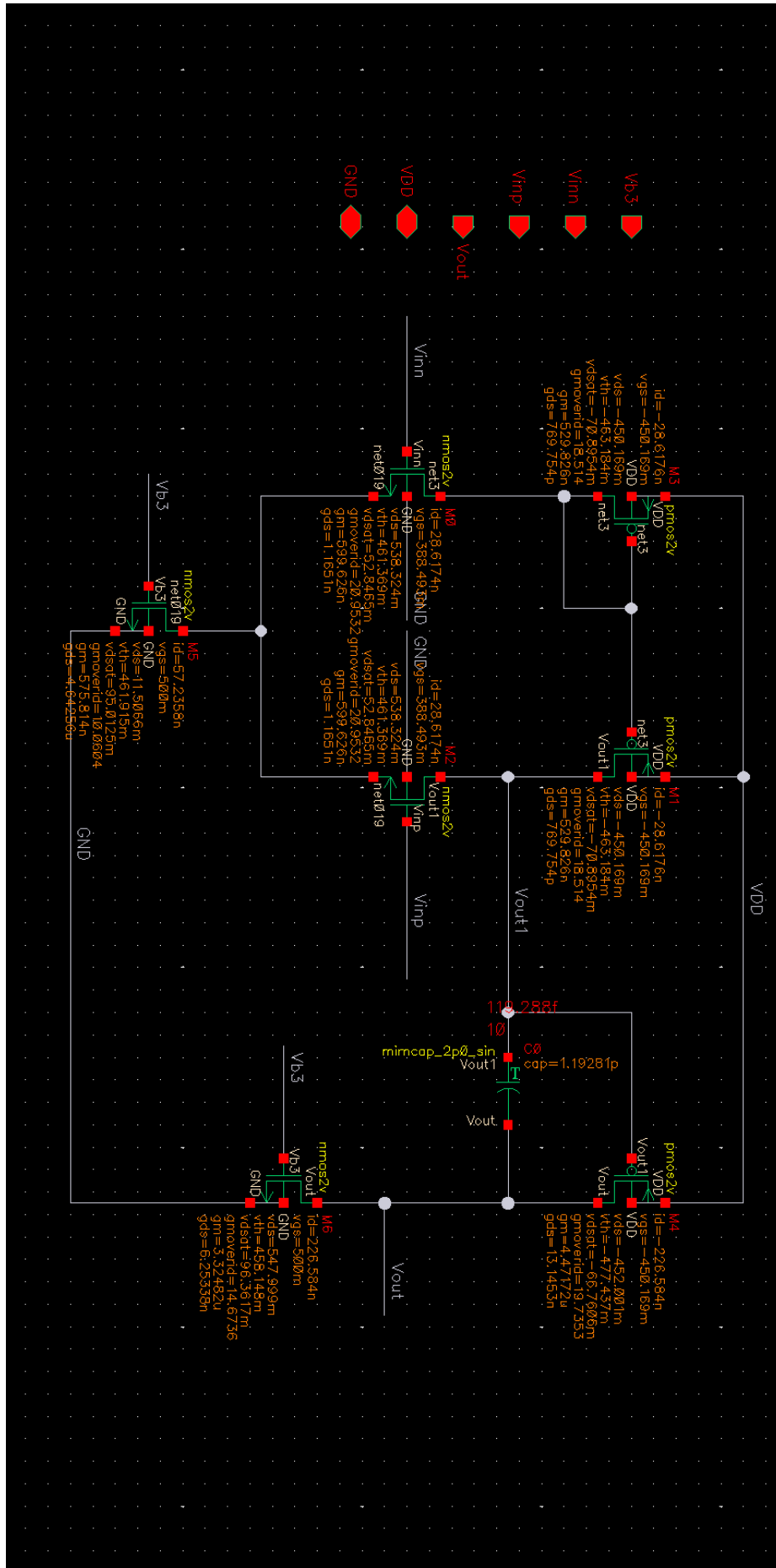
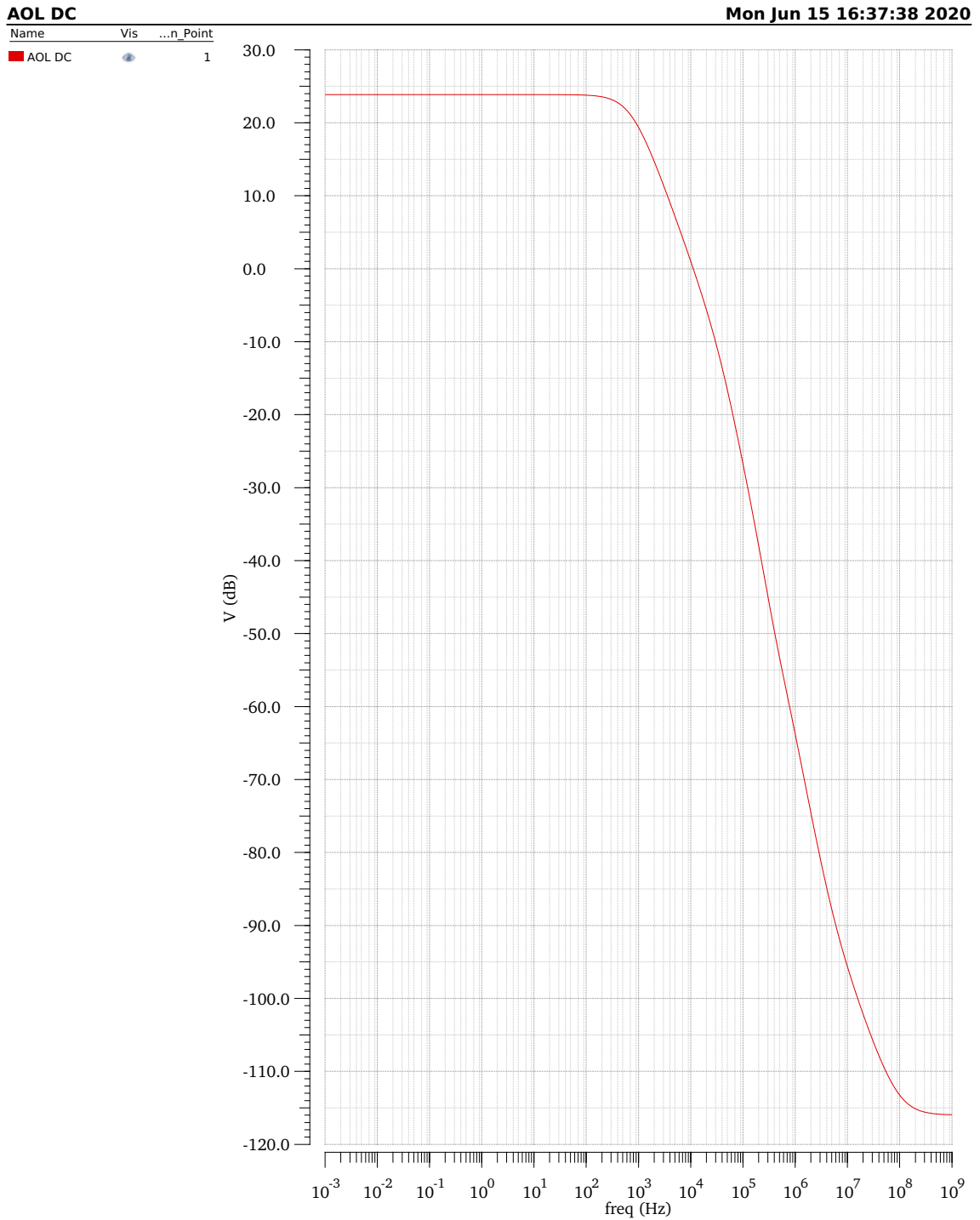Figure 42: g_m/I_D-based two-stage op-amp DC operating points; TSMC 180 nm.
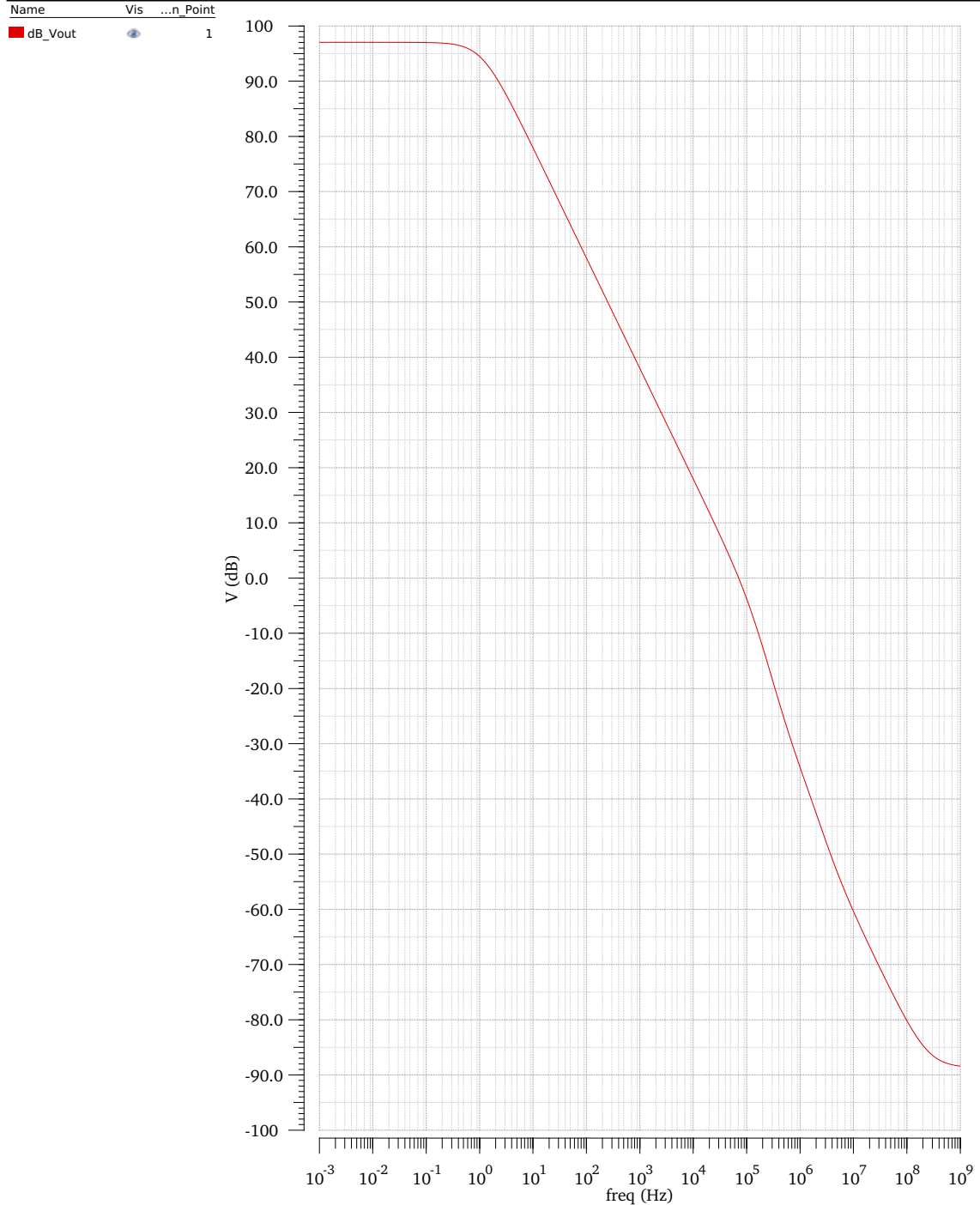
Figure 43: g_m/I_D-based two-stage op-amp gain response; TSMC 180 nm. The magnitude of the AC input test signal is normalised to 1 V. Thus, the V_{out} curve is also the gain response.
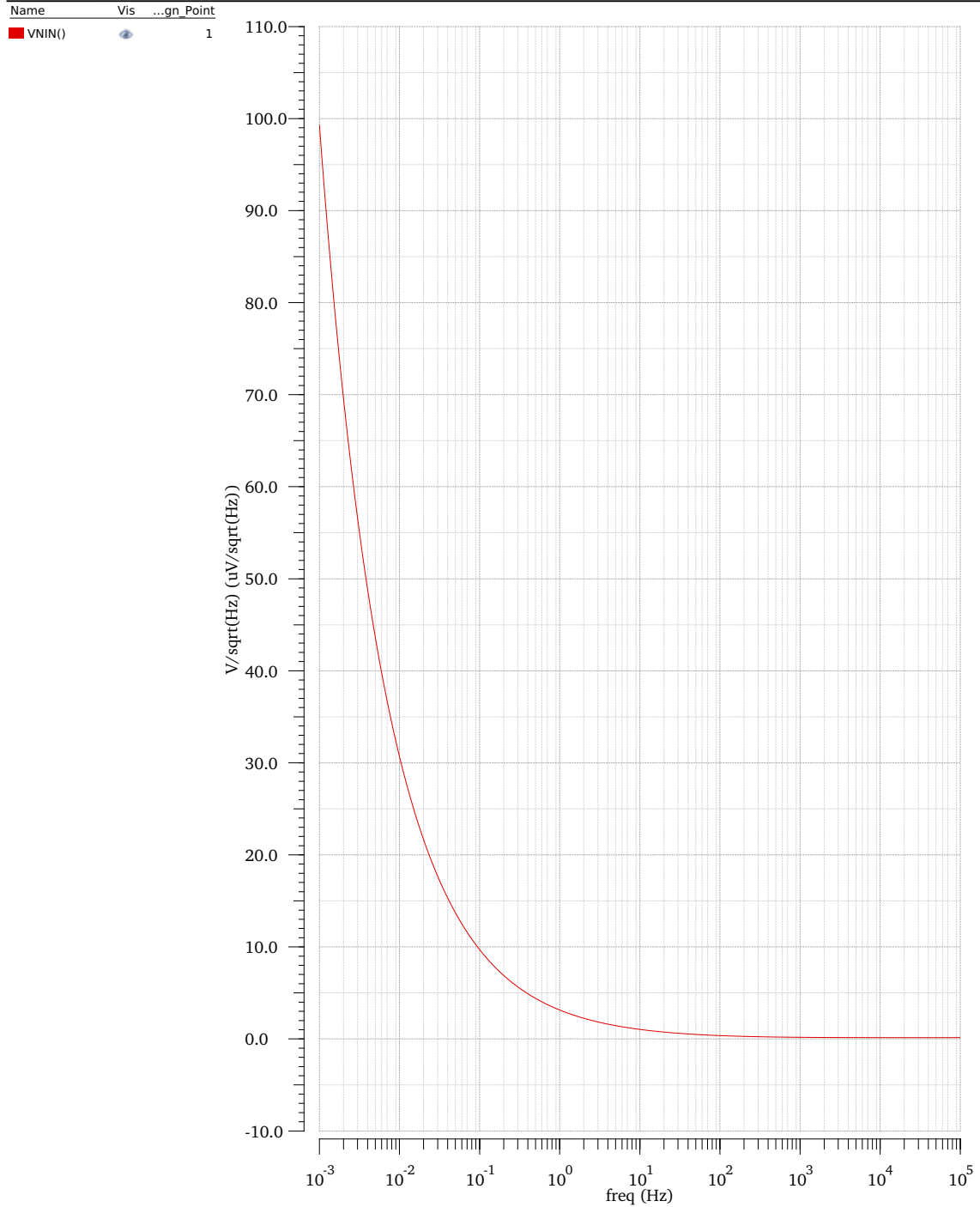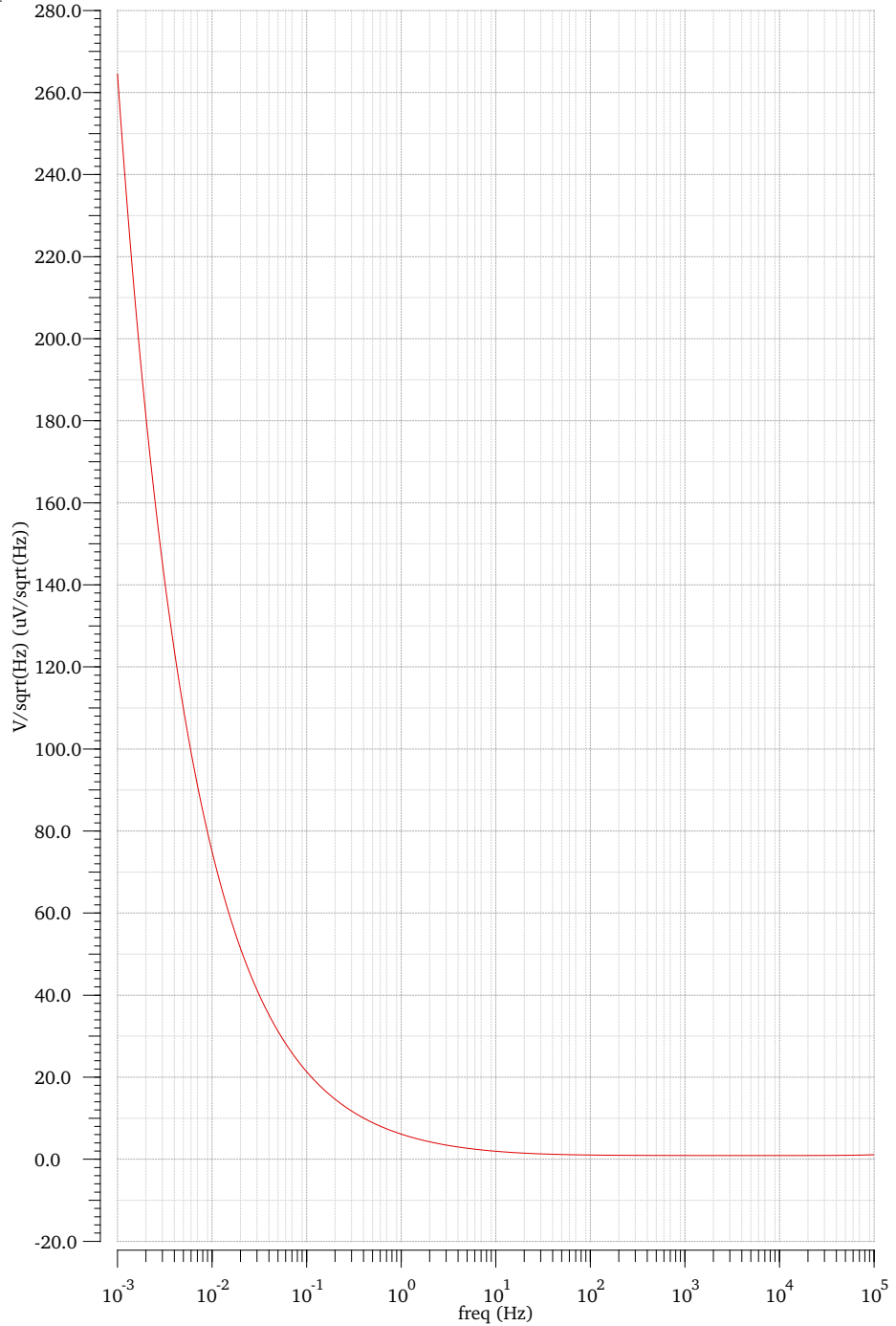
Noise Response

Figure 44: g_m/I_D-based two-stage op-amp input-referred noise response; TSMC 180 nm.

D g_m/I_D-based Folded Cascode Op-amp Cadence Screenshots (TSMC 180 nm)

See below.
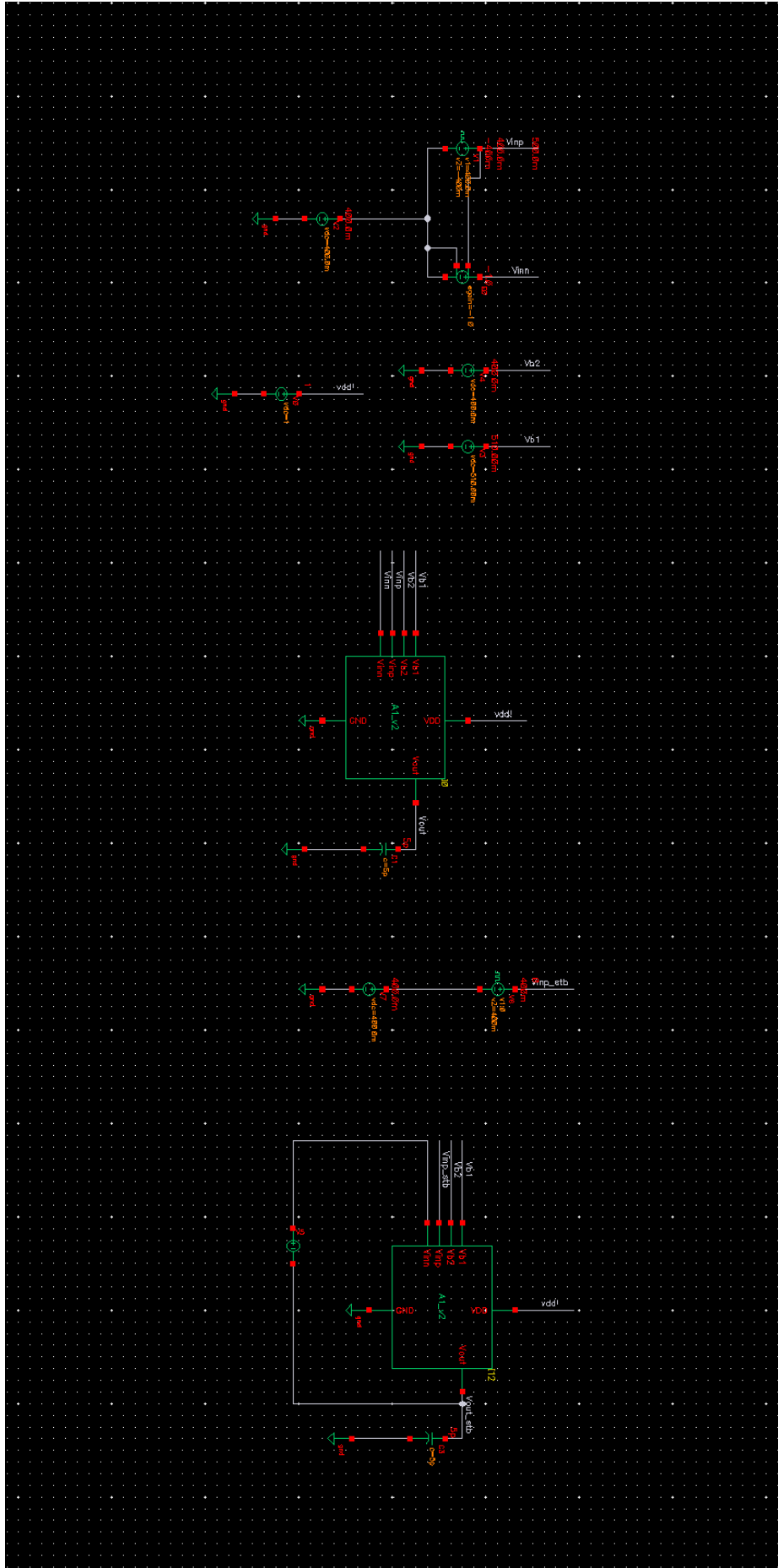
Figure 45: g_m/I_D-based folded cascode op-amp test bench schematic; TSMC 180 nm.
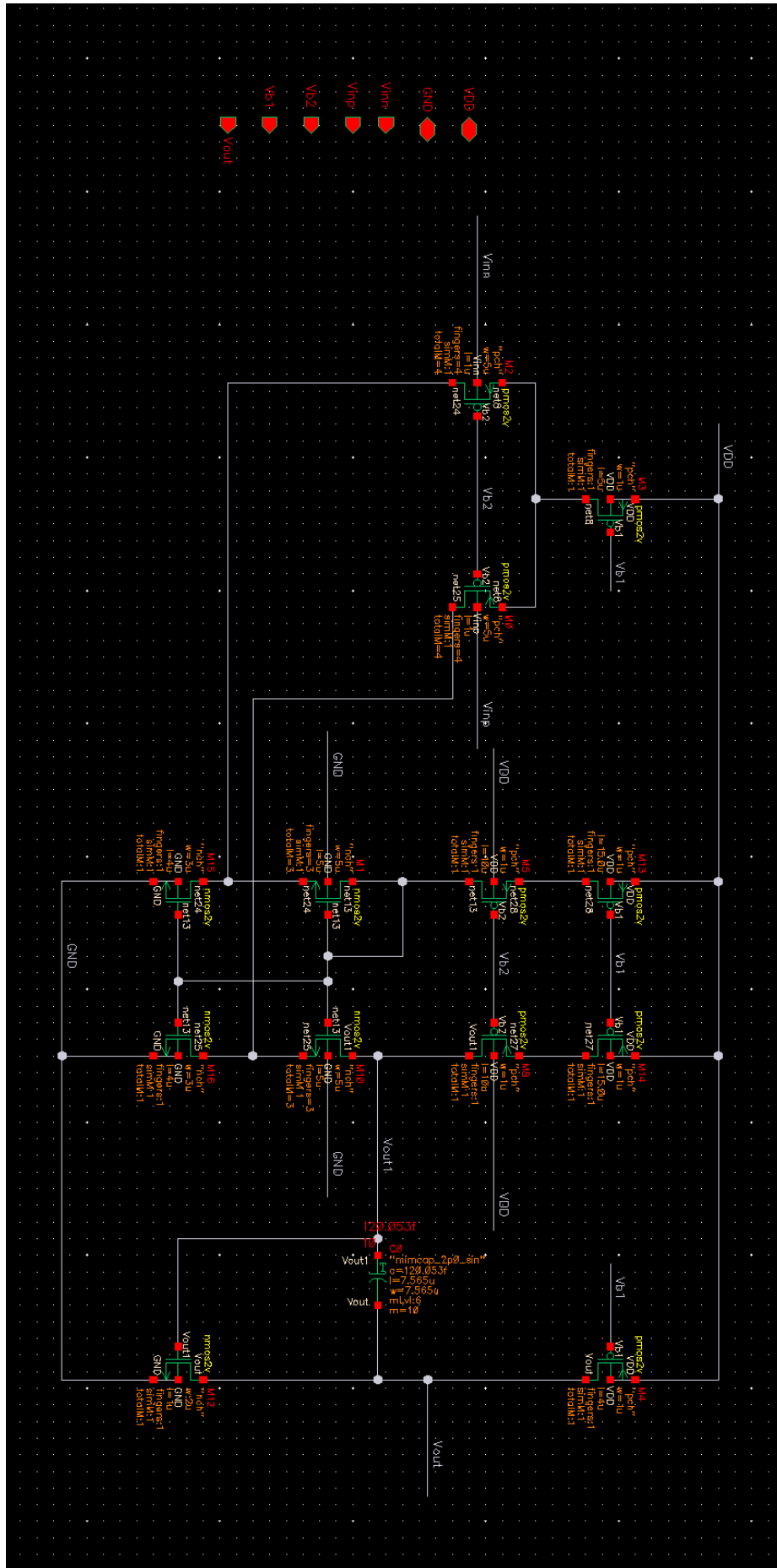
Figure 46: g_m/I_D-based folded cascode op-amp transistor dimensions; TSMC 180 nm.
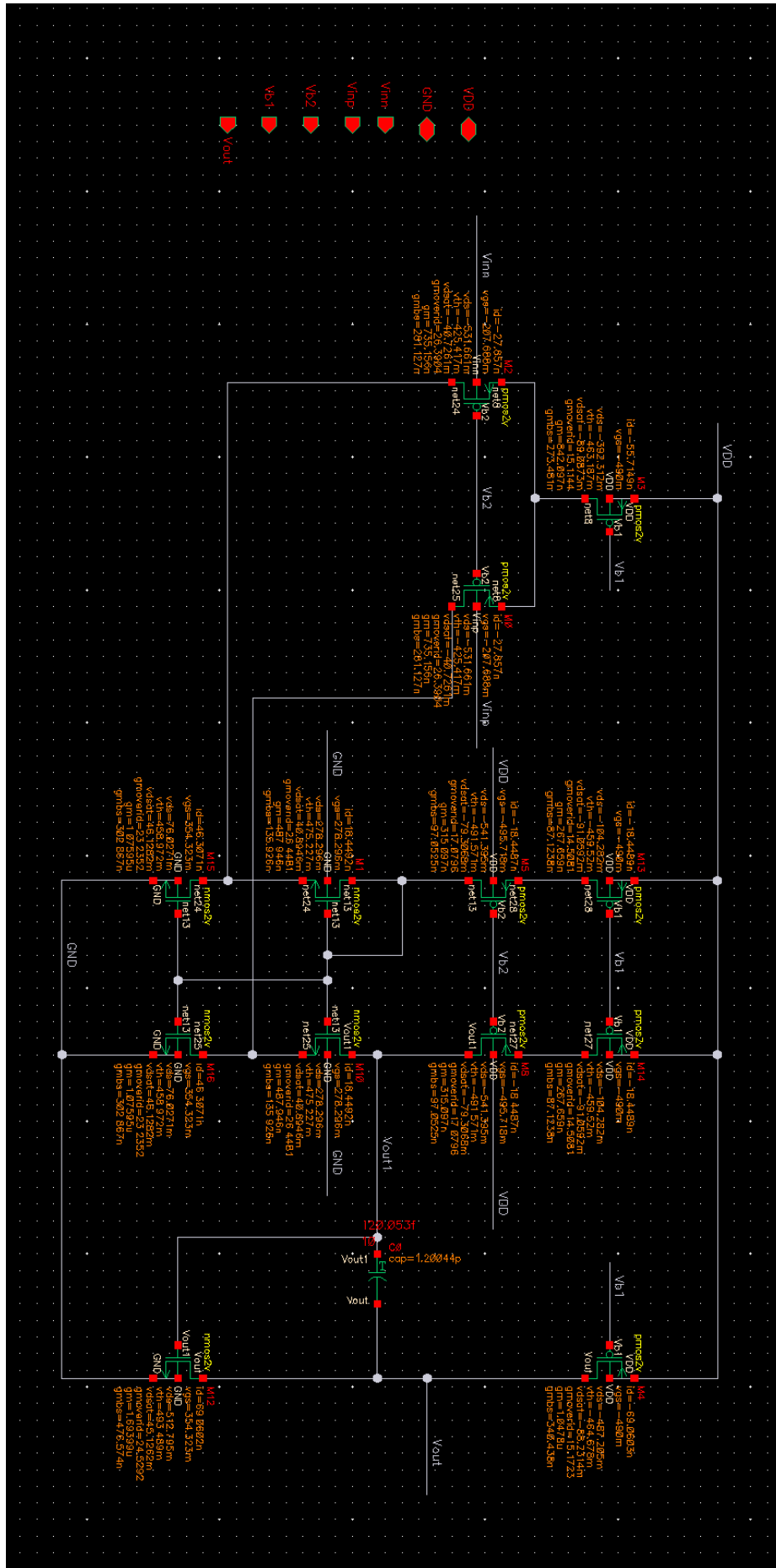
Figure 47: g_m/I_D-based folded cascode op-amp DC operating points; TSMC 180 nm. Note that g_{mb} is also displayed.
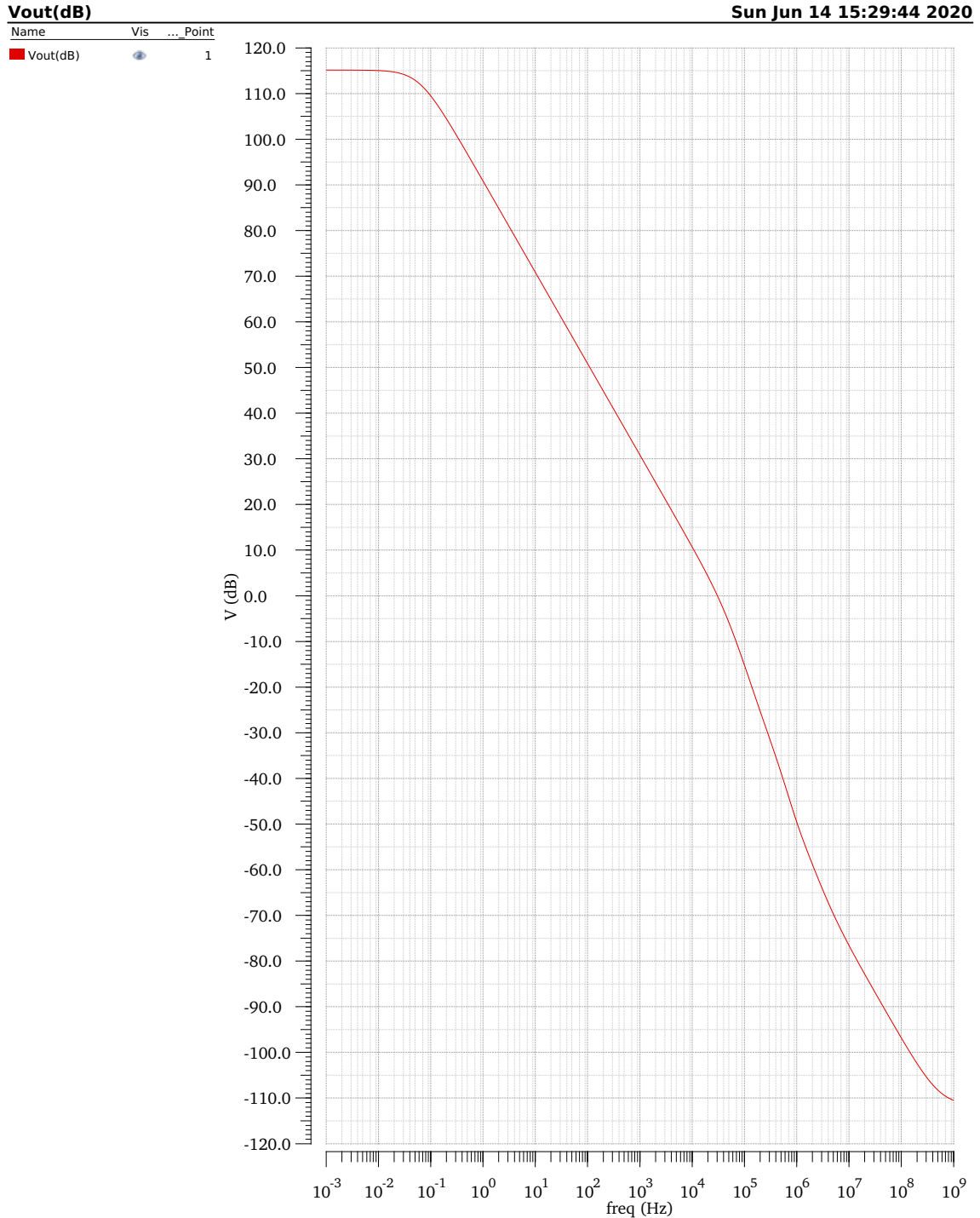
Figure 48: g_m/I_D-based folded cascode op-amp gain response; TSMC 180 nm. The magnitude of the AC input test signal is normalised to 1 V. Thus, the V_{out} curve is also the gain response.

Figure 49: g_m/I_D-based folded cascode op-amp noise response; TSMC 180 nm.

E g_m/I_D-based Two-Stage Op-amp Cadence Screenshots (TSMC 65 nm)

See below.
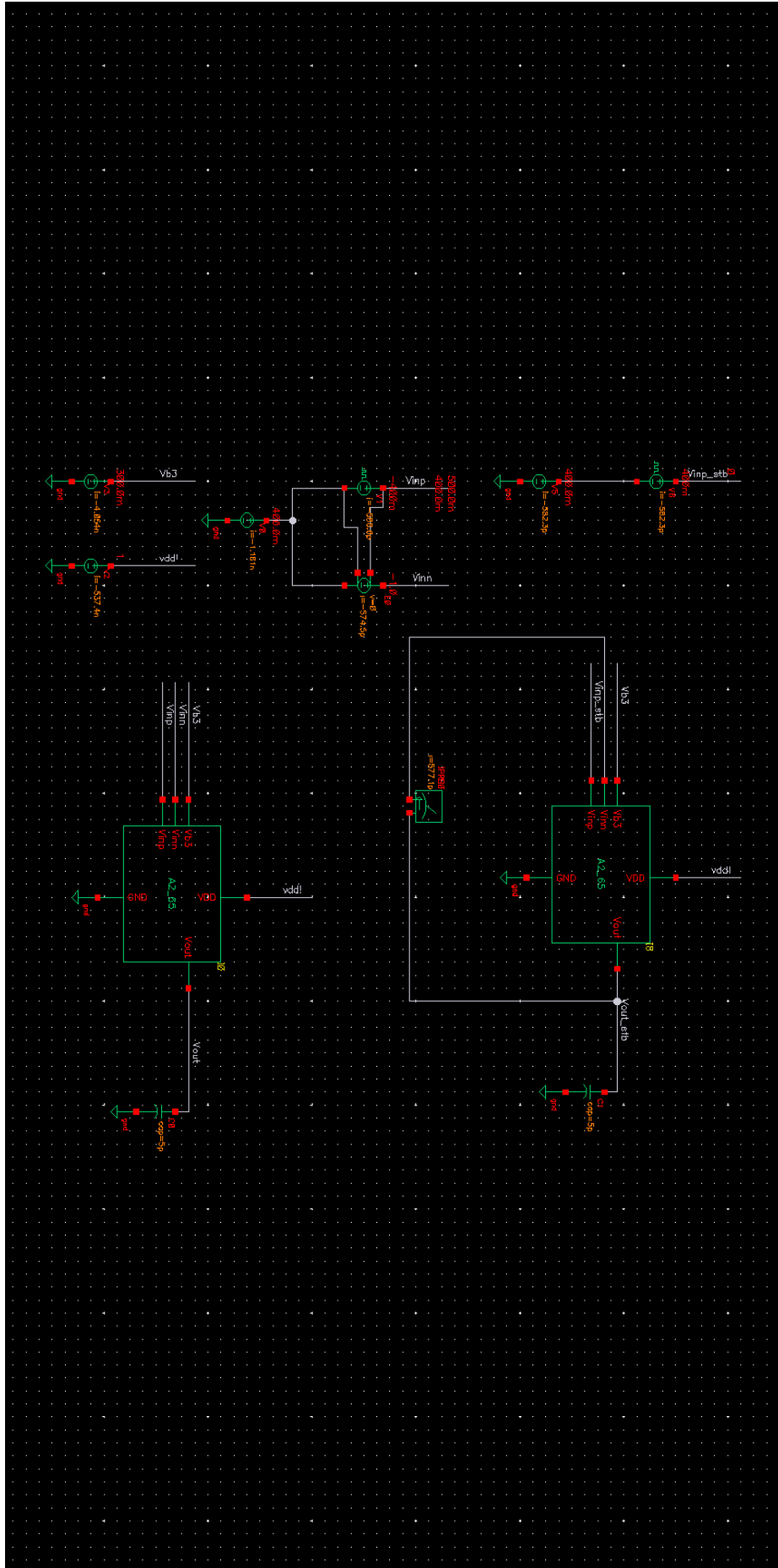
Figure 50: g_m/I_D-based two-stage op-amp test bench schematic; TSMC 65 nm.
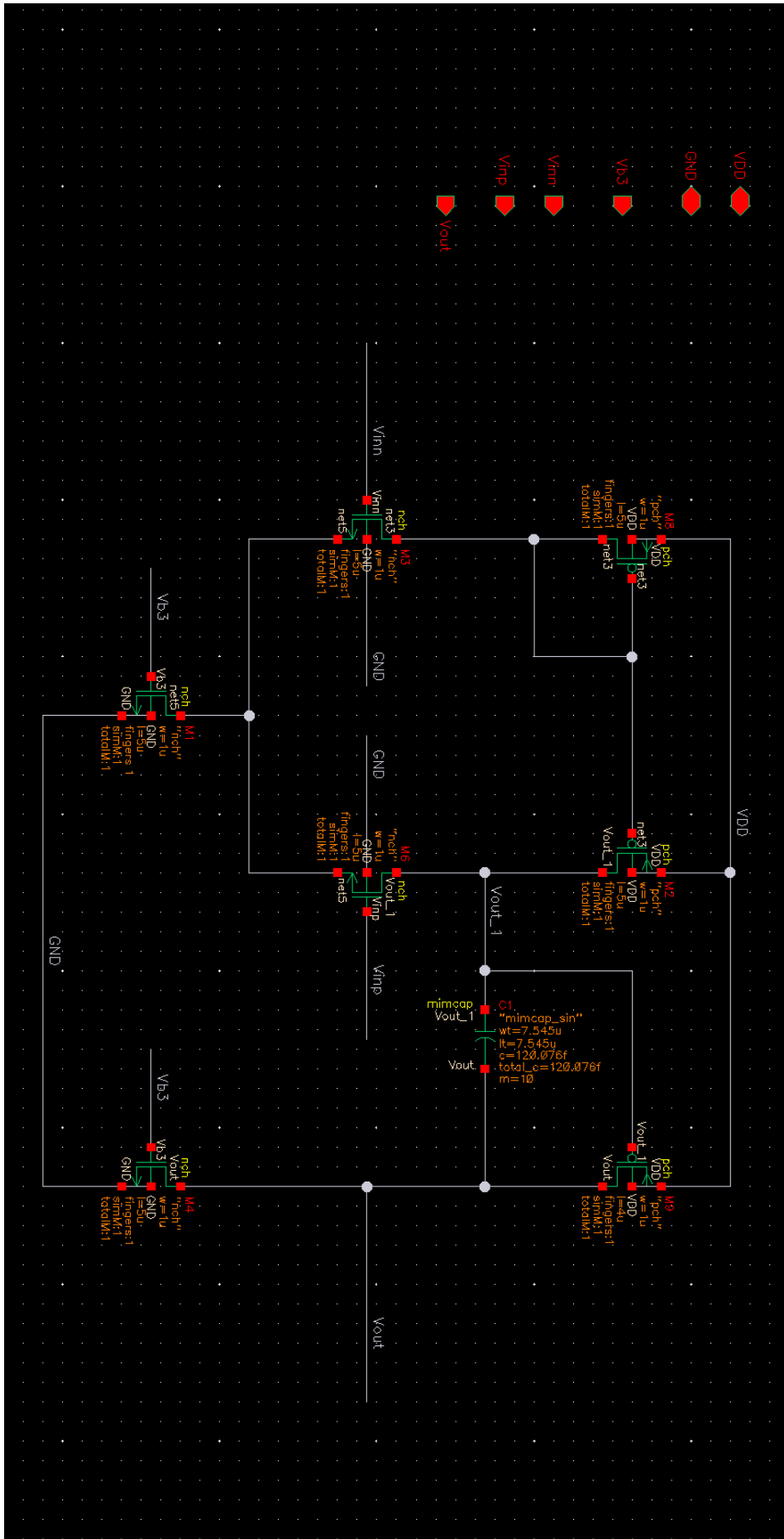
Figure 51: g_m/I_D-based two-stage op-amp transistor dimensions; TSMC 65 nm.
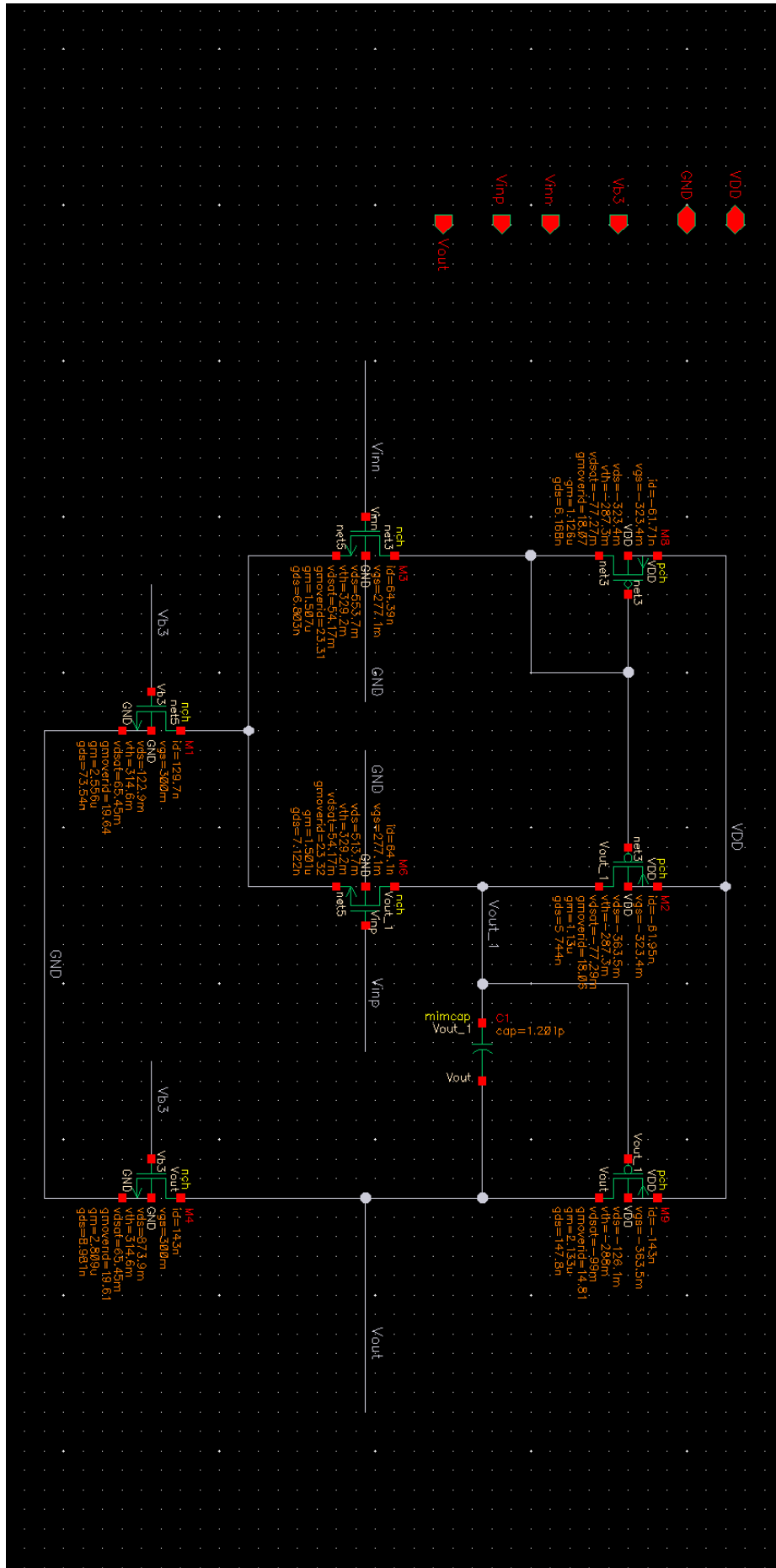
Figure 52: g_m/I_D-based two-stage op-amp DC operating points; TSMC 65 nm.
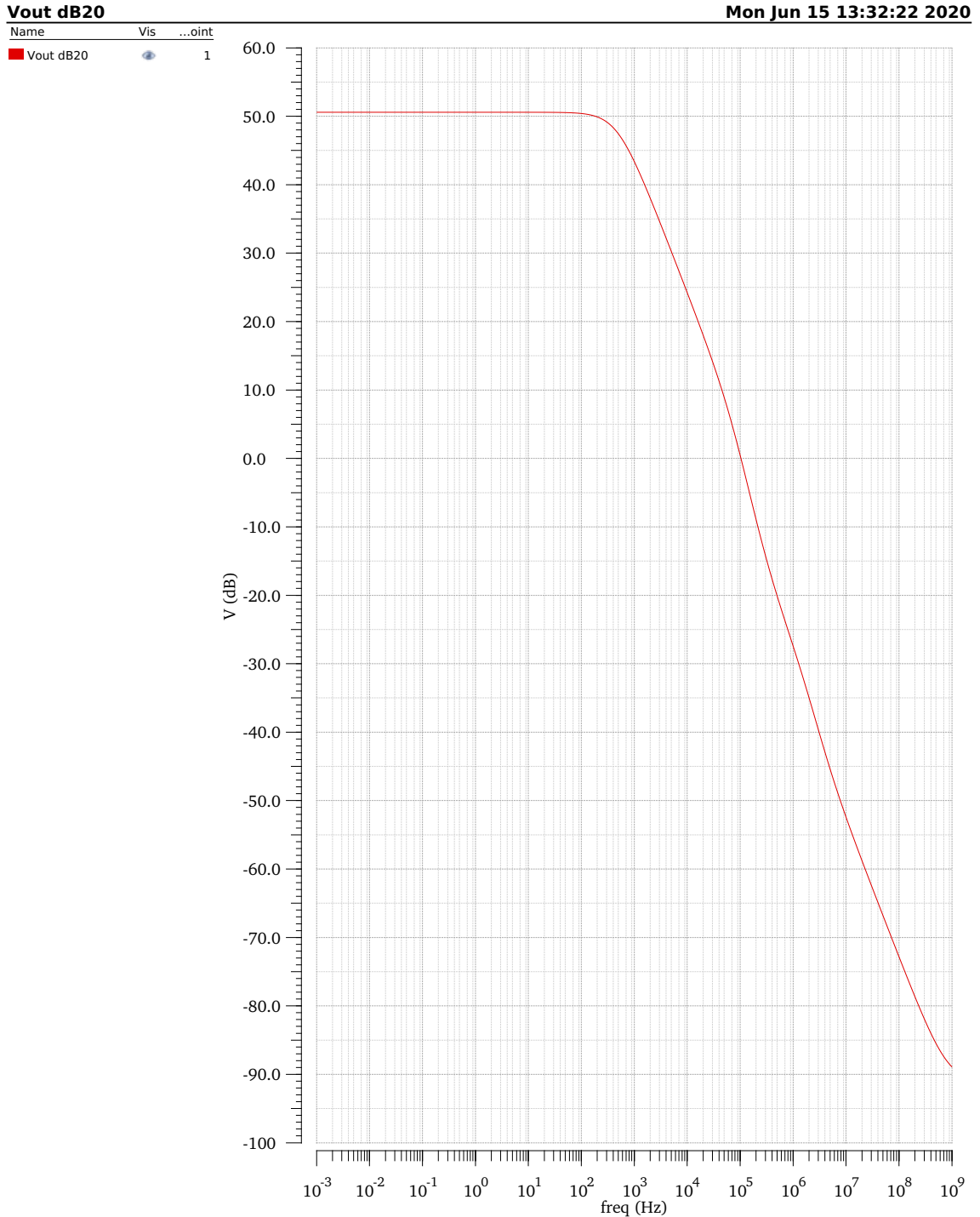
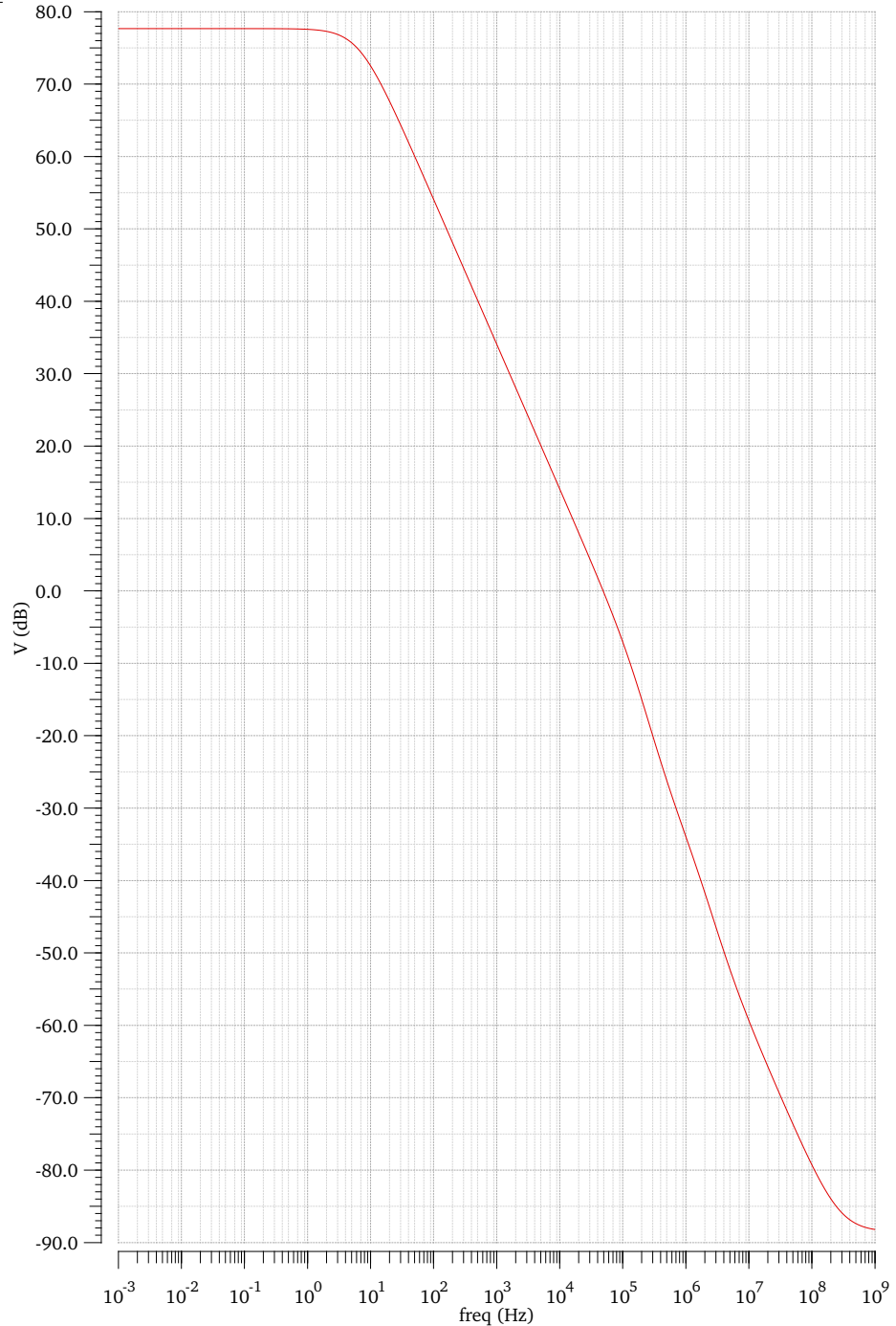Figure 53: g_m/I_D-based two-stage op-amp op-amp gain response; TSMC 65 nm. The magnitude of the AC input test signal is normalised to 1 V. Thus, the V_{out} curve is also the gain response.

Noise Response

Name	Vis	...gn_Point
■ VNIN()	👁	1
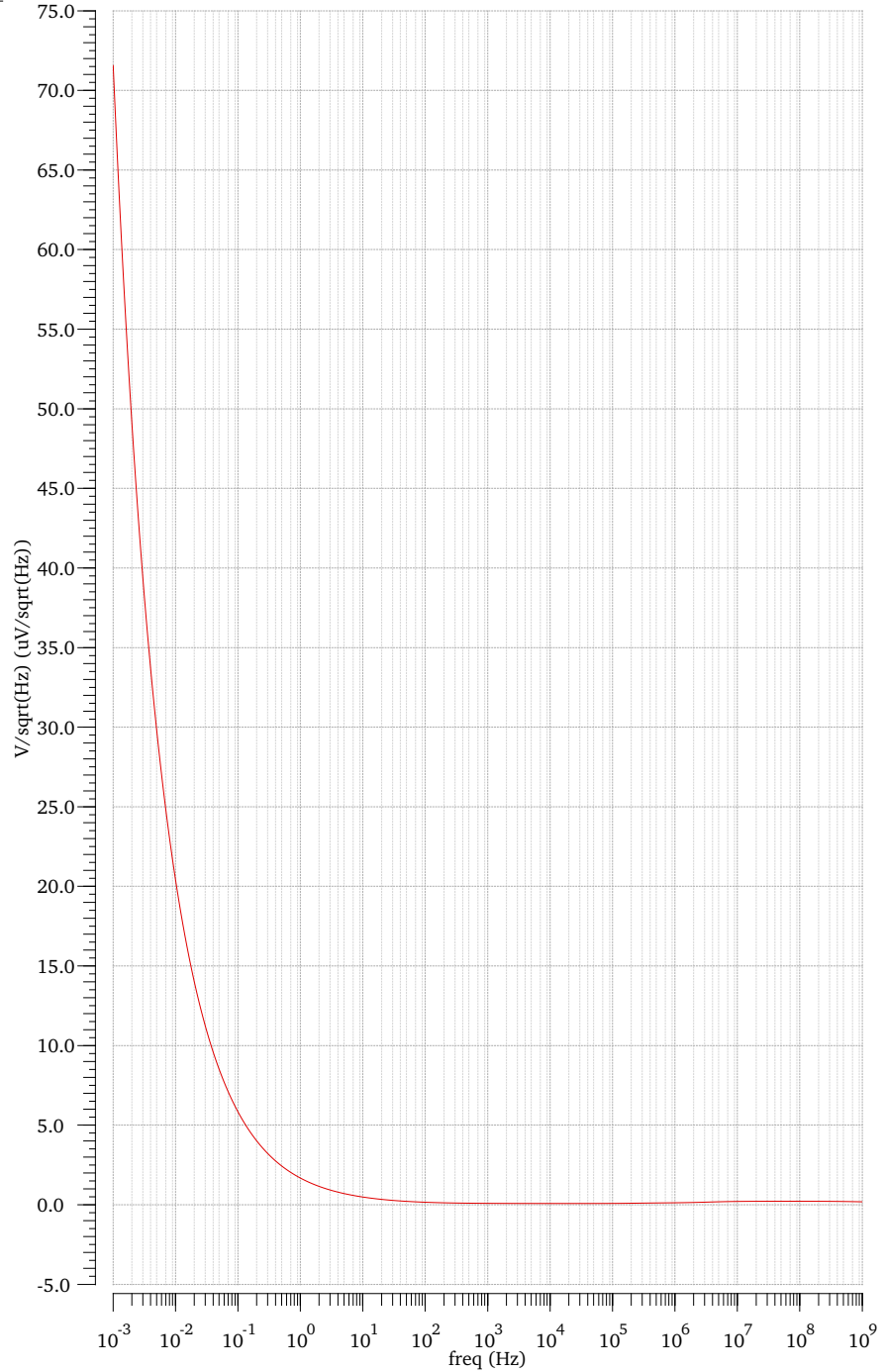
Figure 54: g_m/I_D-based two-stage op-amp op-amp noise response; TSMC 65 nm.

F g_m/I_D-based Folded Cascode Op-amp Cadence Screenshots (TSMC 65 nm)

See below.
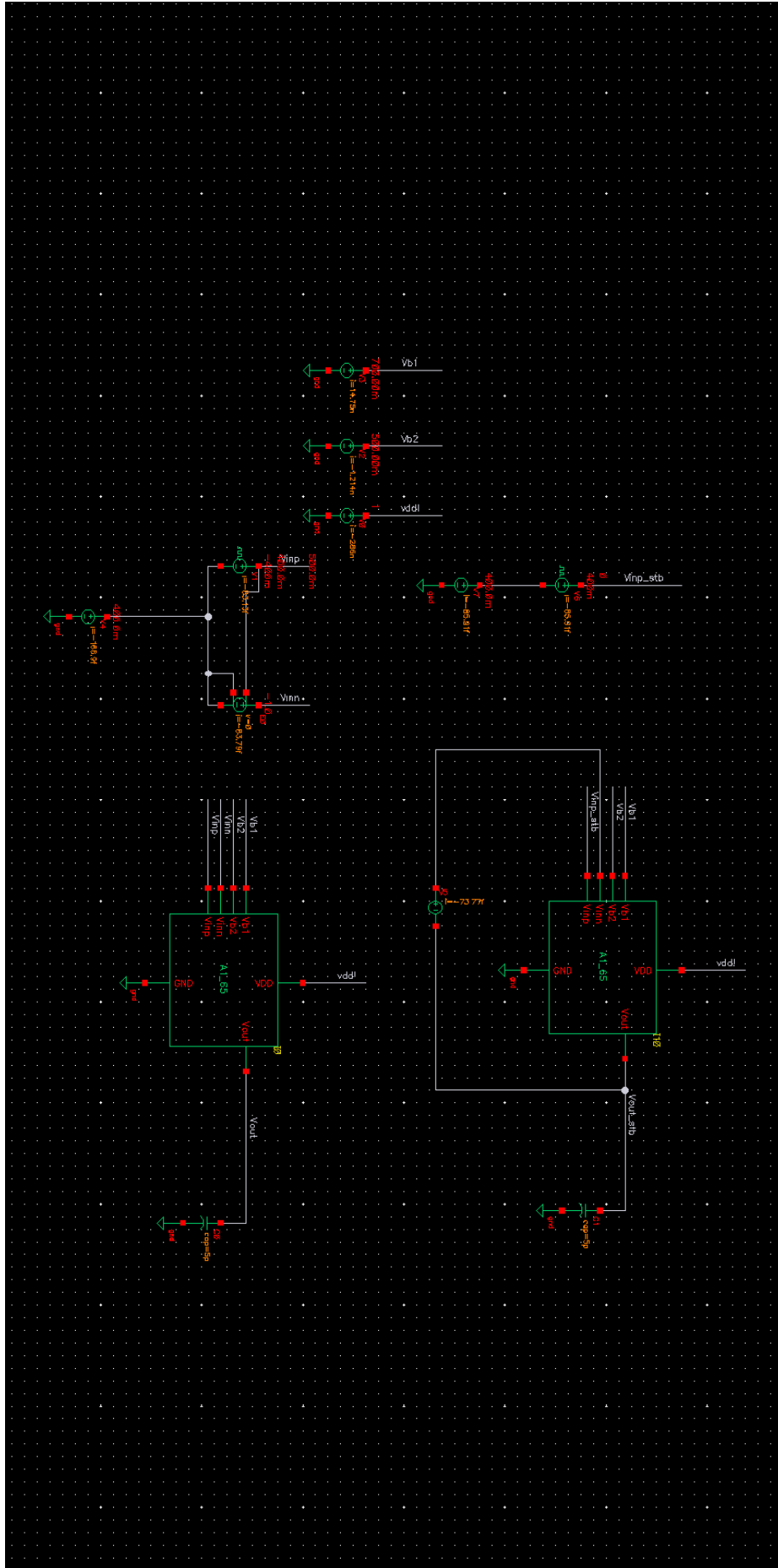
Figure 55: g_m/I_D-based folded cascode op-amp test bench schematic; TSMC 65 nm.
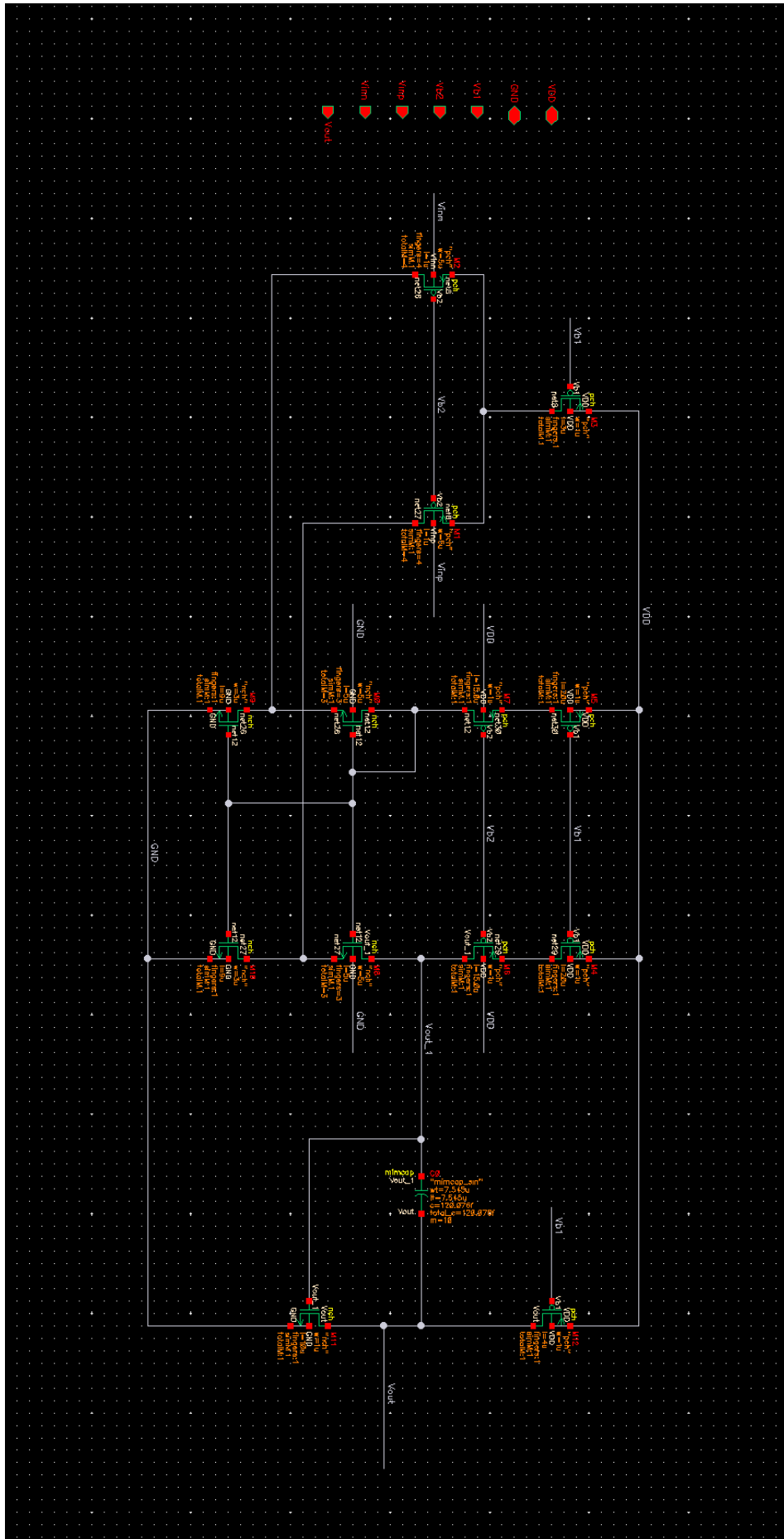
Figure 56: g_m/I_D-based folded cascode op-amp transistor dimensions; TSMC 65 nm.
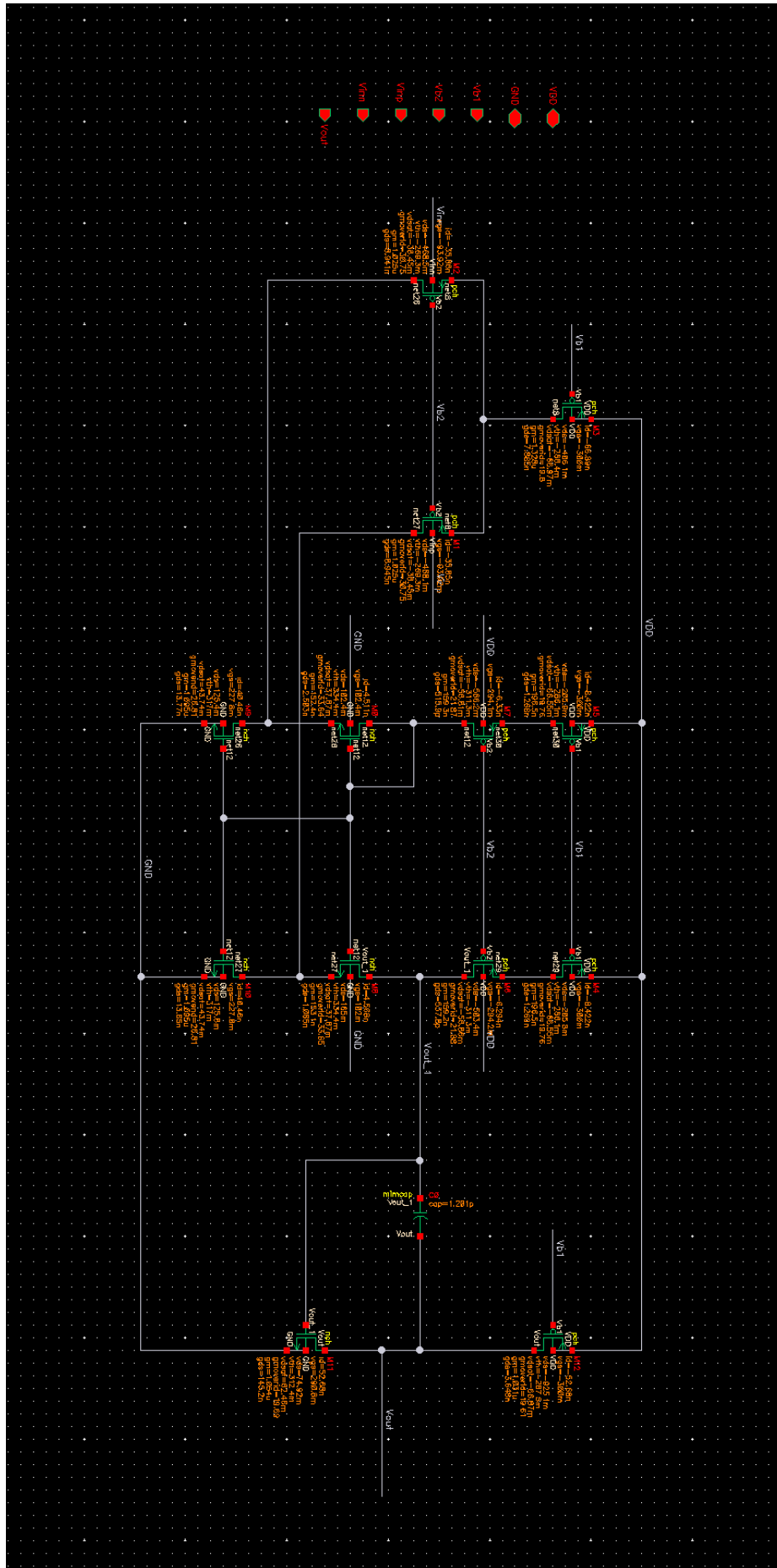
Figure 57: g_m/I_D-based folded cascode op-amp DC operating points; TSMC 65 nm.

Figure 58: g_m/I_D-based folded cascode op-amp op-amp gain response; TSMC 65 nm. The magnitude of the AC input test signal is normalised to 1 V. Thus, the V_{out} curve is also the gain response.
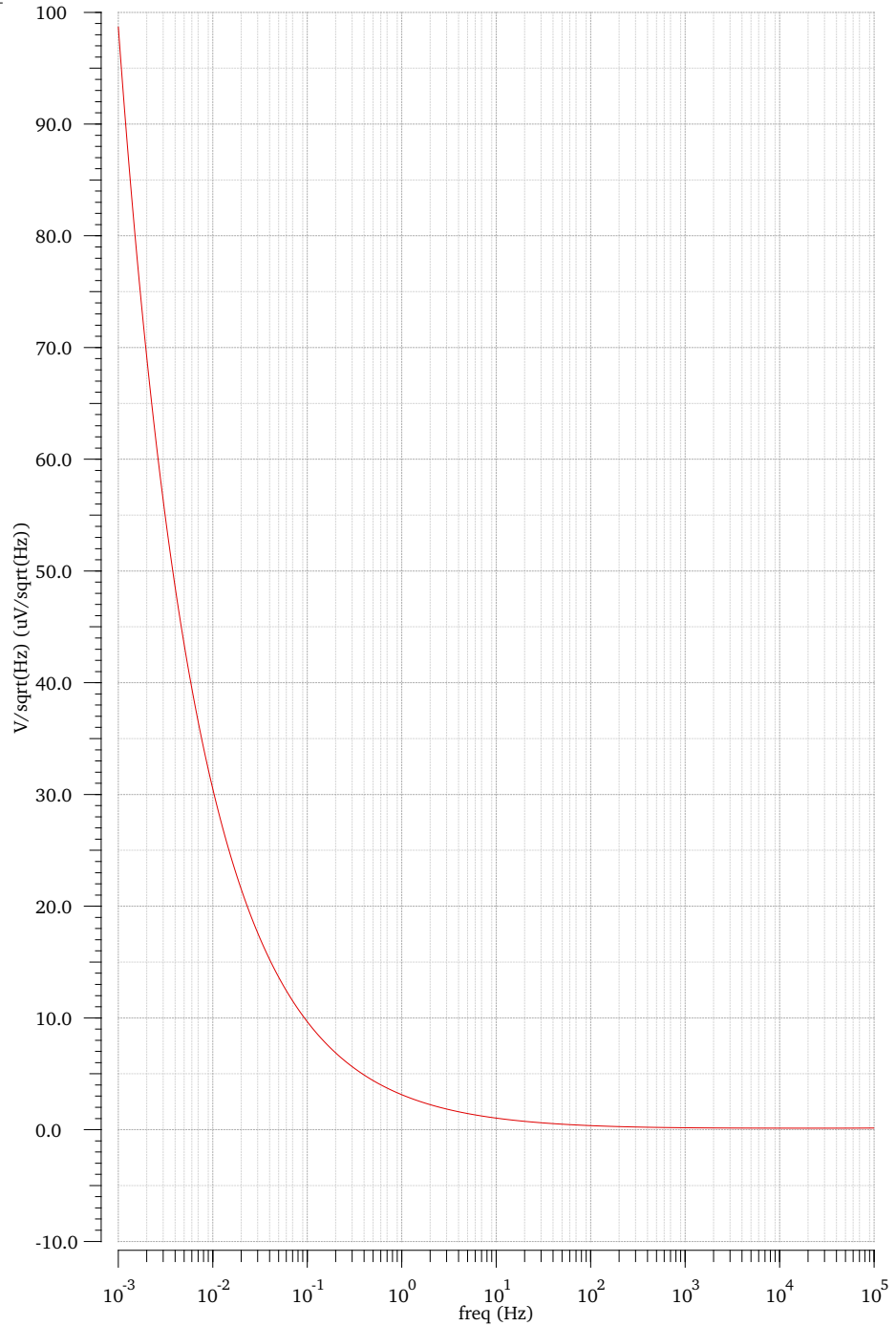
Noise Response

Figure 59: g_m/I_D-based folded cascode op-amp op-amp noise response; TSMC 65 nm.

G High-Level Optimised Potentiostat Test bench Cadence Screenshots (TSMC 180 nm)
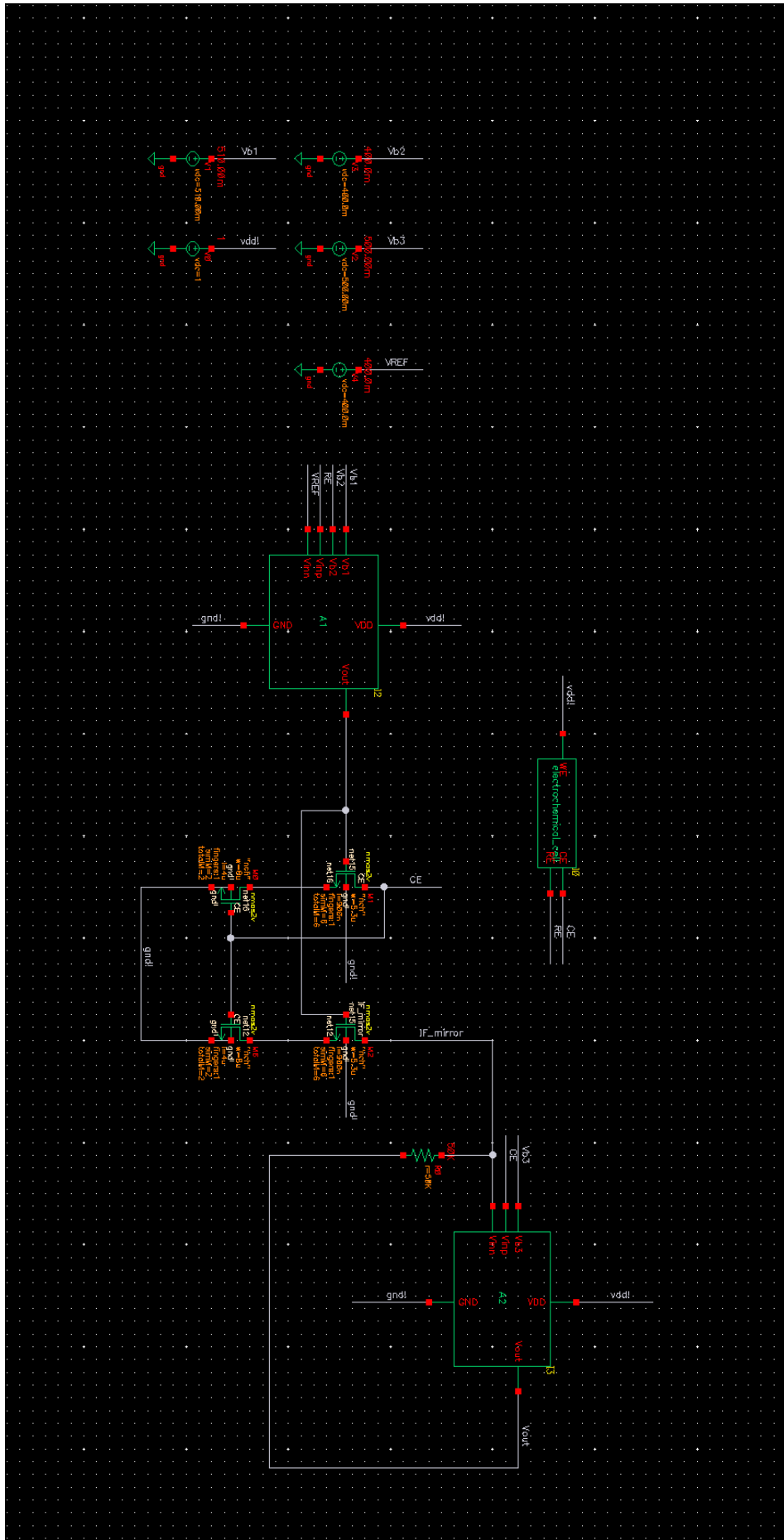
See below.

Figure 60: High-level optimised potentiostat test bench; TSMC 180 nm.

H Original Two-Stage Op-amp Cadence Screenshots (TSMC 180 nm)

See below.
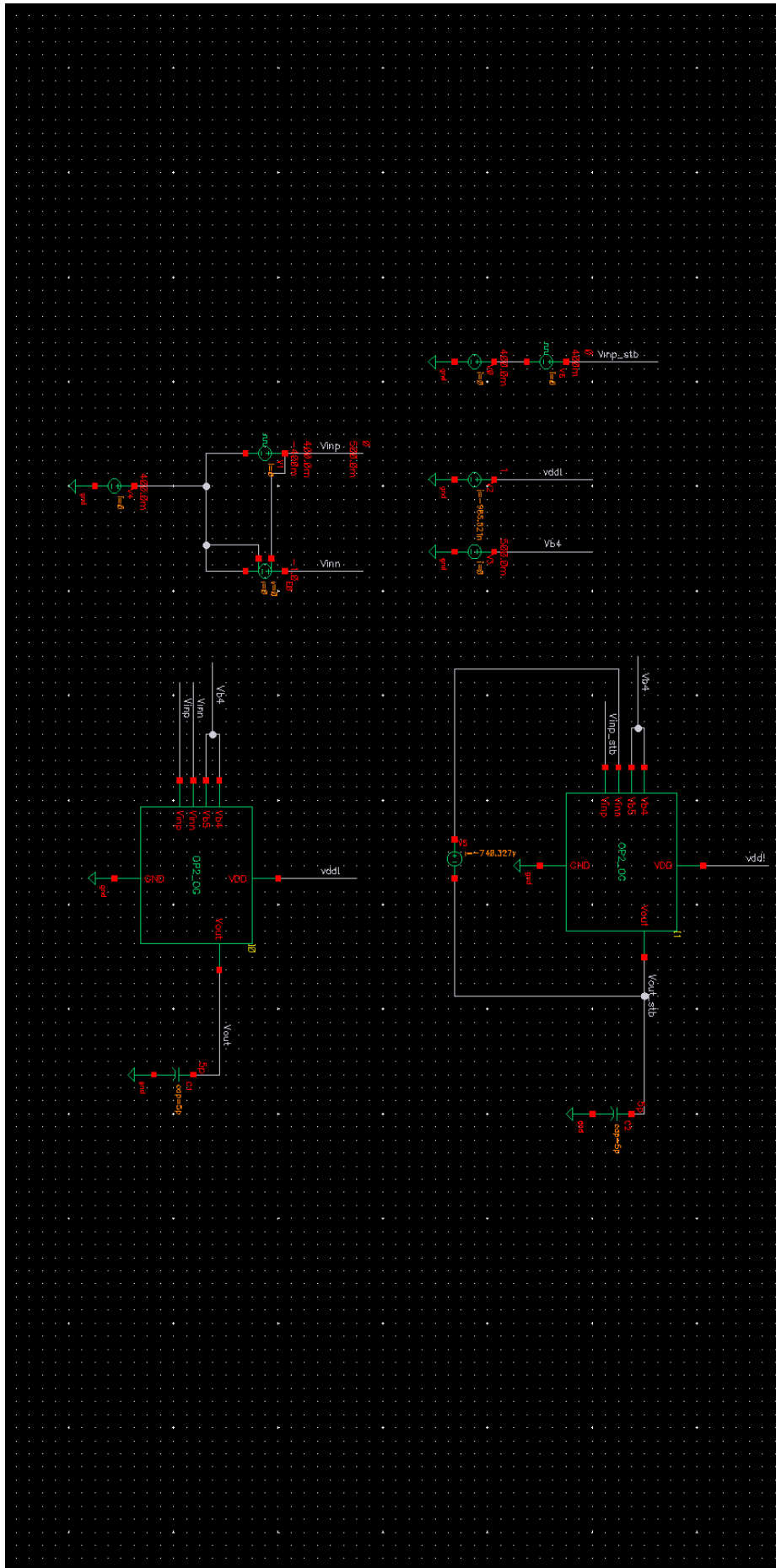
Figure 61: Original two-stage op-amp test bench schematic; TSMC 180 nm.
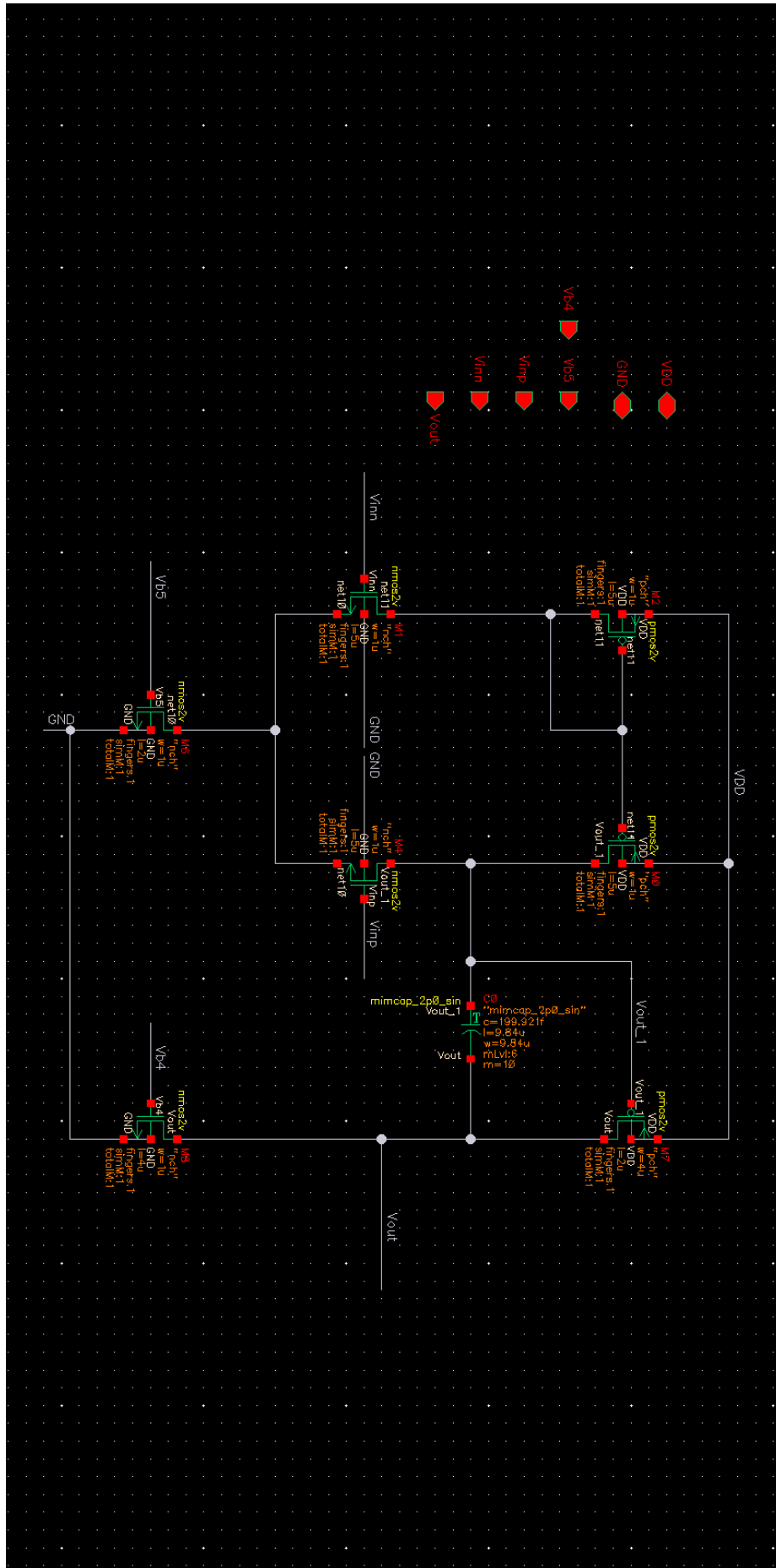
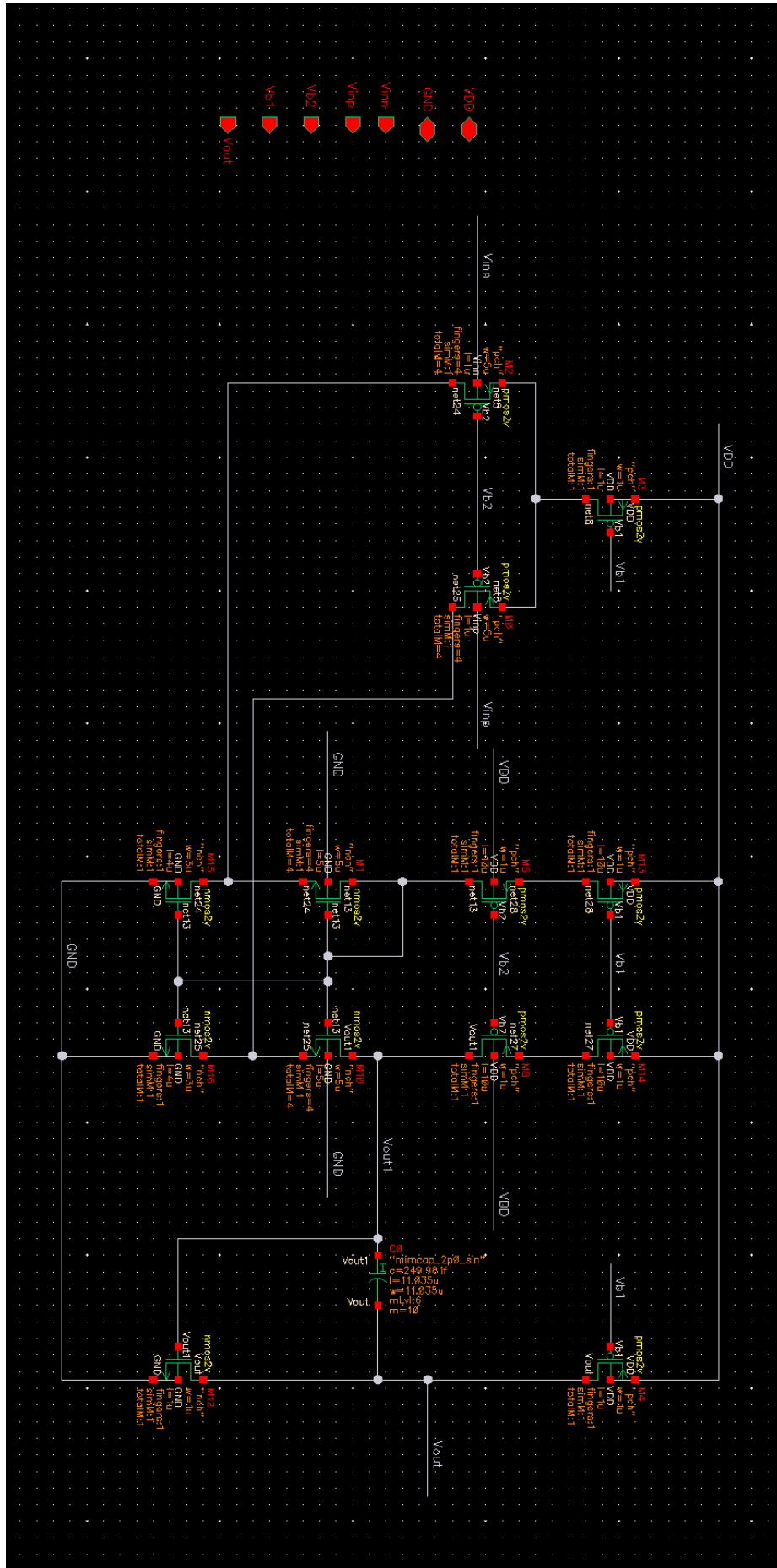Figure 62: Original two-stage op-amp transistor dimensions; TSMC 180 nm.
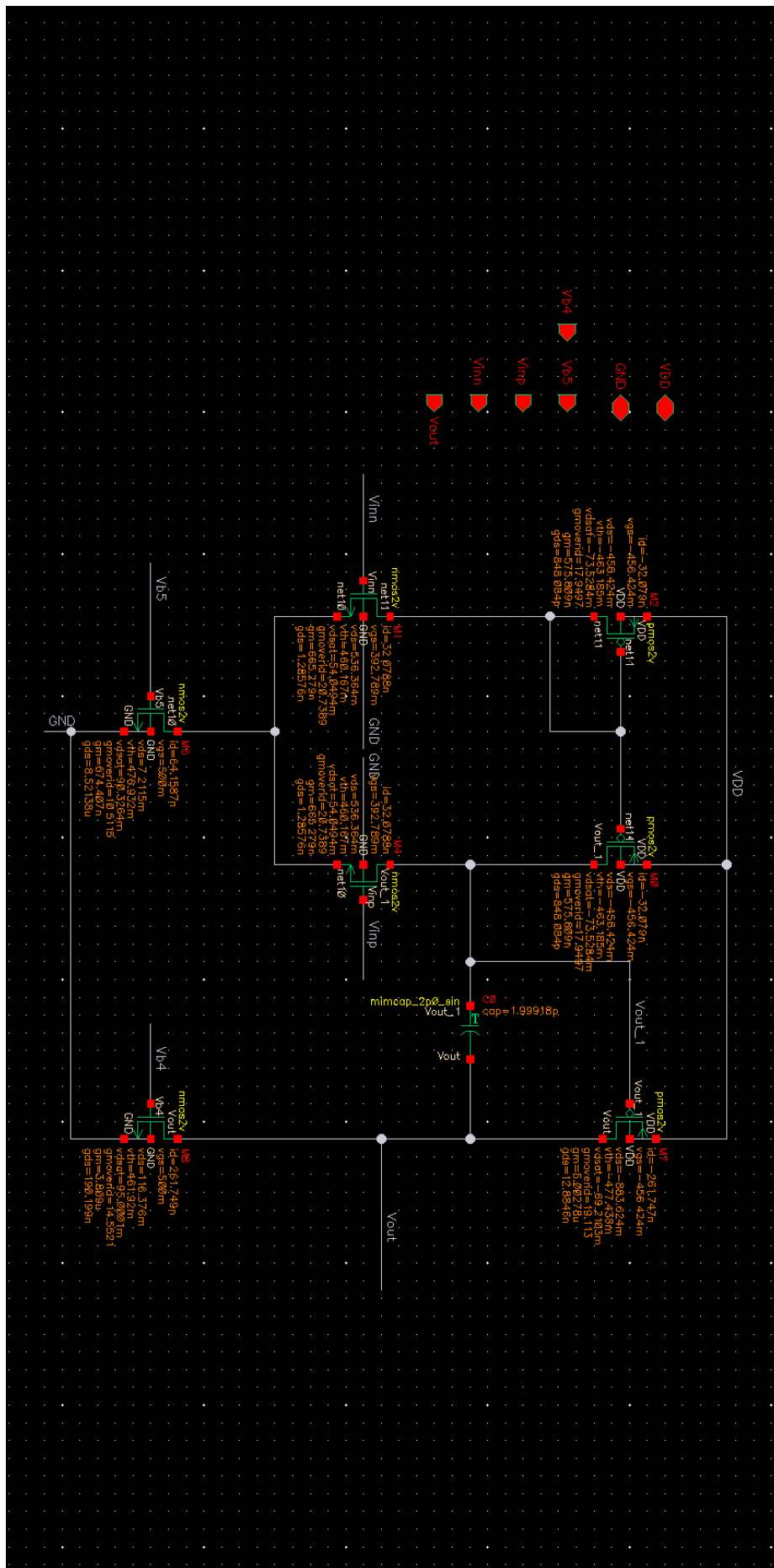
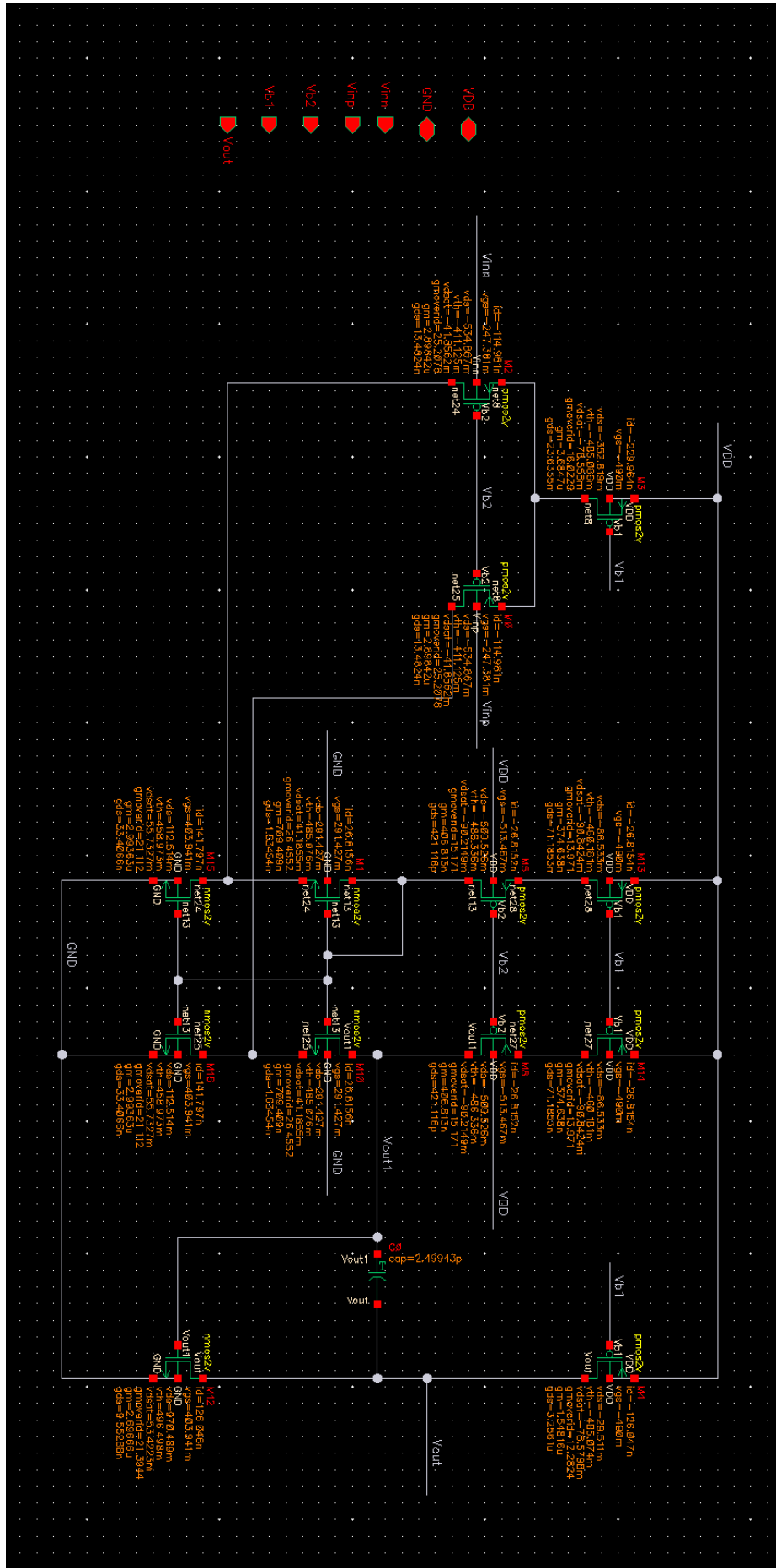Figure 63: Original two-stage op-amp DC operating points; TSMC 180 nm.
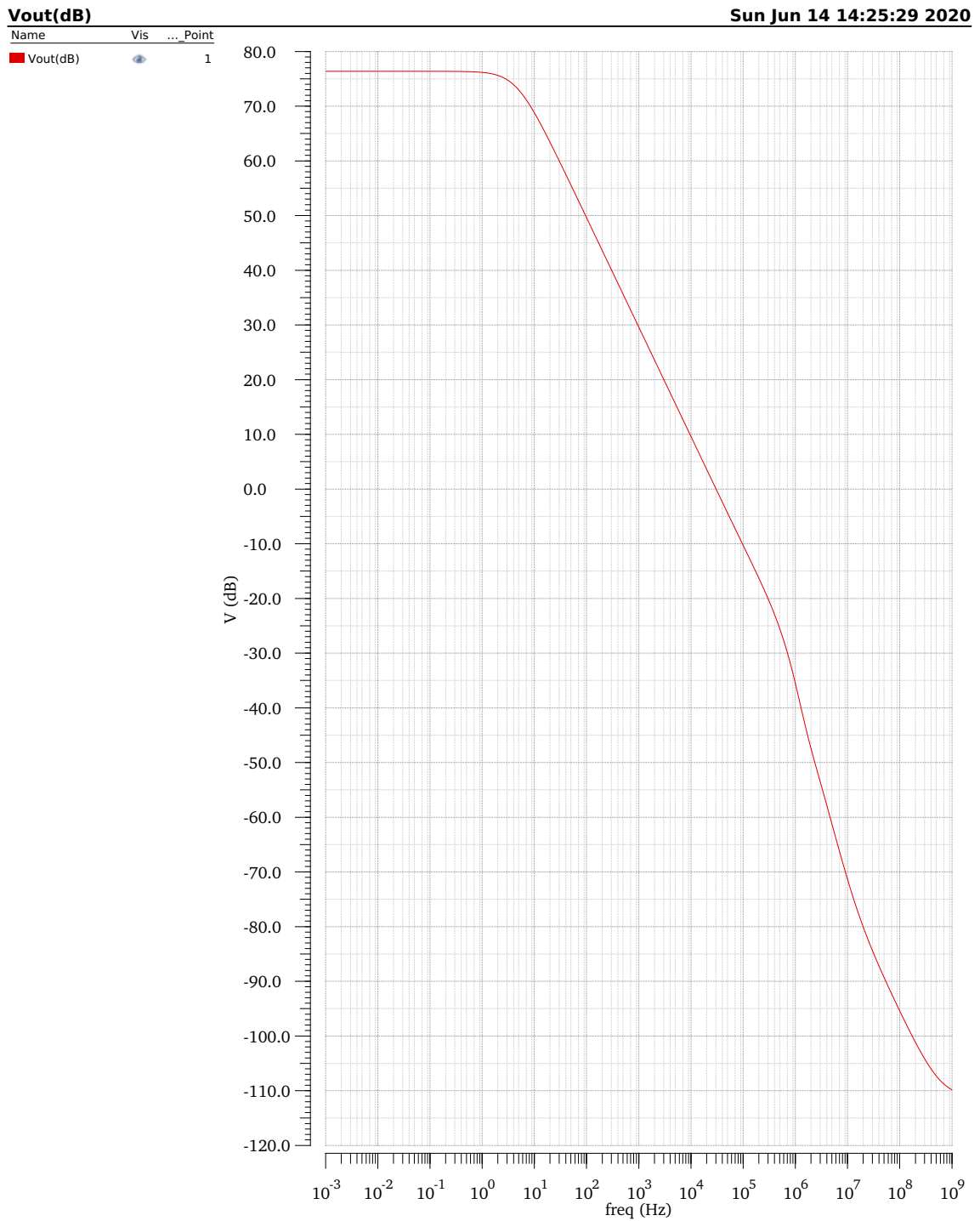
Figure 64: Original two-stage op-amp gain response; TSMC 180 nm. The magnitude of the AC input test signal is normalised to 1 V. Thus, the V_{out} curve is also the gain response.

Noise Response

Figure 65: Original two-stage op-amp input-referred noise response; TSMC 180 nm.

I Original Folded Cascode Op-amp Cadence Screenshots (TSMC 180 nm)

See below.
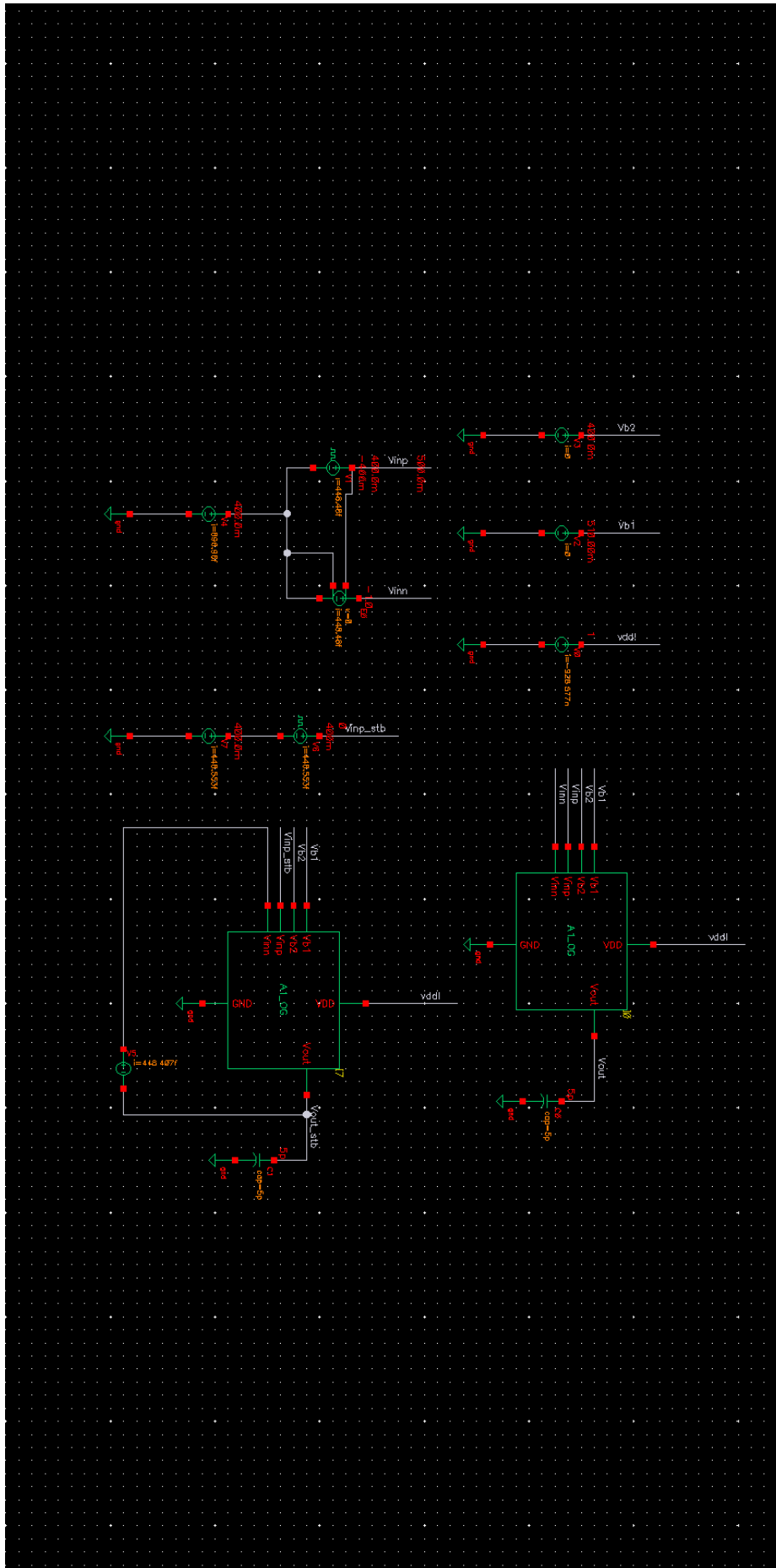
Figure 66: Original folded cascode op-amp test bench schematic; TSMC 180 nm.

Figure 67: Original folded cascode op-amp transistor dimensions; TSMC 180 nm.

Figure 68: Original folded cascode op-amp DC operating points; TSMC 180 nm.

Figure 69: Original folded cascode op-amp op-amp gain response; TSMC 180 nm. The magnitude of the AC input test signal is normalised to 1 V. Thus, the V_{out} curve is also the gain response.
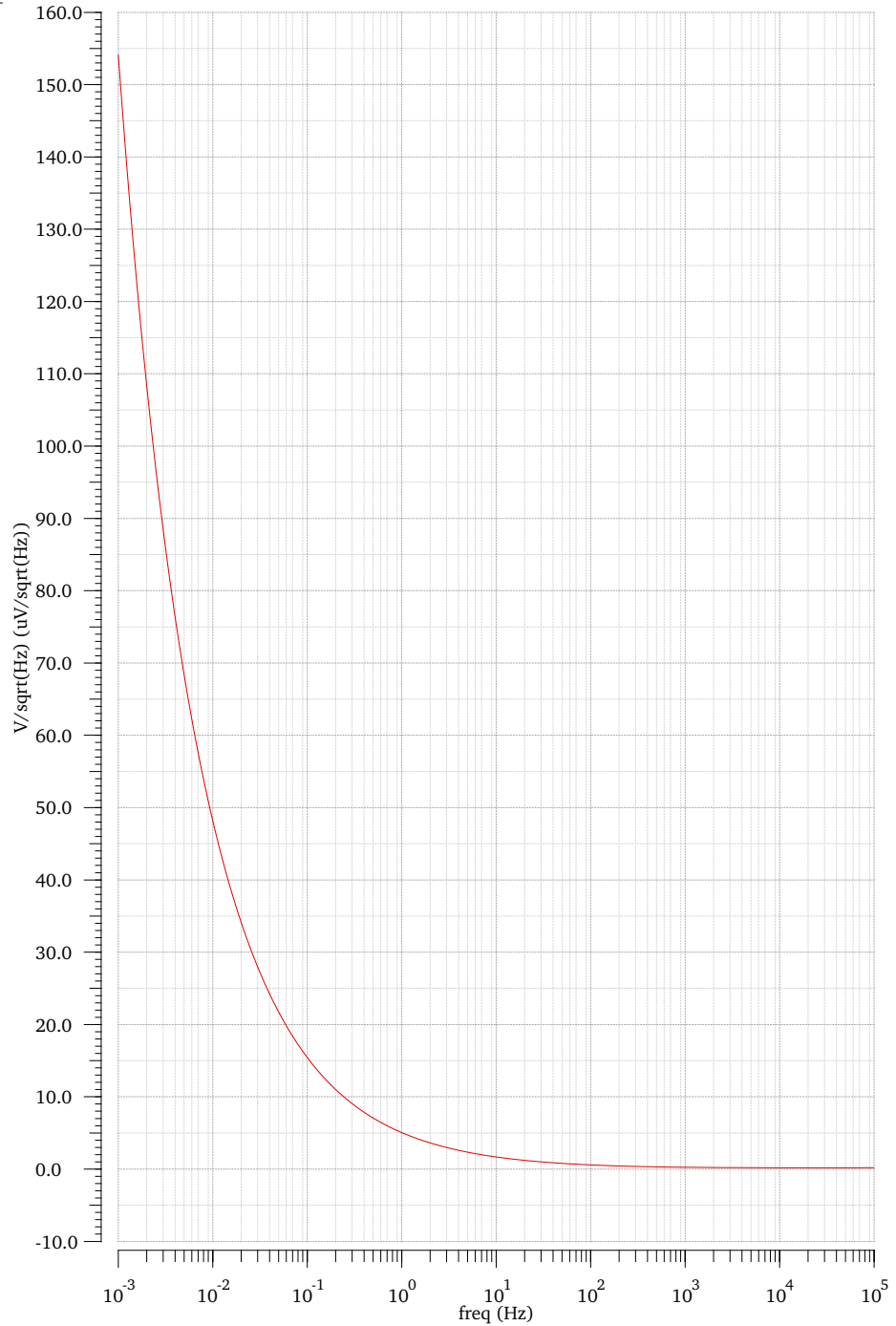
Noise Response

Figure 70: Original folded cascode op-amp input-referred noise response; TSMC 180 nm.

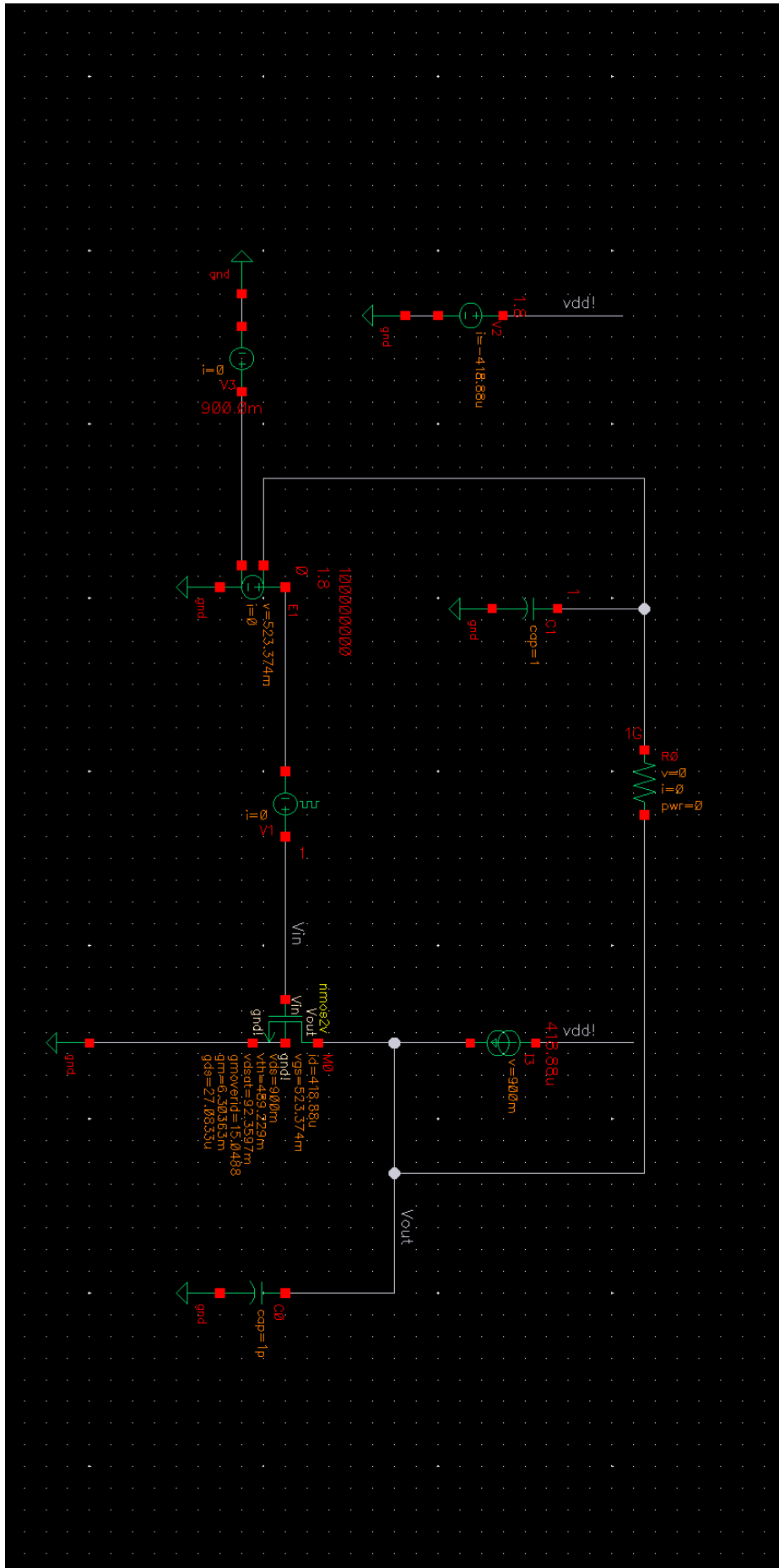## J IGS Cadence Screenshots (TSMC 180 nm)

See below.

Figure 71: IGS test bench schematic and DC operating points; TSMC 180 nm.

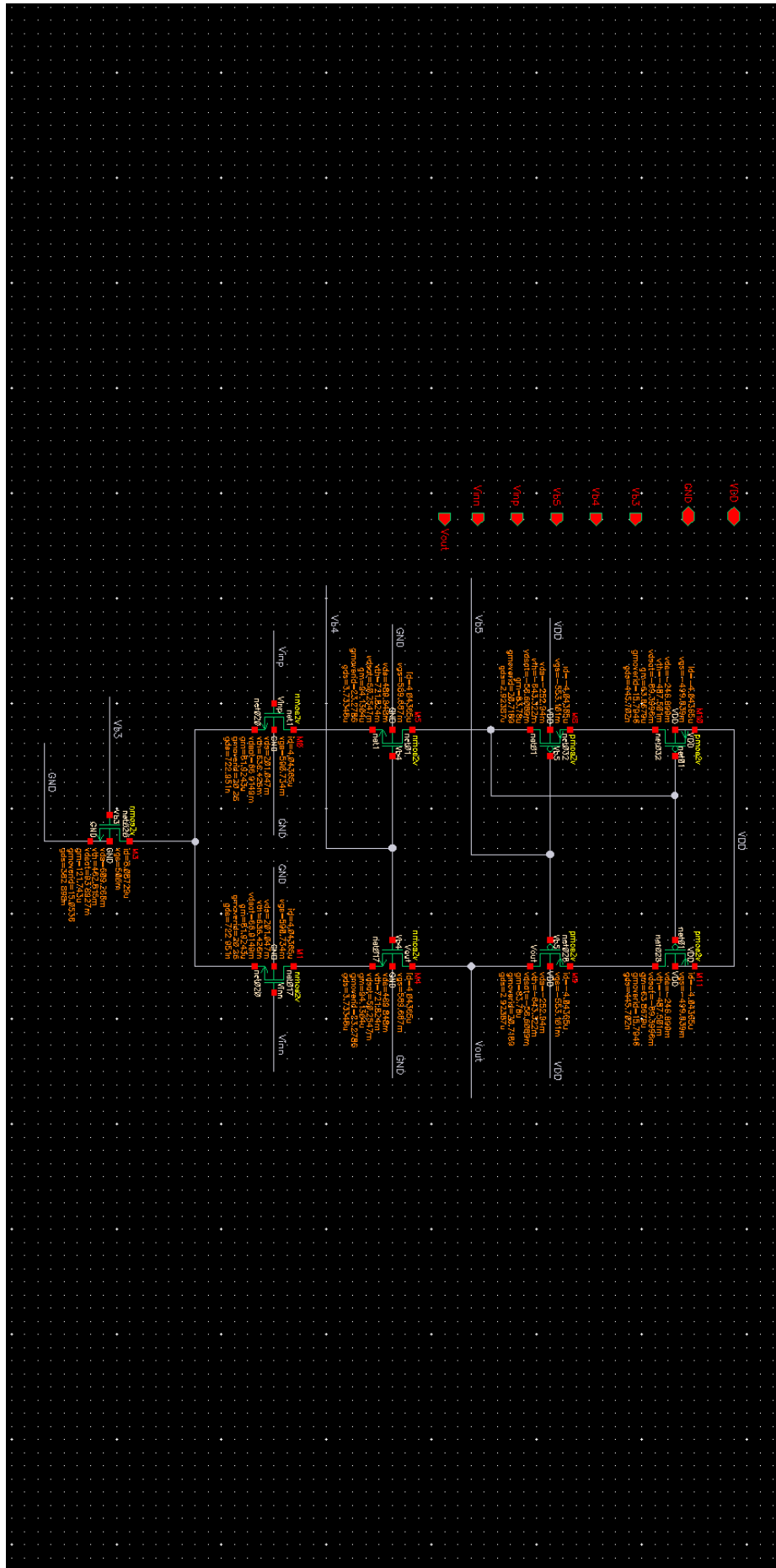# K Ahmadi and Jullien Topology Cadence Screenshots (TSMC 180 nm)

See below.

Figure 72: Potentiostat DC operating points; TSMC 180 nm.