
LAB 3: SHORT CHANNEL AMPLIFIER DESIGN

EE140/240A – Linear Integrated Circuits

Spring 2024

1. Objectives

In lab 2, you learned how to rely on hand-calculations to design an amplifier. This came along with successes and failures, but in the end, you realized that hand calculations give you a great starting place for your design, and you used Cadence to (possibly) sweep around these points. The key point is that our pre-Cadence work aided us in our work in Cadence.

Hand calculation is the preferred method with devices that are long channel. However, the shorter channel devices you have, the more you stray away from square-law models. If you've taken EE 130/230A, you might have seen some of these effects. Many second-order effects are not captured in devices in sub-micron devices, such as subthreshold conduction, drain-induced barrier lowering (DIBL), and velocity saturation. This lab, we will use short channel devices, and learn methods to come up with reasonable designs that do not rely on hand calculations.

The main question you might be asking now: if I can't use hand calculations, what do I use? Through your analysis of hand calculations, you found g_m to be quite a very important parameter. Higher g_m 's relate to better gain, for example. To answer this question, we look at research. In 1995, Flandre et al. devised a powerful design approach for bypassing any device model by relying on pre-computed/simulated transistor-width-independent parameters. You can read the paper here *g_m/I_D Methodology* (you can access this for free with your Berkeley email). It's quite a short read, and understanding this will give you much better appreciation for the method we are going to learn in this lab.

Transistor-width-independent parameters are the key idea driving our design. Particularly, the size-independent parameter known as g_m/I_D is a unique parameter that can be used as a knob to adjust the inversion level and ultimately the speed or energy efficiency of MOS transistors. Think about it this way: high g_m/I_D values (around 20-30) means that lots of g_m is created with little I_D . You can imagine this is good for things like lowering power consumption while increasing gain. This parameter is known as the transconductance efficiency.

Another important parameter is I_D/W , which I'll refer to as normalized drain current (or current density). This allows us to translate our size-independent parameters to the sizes of our devices. There are many other important size-independent parameters, such as g_m/c_{gg} , and g_m/g_{ds}

Ok, maybe I now can tell you the answer to the question you are asking. The essence of all development is using a **look-up tables** for the devices we are using. We can generate look-up tables by probing our gpdk045 devices with all possible $V_{GS}, V_{DS}, V_{BS}, L$ combinations, and measuring the g_m/I_D and other parameters. Now we have lots of data, and our job as analog designers is, when we receive a list of specifications, to generate **software scripts** that return to us the parameters that achieve the specifications. We use these scripts to interact with a 4-dimensional look-up table (LUT), filter out combinations that do not satisfy some constraint, and iterate on a selected value, then denormalize our terms to get them into their size-dependent values.

2. Timeline

This is a 4 week lab, split into 5 total exercises, where the last exercise, we will create a differential amplifier, known as the 5-T workhorse amplifier. I will be releasing the exercises on a weekly basis, and you are highly, highly encouraged to finish them before the next lab. (Note: It will be too easy to fall behind in this lab. My recommendation is to attempt the problem before coming to lab, finish the coding/start with cadence during the lab, and work on your own to finish the exercise). The current tentative plan is listed below; it is subject to change.

- Week 1: Exercises 1 and 2. Introductory exercises to measure gain, unity gain bandwidth, g_m/I_D , V_{GS}
- Week 2: Exercises 3 and 4. Creating scripts that take in gain, bandwidth, etc., to return transistor sizes. We'll take a look at our amplifier in Lab 1 and 2 and make significant improvements to it via this methodology.
- Week 3: Exercise 5. Creating our differential amplifier (5T Workhorse amplifier).
- Week 4: Buffer week. Writing the lab report, a time to catch up if you fell behind.

This lab is tentatively due on March 22nd at 11:59pm. This is planned to be due on the Friday before Spring Break, instead of a Tuesday like the previous two labs. Keep this in mind when we enter week 4, as the lab is due 3 days after.

3. Deliverables

In preparation for the final project, we will create more of a formal report, and the report itself will be much more opened ended than Lab 1 and 2. Your goal is to answer the questions given and prove them through Cadence. You should walk through your design process, as this is equally important to achieving simulation results, from mapping the constraints, to coding in Python/Matlab, to placing down a schematic in Cadence, to verifying the results in ADE Explorer.

I will be checking everyone's progress of the lab each week. This also serves as the attendance of the lab. All but Week 4 will have mandatory attendance. There will be a couple of questions I'll ask each week to make sure you are understanding what is happening, but I'll also have you ask me questions about different design related ideas with these exercises. These check-ins and check-offs will be part of your grade.

4. Grading

Lab 3 is worth 80 points. The rubric is loosely defined at the moment, unlike Lab 1 and 2, as the lab gets a lot more open ended. Here are some criteria I will follow while grading.

- **Correctness of coding.** There will be correct solutions to these some of the exercises. I'll grade these on correctness. There are also open-ended questions. I'll review your code for each exercise that are open-ended. *It'll be easier to grade for me if you provide documentation with your code.*
- **Getting and interpreting meaningful results.** Some questions in the exercises will be proposed as questions, to which you will answer. Some will be requiring you to maximize certain parameters, and you will be graded based on effort in achieving the maximization with some benchmark to help you along.
- **Explanation of your design process.** How did you think about coding up the solution to the design problem? You will all have similar explanations for the design process for the introductory questions, but Exercise 3, 4, and 5 are much more open ended and you must be able to explain why you made certain choices.
- **Style of your lab report.** Attention to details like well-formatted plots, explanations at every major part of the design. Please type your report, and if you need to add in equations or hand-calculations, you can add them in as an image.
- **Check-ins and Check-offs.** These graded on completion. Attending lab each week should guarantee these points. I'll write in each exercise what is needed for check-ins/offs, and what I'll be looking for each week.

5. Background

5.1 Design Flows of Long and Short Channel Devices

The design flow of any circuit starts with a given set of specifications and constraints (e.g., gain, bandwidth and output load). The designer first needs to identify the design variables (e.g., g_m , W/L , etc.) and find their relationships with the given specifications and constraints (e.g., $f_u = g_m/2\pi C_L$). The design space can usually be tightened by reducing the number of variables, that is, by elegant choice of topology (simplest topology with the fewest number of variables considered first) or by describing variables in terms of other variables. The last step in the design is to search within the design space to find the best set of variables that meet the design specs.

In Lab 2, we intentionally used longer channel devices to see can leverage the square-law model in our hand calculations to describe the relationships between the variables. For example, an intrinsic gain stage (IGS) has these relationships:

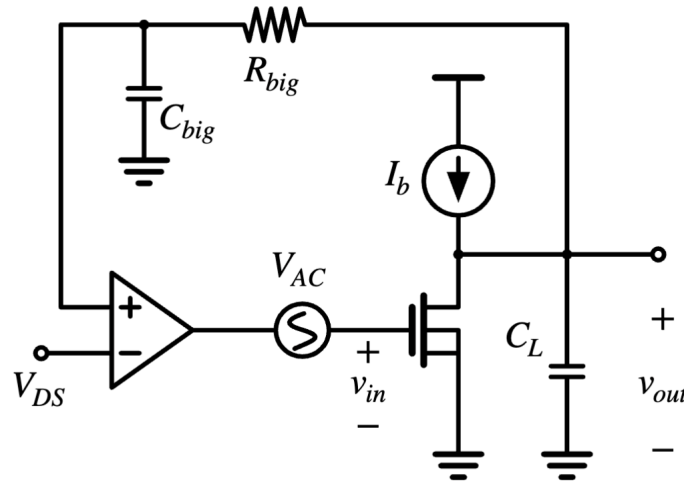


Figure 1: IGS

$$g_m = \sqrt{2\mu C_{ox} \frac{W}{L} I_D}$$

$$r_o = \frac{1}{\lambda I_D}$$

$$A_v = -g_m r_o = -\frac{1}{\lambda} \sqrt{\frac{2\mu C_{ox} W/L}{I_D}}$$

$$f_u = \frac{g_m}{2\pi C_L} = \frac{\sqrt{2\mu C_{ox} I_D W/L}}{2\pi C_L}$$

Therefore, I can give you an gain A_o , f_u , and C_L and you would simultaneously solve W/L and I_D . The design is essentially complete if our model was perfectly square law. Of course, in Lab 2, we made approximations in the square-law model, some of which included channel-length modulation, body-effect, parasitic capacitances, so our modeling wasn't entirely perfect. You might have hand-calculated all your values, but then when applied to our Level 3 MOSFET Model, the results did not meet the specifications.

The broken square-law model is replaced by pre-computed lookup tables. These lookup tables store all the information about our short-channel transistors (such as small-signal parameters g_m, c_{gs}, g_{ds}) across all possible bias points (various chosen V_{GS}, V_{DS}, V_{SB}) and for various device lengths. In short, any small-signal parameter of the device is stored as a function of four variables: V_{GS}, V_{DS}, V_{SB} , and L . These lookup tables are generated for each device only once and by using 4-dimensional parameteric dc simulations whose setup is shown below:

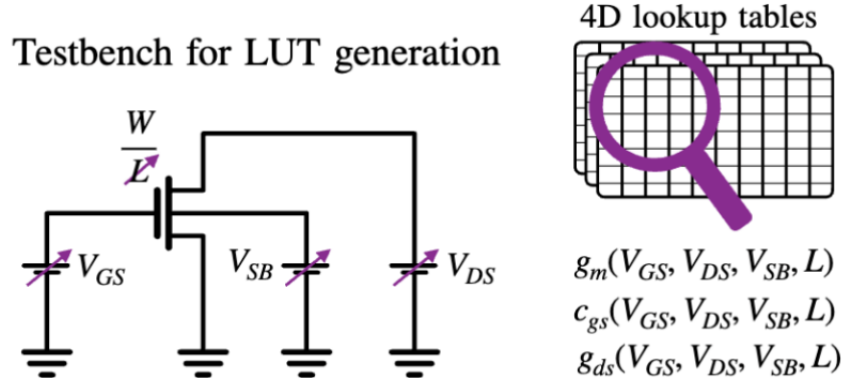


Figure 2: Testbench for LUT generation

As you have noticed, these look-up tables are generated using a single and arbitrary value for the width of the device W , say $1 \mu\text{m}$. Then, you may wonder how we can use these LUTs as these LUTs store small-signal parameters of a device with say $W = 1 \mu\text{m}$. What if my transistor has a different W than the one these look-up tables were generated for? Well, the answer lies in the power of g_m/I_D design method that relies on W -independent variables such as $g_m/I_D, I_D/W, g_m/c_{gg}, g_m/g_{ds}$. It can be shown that these variables are (to the first order) independent of the width of the transistor (curious? Use the square-law model to demonstrate it). Therefore, it does not matter what W is used for the test device in generating the LUTs as long as you use the LUTs for accessing W -independent variables.

5.2 Developing a Systematic Design Methodology

EE140/240A has a library that attempts to systematizes this lookup procedure. Essentially, there is some main lookup function that allows you to lookup a size-independent parameter based on arrays of V_{GS}, V_{DS}, V_{SB} , and L . This library exists for Python and Matlab, and it is inspired by Boris Murmann's scripts written in Matlab. This allows us to follow this design flow.

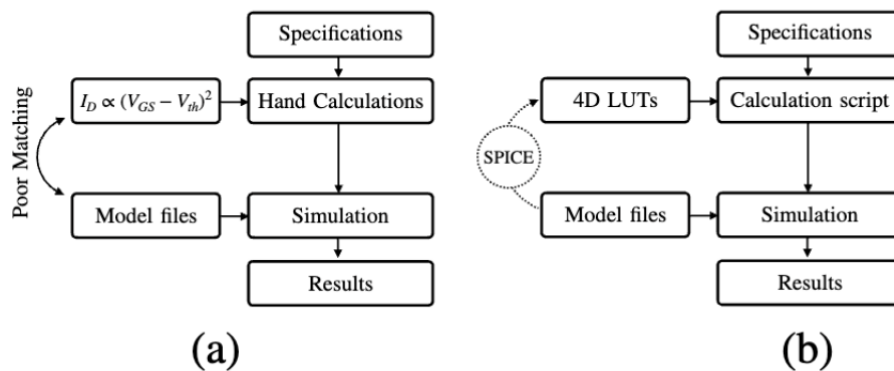


Figure 3: Design Flow

Week 1: Determining and Verifying Lookups

Exercise 1: Getting Used to Lookups

This exercise has a lot of information, but is meant for you to understand how the library works, and give you an understanding of how you can use it to answer design problems.

First, setup your workspace. You are provided the look-up tables for nmos and pmos devices using gpdk045, that is, nmos1v/pmos1v. These lookup tables can be accessed using Python or Matlab and you are free to run whichever you are more comfortable. The lab computers do not support python and can be slow when running Matlab. For that reason, we have provided the data at this link: <http://tinyurl.com/ee140-lab3-data>, so you may work with the data on your personal computer as well. If you choose to use Matlab on the lab computers, run these commands:

```
1 cd sim
2 mkdir gpdk045_LUTs
3 cd gpdk045_LUTs
4 cp -r /home/ff/ee140/fa22/lab2/LUTs/ .
5 export PATH=' $PATH:/share/b/bin '
6 matlab &
```

If you run these commands after sourcing your cadence setup, you need to add Matlab's path to your \$PATH environment variable again by entering:

```
1 export PATH=' $PATH:/share/b/bin '
2 matlab &
```

If you choose to develop on your personal device using Matlab, place nmos_1v.mat, pmos_1v.mat, look_upVGS.m and look_up.m in the same folder as your scripts. If you choose to develop with Python, place nmos_1v.mat, pmos_1v.mat, and lookup.py in the same folder as your code.

Inside the LUTs directory you copied earlier, in addition to the LUTs, there are two important functions look_up and look_upVGS written by Boris Murmann. Open these function scripts and carefully read their description. Now create a new matlab/python script and load the look-up tables of the nmos1v and pmos1v devices:

```
1 % MATLAB
2 nch = importdata('nch_1v.mat');
3 pch = importdata('pch_1v.mat');

1 % Python
2 import numpy as np
3 import matplotlib.pyplot as plot
4 from look_up import *
5 nch = importdata('nch_1v.mat')
6 pch = importdata('pch_1v.mat')
```

NOTE: If you copy code from this PDF to Matlab, the single quote symbol might be incompatible with Matlab's editor, so you have to type the single quote symbols manually.

Explore the data stored in `nch` and `pch`. For example, you can find the length and width of the `nmoslv` test device used for generating these LUTs by entering `nch.L` or `nch['L']` and `nch.W` or `nch['W']`, respectively for MATLAB or python. These dimensions are stored in micrometers. The sweep range for V_{GS} and V_{DS} are respectively obtained by `nch.VGS` and `nch.VDS`. Moreover, you can obtain I_D - V_{GS} curves of the test device (remember with W of `nch.W`) for various L values by the following commands:

```

1 % MATLAB
2 L = 0.1:0.2:1; % Lengths of interest
3 semilogy(nch.VGS, look_up(nch, 'ID', 'VGS', nch.VGS, 'L', L));

1 % Python
2 L = np.arange(0.1, 1, 0.2)
3 plt.semilogy(nch['VGS'], look_up_basic(nch, 'ID', vgs=nch['VGS'], l=L).T)
4 plt.show()

```

Notice we did not specify V_{DS} in the above plots, instead the script specified a default V_{DS} . What is the V_{DS} in the curves you just plotted.

Can you estimate the threshold voltage for the test devices? The threshold voltage is the V_{GS} when the drain current exceeds $1 \mu A$ for a $1 \mu m$ width device.

Similarly, you can obtain I_D , V_{DS} of the test device for various V_{GS} values using the following commands

```

1 % MATLAB
2 VGS = 0.1:0.3:max(nch.VGS);
3 plot(nch.VDS, look_up(nch, 'ID', 'VGS', VGS, 'VDS', nch.VDS))

1 % Python
2 VGS = np.arange(0.1, max(nch['VGS']), 0.3)
3 plt.plot(nch['VDS'], look_up_basic(nch, 'ID', vgs=VGS, vds=nch['VDS']).T)

```

What is the length of the device in the I_D , V_{DS} curves you just plotted? The I_D , V_{GS} and I_D , V_{DS} curves of test `nmoslv` are shown in the figure below for reference.

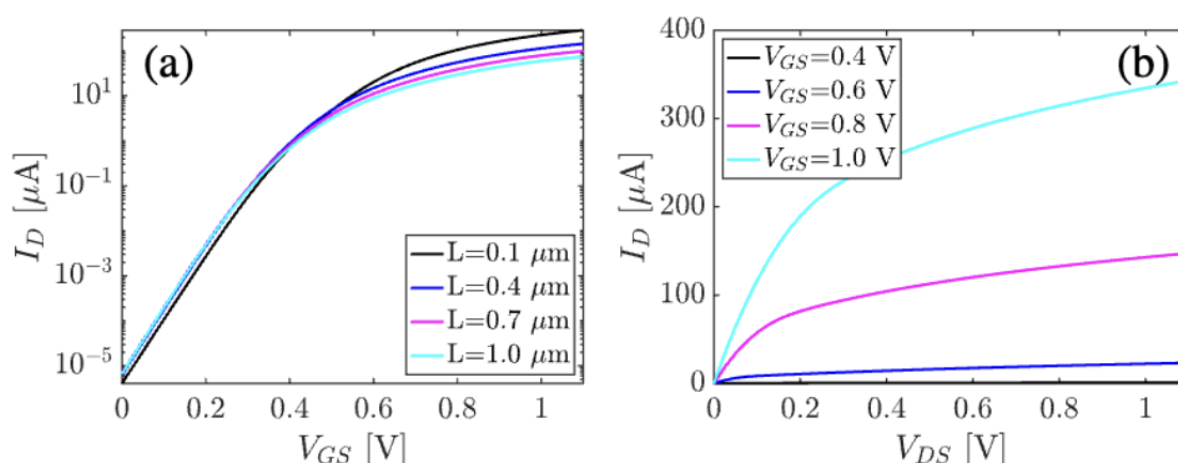


Figure 4: (a) $I_D - V_{GS}$ (b) $I_D - V_{DS}$

As mentioned earlier, almost never do we use LUTs to obtain absolute values of the small signal parameters. Instead, we obtain the ratios of small-signal parameters that are width-independent. You will see later on, how these width-independent ratios are effectively used during the design phase as variables. For now, let's plot three of these width-independent variables to uncover an inherent trade-off in MOS devices.

The transit frequency f_T of a transistor is the frequency at which the magnitude of the current gain of the device falls to unity. This current gain is calculated when the drain of the transistor is at AC ground and the input has a small signal current input i_{in} . Draw the small signal model of an NMOS and prove that

$$f_T = \frac{g_m}{2\pi c_{gg}}$$

$$c_{gg} = c_{gs} + c_{gd} + c_{gb}$$

The classic small signal model of the transistor you used to derive falls short of correctly modeling the transistor's second order effects at frequencies greater than roughly $f_T/10$. Therefore, we limit the operating frequency range of our circuits to be $\leq f_T/10$. Keep in mind that f_T is a measure of a device's speed. That is, devices with a larger f_T allow us to operate at higher frequencies. Both g_m and c_{gg} scale linearly with W , and therefore f_T is a width-independent variable. So are the intrinsic gain $A_v = g_m/g_{ds}$, transconductance efficiency g_m/I_D , and the current density $J_D = I_D/W$ of the MOS devices. In the g_m/I_D design methodology, as the name suggests, g_m/I_D is used as a knob to set the speed or the intrinsic gain of the devices. Let's have a demonstration. Plot f_T and $A_v = g_m/g_{ds}$ of an intrinsic gain stage as a function of g_m/I_D by running the following commands:

```

1  % MATLAB
2  L = [0.05 0.2 0.5 1];
3  gm_ID = 5:0.1:30
4  ft = look_up(nch, 'GM_CGG', 'GM_ID', gm_ID, 'L', L)/2/pi;
5  Av = look_up(nch, 'GM_GDS', 'GM_ID', gm_ID, 'L', L0);
6  plot(gm_ID, ft, '-')
7  hold on
8  yyaxis right
9  plot(gm_ID, AV, ':')

1  % Python
2  L = np.array([0.05, 0.2, 0.5, 1])
3  gm_ID = np.arange(4, 30, 0.1)
4  ft = look_up_vs_gm_id(nch, 'GM_CGG', gm_ID, l=L)/2/np.pi
5  Av = look_up_vs_gm_id(nch, 'GM_GDS', gm_ID, l=L)
6  fig, ax1 = plt.subplots()
7  ax2 = ax1.twinx()
8  ax1.semilogy(gm_ID, ft.T, '-')
9  ax2.plot(gm_ID, Av.T, ':')
10 plt.show()

```

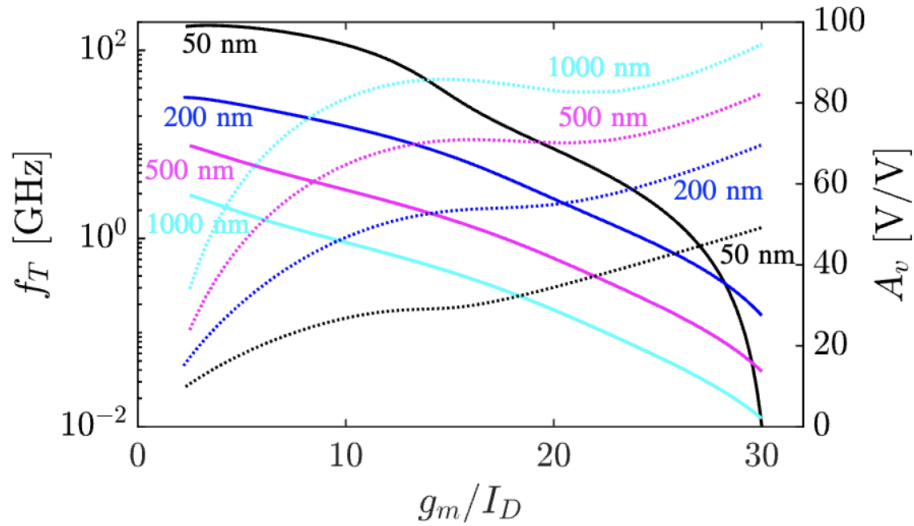


Figure 5: g_m/I_D used as a knob for adjusting the speed and the intrinsic gain of a device. Solid and dotted lines are respectively f_T and A_v .

Four main observations can be made in regards to Figure 5:

1. A_v is proportional to the length of the device. This is expected as increasing the length decreases the channel length modulation and ultimately g_{ds} .
2. f_T is inversely proportional to the length of the device. Again, this is an expected result because reducing the length simultaneously decreases c_{gg} and increases g_m .
3. A_v roughly monotonically increases with g_m/I_D , while at the same time
4. f_T degrades with increasing g_m/I_D .

Therefore, g_m/I_D and L can be used as proxies to adjust the gain and speed, but remember that either the speed or the gain of the device can be maximized, not both by these proxies.

Now we come up with a systematic design methodology:

1. Determine the design specifications.
2. Determine the constraints for the design specifications.
3. Determine g_m .
4. Pick L . Short-channel means high f_T and a small area. A longer channel will give more gain A_v and a larger area.
5. Pick g_m/I_D . A large value will give low power and high gain, a small value will give high f_T and small area.
6. Determine I_D from step 3 and 5.
7. Determine W using I_D/W denormalization.

Consider the intrinsic gain stage. Calculate the intrinsic gain, f_T , W and required V_{GS} of the nmos1v device with the following assumptions: $g_m/I_D = 15 \text{ S/A}$, $L = 100 \text{ nm}$, $C_L = 1 \text{ pF}$, $f_u = 1 \text{ GHz}$, and $V_{DS} = 0.55 \text{ V}$.

The feedback network shown in gray is used to properly set the gate-source voltage of the nmos such that the output DC value is at $V_{DS} = 0.55 \text{ V}$. The V_{DD} of the circuit is 1.1 V . For AC signals, you can ignore the feedback loop. You can use an ideal voltage-controlled voltage source, vcvs, from the analogLib library to implement the DC settling feedback amplifier. To avoid simulation convergence issues you can set the Maximum Output Voltage and the Minimum Output Voltage of the vcvs source to 1.1 V and 0 V , respectively. The Voltage gain of the vcvs should be ≥ 1000 .

Here are all the tips that can help you figure out this problem. These are in the order you should think about them in.

- $g_m/I_D, L, V_{DS}, V_{SB} = 0$ are known. Write them all down in code.
- Since these values are known, you can directly use the look_up function to obtain f_T and the intrinsic gain of the device. Don't forget the division by 2π .
- The intrinsic gain is $g_m r_o$. Recall that $r_o = \frac{1}{g_{ds}}$. How can you use a look_up to get the gain knowing this?
- g_m can be found from the choice of f_u and C_L .
- You'll need to perform a lookup to get current density $J_D = I_D/W$. Then you can get the width knowing that $W = I_D/J_D$. How can you get I_D from a term you did a lookup on that includes I_D ?
- Finally, V_{GS} can be performed using a lookup with the known values.

Now, simulate the performance of your IGS with the parameter values calculated above. You should run both DC and AC simulations to confirm the bias points of the amplifier as well as its frequency response. Save a copy of the bode plot of the transfer function of your amplifier and create a table listing its designed and simulated gain, unity gain bandwidth, g_m/I_D , and V_{GS} .

Exercise 2: Another IGS Example

We will now try another exercise with less guidance. The previous exercise was fully constrained with known g_m/I_D and L . There was no design involved, and we merely analyzed the given IGS. In this exercise, we relax the constraints by setting L as a design variable.

Consider the same intrinsic gain stage. Assuming a constant $g_m/I_D = 5V^{-1}$, sweep over $nch.L$, find the achievable f_T and A_v by the IGS as well as the corresponding W for the `nmos1v` device. As in Exercise 1: $C_L = 1\text{ pF}$, $V_{DD} = 1.1\text{ V}$ and $V_{DS} = 0.55\text{ V}$, but $f_u = 500\text{ MHz}$.

Plot $\log(f_T)$ vs. L and A_v vs. L . Find the maximum value of L that satisfies $f_u \leq f_T/10$. As you can see, there are many solutions $\{L, W, I_D\}$ that satisfy $f_u = 500\text{ MHz}$ constraint (with different intrinsic gains, of course). Choose two of these solutions and confirm with simulations both of them obtain $f_u = 500\text{ MHz}$. Save the simulated AC response of your designs and compare their performance in a table.