Design of the Folded Cascode OTA

Version 1

Christian Enz

2025-05-22

Table of contents

1 Introduction

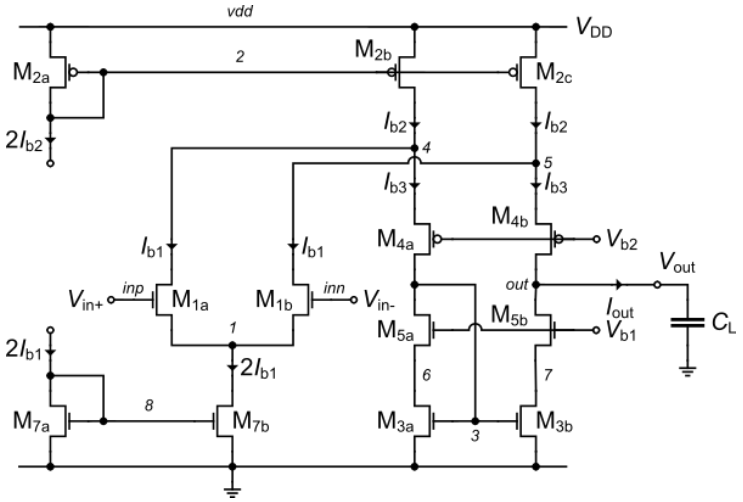

Figure 1.1: Schematic of the folded cascode differential OTA [1].

> **Note**
>
> Note that all nMOS transistors in Figure 1.1 have an odd number and all the pMOS transistors an even number which explains the numbering process (this is the reason why there is no M_6 transistor!).

This notebook presents the design of the folded cascode differential OTA [1] shown in Figure 1.1 for the same specifications used for the other OTAs. The folded cascode OTA is similar to the telescopic OTA except that the current I_{b3} flowing in the current mirror M_{3a}-M_{3b} can be different than the current I_{b1} flowing in the differential pair. How to choose the ratio $\alpha \triangleq I_{b3}/I_{b1}$? If the current I_{b3} is made equal to current I_{b1} ($\alpha = 1$), then if the differential input voltage is largely positive, the differential pair will saturate and all the bias current $2\,I_{b1}$ is then steered into M_{1a}. This means that all the bias current I_{b2} is flowing into M_{1a} and no current is available at the input of the cascode current mirror M_{3a}-M_{3b}. This situation should be avoided because it introduces an additional delay to charge the parasitic capacitance at node 3 and bring the current mirror M_{3a}-M_{3b} back to normal operation. Usually, the ratio α is taken $1.2 < \alpha < 2$. In order to minimize the current consumption, we will choose $\alpha = 1.2$. This means that the folded cascode OTA unavoidably consumes more current than the telescopic OTA (more than twice).

We will start with a detailed analysis of the OTA which will allow to derive all the design equations that will be used in the design phase. The OTA is then designed for a given set of specifications for the chosen IHP SG13G2 130nm BiCMOS technology [2]. We have selected this technology because IHP provides an open source PDK which is then used for the validation of the design by simulation with ngspice [3] using the PSP compact model [4] provided by the open source PDK [2].

2 Analysis

2.1 Small-signal analysis

In a 1st-order analysis, we can neglect the capacitances at the low impedance cascode nodes 4, 5, 6 and 7 and only account for the capacitances at high impedance nodes 3 (gate of current mirror M_{3a}-M_{3b}) and out (output node). The circuit becomes similar to that of the simple OTA, with the dominant pole ω_0 at the output node (out) and the non-dominant pole ω_p at the current mirror node 3. The transfer function also has a pole-zero doublet. Its transfer function is then given by

$$A_{dm}(s) \triangleq \frac{\Delta V_{out}}{\Delta V_{in}} \cong A_{dc} \cdot \frac{1 + s/\omega_z}{(1 + s/\omega_0)(1 + s/\omega_p)} \cong \frac{A_{dc}}{1 + s/\omega_0} \cong \frac{\omega_u}{s}, \tag{2.1}$$

where

$$A_{dc} \cong \frac{G_{m1}}{G_o}, \tag{2.2}$$

$$G_o \cong \frac{G_{ds3}\, G_{ds5}}{G_{ms5}} + \frac{(G_{ds1} + G_{ds2})\, G_{ds4}}{G_{ms4}}, \tag{2.3}$$

$$\omega_0 \cong \frac{G_o}{C_L}, \tag{2.4}$$

$$\omega_p \cong \frac{G_{m3}}{C_3}, \tag{2.5}$$

$$\omega_z = 2\,\omega_p, \tag{2.6}$$

$$\omega_u = A_{dc} \cdot \omega_0 \cong \frac{G_{m1}}{C_L}. \tag{2.7}$$

2.2 Noise Analysis

At low-frequency the noise of the cascode transistors M_{4a}-M_{4b} and M_{5a}-M_{5b} can be neglected and the noise analysis is then identical to that of the simple OTA. The PSD of the output noise current is given by

$$S_{nout} \cong 2\,(S_{I_{n1}} + S_{I_{n2}} + S_{I_{n3}}) \tag{2.8}$$

which can be expressed in terms of the output noise conductance

$$S_{nout} = 4kT \cdot G_{nout}, \tag{2.9}$$

where

$$G_{nout} \cong 2\,(G_{n1} + G_{n2}G_{n3}), \tag{2.10}$$

with

$$G_{ni} = \gamma_{ni} \cdot G_{mi} + G_{mi}^2 \cdot \frac{\rho_i}{W_i\, L_i\, f} \quad \text{for } i = 1, 2, 3. \tag{2.11}$$

The input-referred noise resistance is then given by

$$R_{nin} \triangleq \frac{G_{nout}}{G_{m1}^2} = \frac{2\,(G_{n1} + G_{n2} + G_{n3})}{G_{m1}^2} = \frac{2\,G_{n1}}{G_{m1}^2} \cdot \left(1 + \frac{G_{n2}}{G_{n1}} + \frac{G_{n3}}{G_{n1}}\right) \tag{2.12}$$

which we rewrite as

$$R_{nin} = \frac{2\,G_{n1}}{G_{m1}^2} \cdot (1+\eta) \tag{2.13}$$

with

$$\eta = \frac{G_{n2}}{G_{n1}} + \frac{G_{n3}}{G_{n1}}. \tag{2.14}$$

η represents the contribution of the current source M_{2b}-M_{2c} and the current mirror M_{3a}-M_{3b} referred to the input and normalized to the contribution of the differential pair M_{1a}-M_{1b}. Of course during the design phase we will try to minimize η.

2.2.1 Input-referred thermal noise

The input-referred thermal noise resistance is given by

$$R_{nth} = \frac{2\,\gamma_{n1}}{G_{m1}} \cdot (1+\eta_{th}), \tag{2.15}$$

where

$$\eta_{th} = \frac{\gamma_{n2}}{\gamma_{n1}} \cdot \frac{G_{m2}}{G_{m1}} + \frac{\gamma_{n3}}{\gamma_{n1}} \cdot \frac{G_{m3}}{G_{m1}} \tag{2.16}$$

represents the contribution to the input-referred thermal noise of the current source M_{2b}-M_{2c} and the current mirror M_{3a}-M_{3b} relative to that of the differential pair M_{1a}-M_{1b}. In case $G_{m1} \gg G_{m2}$ and $G_{m1} \gg G_{m3}$, then $\eta_{th} \ll 1$ and the thermal noise is dominated by the input differential pair. Eqn. (2.15) can then be simplified to

$$R_{nth} \cong \frac{2\,\gamma_{n1}}{G_{m1}}. \tag{2.17}$$

The OTA thermal noise excess factor is defined as

$$\gamma_{ota} \triangleq G_m \cdot R_{nth} \tag{2.18}$$

with $G_m = G_{m1}$ the OTA equivalent transconductance. The OTA thermal noise excess factor then writes

$$\gamma_{ota} = 2\,\gamma_{n1} \cdot (1+\eta_{th}). \tag{2.19}$$

In the case $G_{m1} \gg G_{m2}$ and $G_{m1} \gg G_{m3}$, then $\eta_{th} \ll 1$ and the noise is dominated by the input differential pair M_{1a}-M_{1b}. The OTA thermal noise excess factor can then be simplified as

$$\gamma_{ota} \cong 2\,\gamma_{n1}. \tag{2.20}$$

2.2.2 Input-referred flicker noise

The input-referred flicker noise is given by

$$R_{nfl} = \frac{2}{f}\left[\frac{\rho_n}{W_1\,L_1} + \left(\frac{G_{m2}}{G_{m1}}\right)^2 \frac{\rho_p}{W_2\,L_2} + \left(\frac{G_{m3}}{G_{m1}}\right)^2 \frac{\rho_n}{W_3\,L_3}\right] \tag{2.21}$$

which can be rewritten as

$$R_{nfl} = \frac{2\rho_n}{W_1\,L_1\,f} \cdot (1+\eta_{fl}) \tag{2.22}$$

where

$$\eta_{fl} = \frac{\rho_p}{\rho_n}\left(\frac{G_{m2}}{G_{m1}}\right)^2 \frac{W_1\,L_1}{W_2\,L_2} + \left(\frac{G_{m3}}{G_{m1}}\right)^2 \frac{W_1\,L_1}{W_3\,L_3} \tag{2.23}$$

represents the contribution to the input-referred flicker noise of the current source M_{2b}-M_{2c} and the current mirror M_{3a}-M_{3b} relative to that of the differential pair M_{1a}-M_{1b}.

The corner frequency is the frequency at which the flicker noise becomes equal to the thermal noise

$$R_{nfl}(f = f_k) = R_{nth} \tag{2.24}$$

which is given by

$$f_k = \frac{1}{R_{nth}} \cdot \frac{2\rho_n}{W_1\, L_1} \cdot (1 + \eta_{fl}) = \frac{G_{m1}\, \rho_n}{\gamma_{n1}\, W_1\, L_1} \cdot \frac{1 + \eta_{fl}}{1 + \eta_{th}}. \tag{2.25}$$

The transconductance G_{m1} is set by the constraints either on thermal noise or on bandwidth (GBW product). The corner frequency f_k can be reduced by increasing $W_1\, L_1$ but also $W_2\, L_2$ and $W_3\, L_3$ at the same time to conserve the same η_{fl} factor. Assuming that $G_{m2}/G_{m1} \ll 1$ and $G_{m1} \gg G_{m3}$, as required by the constraints on minimizing the contribution of the current source M_{2b}-M_{2c} and current mirror M_{3a}-M_{3b} to the input-referred offset and thermal noise, then $\eta_{th} \ll 1$ and $\eta_{fl} \ll 1$ and the corner frequency f_k is then mainly set by the differential pair transconductance and gate transistor area

$$f_k \cong \frac{G_{m1}\, \rho_n}{\gamma_{n1}\, W_1\, L_1}. \tag{2.26}$$

2.3 Input-referred offset voltage

The offset analysis is similar to that of the simple OTA because the contribution of the mismatch of the cascode transistors can be neglected. The random offset current is then mainly due to the mismatch between M_{1a} and M_{1b}, M_{2b} and M_{2c} and M_{3a} and M_{3b}. The variance of the output offset current is then given by

$$\sigma^2_{I_{os}} \cong \sigma^2_{\Delta I_{D1}} + \sigma^2_{\Delta I_{D2}} + \sigma^2_{\Delta I_{D3}} = I^2_{b1} \cdot \sigma^2_{\frac{\Delta I_{D1}}{I_{D1}}} + I^2_{b2} \cdot \sigma^2_{\frac{\Delta I_{D2}}{I_{D2}}} + I^2_{b3} \cdot \sigma^2_{\frac{\Delta I_{D3}}{I_{D3}}}, \tag{2.27}$$

with

$$\sigma^2_{\frac{\Delta I_{Di}}{I_{Di}}} = \sigma^2_{\beta_i} + \left(\frac{G_{mi}}{I_b}\right)^2 \sigma^2_{V_{Ti}} \quad \text{for } i = 1, 2, 3, \tag{2.28}$$

where

$$\sigma^2_{\beta_i} = \frac{A^2_\beta}{W_i\, L_i} \quad \text{for } i = 1, 2, 3 \tag{2.29}$$

are the β-mismatches and

$$\sigma^2_{V_{Ti}} = \frac{A^2_{V_T}}{W_i\, L_i} \quad \text{for } i = 1, 2, 3 \tag{2.30}$$

are the V_T-mismatches.

The variance of the output offset current then becomes

$$\sigma^2_{I_{os}} = I^2_{b1} \cdot \sigma^2_{\beta_1} + I^2_{b2} \cdot \sigma^2_{\beta_2} + I^2_{b3} \cdot \sigma^2_{\beta_3} + G^2_{m1} \cdot \sigma^2_{V_{T1}} + G^2_{m2} \cdot \sigma^2_{V_{T2}} + G^2_{m3} \cdot \sigma^2_{V_{T3}}. \tag{2.31}$$

The variance of the input-referred offset voltage is obtained by dividing the variance of the output offset current by G^2_{m1} resulting in

$$\sigma^2_{V_{os}} = \sigma^2_{V_T} + \sigma^2_\beta. \tag{2.32}$$

$\sigma^2_{V_T}$ is the V_T-mismatch given by

$$\sigma^2_{V_T} = \sigma^2_{V_{T1}} \cdot (1 + \xi_{V_T}) \tag{2.33}$$

where

$$\xi_{V_T} = \left(\frac{G_{m2}}{G_{m1}}\right)^2 \cdot \frac{\sigma^2_{V_{T2}}}{\sigma^2_{V_{T1}}} + \left(\frac{G_{m3}}{G_{m1}}\right)^2 \cdot \frac{\sigma^2_{V_{T3}}}{\sigma^2_{V_{T1}}} \tag{2.34}$$

represents the V_T-mismatch contributions to the input-referred offset of the current source M_{2b}-M_{2c} and current mirror M_{3a}-M_{3b} relative to that of the differential pair with

$$\sigma_{V_{T1}} = \frac{A^2_{V_{Tn}}}{W_1\, L_1}, \tag{2.35}$$

$$\sigma_{V_{T2}} = \frac{A^2_{V_{Tp}}}{W_2\, L_2}, \tag{2.36}$$

$$\sigma_{V_{T3}} = \frac{A^2_{V_{Tn}}}{W_3\, L_3}, \tag{2.37}$$

σ^2_β is the β-mismatch given by

$$\sigma^2_\beta = \left(\frac{I_{b1}}{G_{m1}}\right)^2 \cdot \sigma^2_{\beta_1} \cdot (1 + \xi_\beta) \tag{2.38}$$

where

$$\xi_\beta = \left(\frac{I_{b2}}{I_{b1}}\right)^2 \frac{\sigma^2_{\beta_2}}{\sigma^2_{\beta_1}} + \left(\frac{I_{b3}}{I_{b1}}\right)^2 \frac{\sigma^2_{\beta_3}}{\sigma^2_{\beta_1}} \tag{2.39}$$

represents the β-mismatch contributions to the input-referred offset of the current source M_{2b}-M_{2c} and current mirror M_{3a}-M_{3b} relative to that of the differential pair with

$$\sigma^2_{\beta_1} = \frac{A^2_{\beta_n}}{W_1\, L_1}, \tag{2.40}$$

$$\sigma^2_{\beta_2} = \frac{A^2_{\beta_p}}{W_2\, L_2}, \tag{2.41}$$

$$\sigma^2_{\beta_3} = \frac{A^2_{\beta_n}}{W_3\, L_3}. \tag{2.42}$$

Replacing in (2.34) and (2.39) results in

$$\xi_{V_T} = \left(\frac{G_{m2}}{G_{m1}}\right)^2 \left(\frac{A_{V_{Tp}}}{A_{V_{Tp}}}\right)^2 \frac{W_1\, L_1}{W_2\, L_2} + \left(\frac{G_{m3}}{G_{m1}}\right)^2 \frac{W_1\, L_1}{W_3\, L_3} \tag{2.43}$$

and

$$\xi_\beta = \left(\frac{I_{b2}}{I_{b1}}\right)^2 \left(\frac{A_{\beta_p}}{A_{\beta_n}}\right)^2 \frac{W_1\, L_1}{W_2\, L_2} + \left(\frac{I_{b3}}{I_{b1}}\right)^2 \frac{W_1\, L_1}{W_2\, L_2}. \tag{2.44}$$

Similarly to the flicker noise, the input-referred offset (variance or standard deviation) can be reduced by increasing the area $W_1\, L_1$ of M_{1a}-M_{1b} but at the same time also increasing the area $W_2\, L_2$ of the current source M_{2b}-M_{2c} and also the area $W_3\, L_3$ of the current mirror M_{3a}-M_{3b}.

3 Design

3.1 Specifications

The OTA specifications are given in Table 3.1.

> **Warning**
>
> An important limitation of the IHP SC13G2 BiCMOS technology [2] is the high output conductance and hence limited intrinsic gain of nMOS transistors. Because of this, the specification on the DC gain, that was initially set to 80 dB, has been downgraded to 70 dB.

> **Note**
>
> The specifications given in Table 3.1 are simplified specifications. They are mainly targeting the achievement of a certain gain-bandwidth product GBW and DC gain at lowest current consumption. The GBW sets the differential pair transconductance while the DC gain sets the output conductance. There is an additional specification on the random input-referred offset voltage which, if not met, might eventually require to increase the transistors area. There are no specifications on thermal noise since the transconductance is set by the GBW. There are no specifications on the flicker noise but if the corner frequency was set lower this would required to increase the transistors area. There are also no specifications on the slew-rate, which might be small because of the low-power objective. Finally, there are many more specifications such as CMRR, PSRR, input common-mode voltage range, output-voltage swing, etc... that are not discussed in this example.

Table 3.1: OTA specifications.

Specification	Symbol	Value	Unit
Minimum DC gain	A_{dc}	70	dB
Minimum gain-bandwidth product	GBW	1	MHz
Load capacitance	C_L	1	pF
Maximum input-referred random offset voltage	V_{os}	10	mV
Phase margin	PM	60	$^\circ$

3.2 Process

We will design the cascode gain stage for the open source IHP 13G2 BiCMOS process [2]. The physical parameters are given in Table 3.2, the global process parameters in Table 3.3 and finally the MOSFET parameters in Table 3.4.

> **Warning**
>
> The matching parameters for IHP 130nm are unknown. We will use those from a generic 180nm technology.

Table 3.2: Physical parameters

Parameter	Value	Unit
T	300	K
U_T	25.865	mV

Table 3.3: Process parameters.

Parameter	Value	Unit	Comment
t_{ox}	2.24	nm	SiO_2 oxyde thickness
C_{ox}	15.413	$\frac{fF}{\mu m^2}$	Oxyde capacitance per unit area
V_{DD}	1.2	V	Nominal supply voltage
L_{min}	130	nm	Minimum drawn gate length
W_{min}	150	nm	Minimum drawn gate width
z_1	340	nm	Minimum outer diffusion width
z_2	389	nm	Minimum diffusion width between two fingers

Table 3.4: Transistors parameters.

Parameter	nMOS	pMOS	Unit
Length and width correction parameters for current			
DL	59	51	nm
DW	-20	30	nm
Length and width correction for intrinsic and overlap capacitances			
$DLCV$	93	146	nm
$DWCV$	-10	15	nm
Length and width correction parameter for fringing capacitances			
$DLGCV$	34	96	nm
$DWGCV$	10	-15	nm
Long-channel sEKV parameters parameters			
n	1.22	1.23	-
$I_{spec\square}$	708	245	nA
V_{T0}	246	365	mV
Short-channel sEKV parameters parameters			
L_{sat}	7.1	24.9	nm
λ	1.375	6.078	$\frac{V}{\mu m}$
Junction capacitances parameters			
C_J	0.976	0.863	$\frac{fF}{\mu m^2}$
C_{JSWSTI}	0.025	0.032	$\frac{fF}{\mu m}$
C_{JSWGAT}	0.03	0.027	$\frac{fF}{\mu m}$
Overlap capacitances parameters			
C_{GSo}	0.453	0.443	$\frac{fF}{\mu m}$
C_{GDo}	0.453	0.443	$\frac{fF}{\mu m}$
C_{GBo}	0	0.022	$\frac{fF}{\mu m}$

Table 3.4: Transistors parameters.

Parameter	nMOS	pMOS	Unit
Fringing capacitances parameters			
C_{GSf}	0.2	0.1	$\frac{fF}{\mu m}$
C_{GDf}	0.2	0.1	$\frac{fF}{\mu m}$
Flicker noise parameters			
K_F	2.208e-24	1.2e-23	VAs
AF	1	1	-
ρ	0.008642	0.04697	$\frac{Vm^2}{As}$
Matching parameters			
A_{VT}	5	5	$mV \cdot \mu m$
A_β	1	1	$\% \cdot \mu m$

3.3 Design procedure

> **!** Important
>
> For this process, the transistor dimensions are rounded to 10nm. We also will ignore the length and width reduction parameters DL and DW. The main reason is that most of the transistor length and width are sufficiently large that ignoring these parameters has little impact.

We start sizing the differential pair M_{1a}-M_{1b}.

3.3.1 Sizing the differential pair M_{1a}-M_{1b}

In this example there is no specification on the thermal noise. Therefore the transconductance G_{m1} is set by the gain-bandwidth product according to

$$GBW = \frac{G_{m1}}{2\pi C_{out}}, \tag{3.1}$$

where G_{m1} is the gate transconductance of M_{1a} and M_{1b} and C_{out} the total output capacitance

$$C_{out} = C_o + C_L \tag{3.2}$$

with C_o the parasitic capacitance at the output node and C_L the load capacitance.

In order to minimize the input-referred noise and offset, the input differential pair should be biased in weak inversion. The transconductance G_{m1} in deep weak inversion is then given by

$$G_{m1} = \frac{I_{b1}}{nU_T}. \tag{3.3}$$

The bias current I_{b1} is the current flowing in each transistor M_{1a} and M_{1b} when the input differential voltage is zero. The bias current provided by M_{5b} is therefore $2I_{b1}$. The bias current must satisfy the following inequality:

$$I_{b1} \geq 2\pi n_{0n} U_T C_{out} GBW_{min}. \tag{3.4}$$

which for the given specifications gives $I_{b1,min} = 198\ nA$. The corresponding slew-rate is then equal to $SR_{min} = 198\ mV/\mu s$ which we will consider as sufficient.

> **!** Important
>
> If the slew-rate is not sufficient, the bias current I_b should be increased resulting in a higher current and power consumption. Other options include the use of a class AB OTA [5] or a dynamic/adaptive biasing OTA [6].

To have some margin to account for the additional parasitic capacitance at the output C_o due mostly to the junction capacitances that add to the load capacitance C_L, we set $I_{b1} = 250\ nA$ and the inversion coefficient to $IC_1 = 0.1$. The transconductance can be calculated from the G_m/I_D function as $G_{m1} = 7.258\ \mu A/V$. This leads to a gain-bandwidth product $GBW = 1.2\ MHz$, which is slightly higher than the target specification offering some margin. Knowing the drain current I_{D1} and the inversion coefficient, we can calculate the specific current $I_{spec1} = 2.500\ \mu A$ and the aspect ration $W_1/L_1 = 3.5$.

The degree of freedom left (W_1 or L_1) can be determined by constraints either on the dc gain, the offset voltage or the flicker noise.

Before finalizing the sizing of the differential pair, we first will size the current mirror M_{3a}-M_{3b} and the current sources M_{2a}-M_{2b}-M_{2c}.

3.4 Sizing of M_{3a}-M_{3b}

The current mirror M_{3a}-M_{3b} should be biased as much in strong inversion as the voltage constraint allows for. We can set its V_{GS} voltage to the desired quiescent output voltage which we set to $V_{GS3} = V_{outq} = V_{DD}/2 = 600\ mV$, which corresponds to an inversion coefficient $IC_3 = 28.080$ and a saturation voltage $V_{DSsat3} = 293\ mV$. We see that choosing a large IC and hence a large V_{GS} voltage comes at the cost of a large saturation voltage and hence a reduced output voltage swing. As discussed in the introduction we have chosen $\alpha = 1.200$ so that $I_{b3} = 300\ nA$ and $I_{b2} = I_{b1} + I_{b3} = 550\ nA$. Having set the inversion coefficient and bias current we can derive the specific current $I_{spec3} = 10.68\ nA$ and the aspect ration $W_3/L_3 = 0.015$. Having IC and I_D, we can deduce the transconductance $G_{m3} = 1.633\ \mu A/V$.

We will now have to make sure that the non-dominant pole f_p at node 3 is sufficiently higher than the gain-bandwidth product GBW to insure the desired phase margin. The non-dominant pole is given

$$\omega_p = \frac{G_{m3}}{C_3}, \tag{3.5}$$

where C_3 is given by

$$C_3 = 2(C_{GS3} + C_{GB3}) \tag{3.6}$$

Assuming that M_{3a}-M_{3b} are in saturation, the gate-to-source capacitance C_{GS3} is given by

$$C_{GS3} \cong W_3\, L_3\, C_{ox} \cdot c_{gsi} + (C_{GSon} + C_{GSfn}) \cdot W_3, \tag{3.7}$$

where c_{gsi} is the intrinsic gate-to-source capacitance normalized to the total gate area $W\, L\, C_{ox}$, which is typically equal to 2/3 in strong inversion and is proportionnal to IC in weak inversion.

The gate-to-bulk capacitance C_{GB3} is given by

$$C_{GB3} \cong\ W_3\, L_3\, C_{ox} \cdot c_{gbi} + C_{GBon} \cdot W_3, \tag{3.8}$$

where c_{gbi} is the gate-to-bulk intrinsic capacitance normalized to the total gate area $W\, L\, C_{ox}$ and given by

$$c_{gbi} = \frac{n-1}{n} \cdot c_{gsi}. \tag{3.9}$$

The capacitance at node 3 then scales with W_3 and L_3 according to

$$C_3 = W_3\,L_3 \cdot C_{WL} + W_3 \cdot C_W, \tag{3.10}$$

with

$$C_{WL} = 2\,C_{ox} \cdot (c_{gsi} + c_{gbi}), \tag{3.11}$$
$$C_W = 2(C_{GSon} + C_{GSfn} + G_{GBon}). \tag{3.12}$$

We already have set the inversion coefficient $IC_3 = 28.080$, from which we can calculate $c_{gsi} = 0.600$ and $c_{gbi} = 0.072$. Since the W/L has already been set by the transconductance and the current, we can derive W_3 and L_3 for achieving a given capacitance C_3 according to

$$W_3 = \frac{-C_W \cdot W_3/L_3 + \sqrt{W_3/L_3} \cdot \sqrt{4\,C_3\,C_{WL} + C_W^2 \cdot W_3/L_3}}{2\,C_{WL}}, \tag{3.13}$$

$$L_3 = \frac{W_3}{W_3/L_3}. \tag{3.14}$$

Setting the non-dominant pole f_p to 10 times the GBW, we get $C_3 = 26$ fF, $W_3 = 140\ nm$ and $L_3 = 9.28\ \mu m$.

We see that W_3 is smaller than the minimum width. If we don't want to increase $W_3\,L_3$ and hence C_3, we need to reduce the inversion coefficient IC_3. We can find the IC such that $W_3 = W_{min}$ for the given f_p by looking at Figure 3.1.

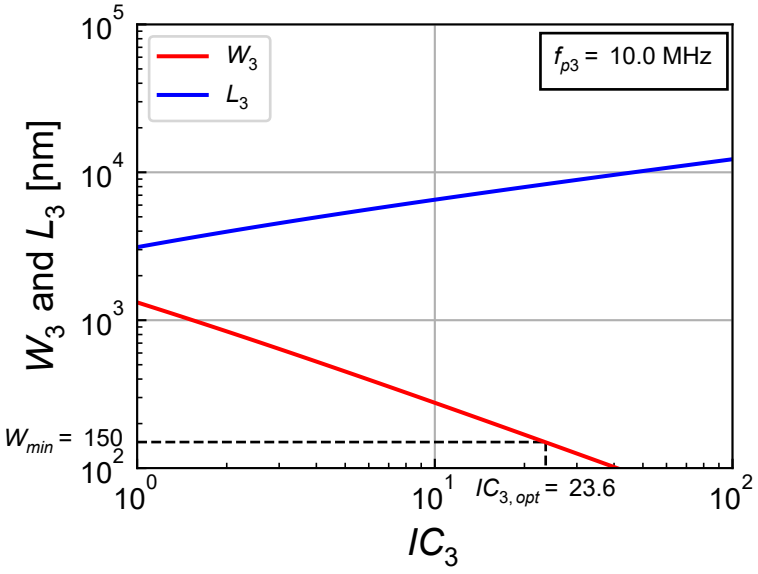

Figure 3.1: Length L_3 and width W_3 of M_3 versus IC_3 for a given f_{p3}/GBW ratio.

We get $IC_3 = 23.6$, $W_3 = 150\ nm$, $L_3 = 8.35\ \mu m$, $C_3 = 14.3\ fF$, $G_{m3} = 1.767\ \mu A/V$ and $I_{spec3} = 12.73\ nA$. This leads to leading to the desired ratio $f_p/GBW = 10.0$.

The inversion coefficient of M_{3a}-M_{3b} is therefore reduced from 28.1 to 23.6 in order to maintain a reasonable parasitic capacitance at node 3 to make sure that f_p is 10 times higher than the GBW.

Note that the V_{GS3} voltage has also dropped to $V_{GS3} = 569\ mV$, so that the quiescent output voltage is now lower than $V_{outq} = V_{DD}/2 = 600\ mV$. We might need to introduce an offset voltage to shift the output back to V_{outq} (in the high gain region). On the other hand the saturation voltage is now also lower $V_{DSsat3} = 272\ mV$ improving the output swing.

3.5 Sizing of M_{4a}-M_{4b} and M_{5a}-M_{5b}

The cascode transistors are sized according to the desired DC gain given by

$$A_{dc} = \frac{G_{m1}}{G_o} \tag{3.15}$$

where G_o is the conductance at the output node given by

$$G_o \cong G_{o35} + G_{o124} \tag{3.16}$$

where

$$G_{o35} \triangleq \frac{G_{ds3}\, G_{ds5}}{G_{ms5}}, \tag{3.17}$$

$$G_{o124} \triangleq \frac{(G_{ds1} + G_{ds2})\, G_{ds4}}{G_{ms4}}. \tag{3.18}$$

The minimum DC gain specification is given by $A_{dc} = 3.2\text{e+}03$ or $A_{dc} = 70$ dB. We can then deduce the output conductance $G_o = 2295.104\ \frac{pA}{V}$. We will split the output conductance G_o equally between the nMOS and pMOS cascodes.

To minimize the saturation voltage and maximize the current efficiency, the cascode transistors M_{4a}-M_{4b} and M_{5a}-M_{5b} are biased in weak inversion. We choose their inversion coefficient as $IC_5 = IC_4 = 0.1$, which gives a saturation voltage $V_{DSsat5} = V_{DSsat4} = 105\ mV$. Having set the inversion coefficient and knowing the bias current, we can deduce the specific current $I_{spec4} = 3.0\ \mu A$ and the aspect ratio $W_4/L_4 = 12.265$ for M_{4a}-M_{4b} and $I_{spec5} = 3.0\ \mu A$ and $W_5/L_5 = 4.235$ for M_{5a}-M_{5b}.

We can now calculate the source transconductances that are needed for the calculation of the output conductances $G_{ms4} = 10.625\ \mu A/V$ and $G_{ms5} = 10.625\ \mu A/V$. Having already the length of M_{3a}-M_{3b}, we can estimate its output conductance $G_{ds3} = 26.130\ nA/V$. We can then deduce the output conductance of M_{5a}-M_{5b} $G_{ds5} = 466.6\ nA/V$, which corresponds to a cascode gain $G_{ms5}/G_{ds5} = 22.8$. We can then deduce the length of M_{5a}-M_{5b} from G_{ds5} as $L_5 = 468\ nm$.

Because of the high output conductance of nMOS transistor we increase the length of M_{5a}-M_{5b} to $L_5 = 2.00\ \mu m$. Keeping the same W/L we get the width $W_5 = 8.47\ \mu m$.

We can recalculate the contribution of M_{3a}-M_{3b} and M_{5a}-M_{5b} to the total output conductance $G_{o35} = 268.3\ pA/V$ and deduce the remaining contribution allocated to M_{1a}-M_{1b}, M_{2b}-M_{2c} and M_{4a}-M_{4b} $G_{o124} = 2.0\ nA/V$.

At this point we don't know the lengths neither of M_{1a}-M_{1b} nor of M_{2b}-M_{2c} and M_{4a}-M_{4b}. We can choose that $G_{ds1} + G_{ds2} = G_{ds4}$ and $G_{ds1} = G_{ds2}$ so that $G_{ds1} = G_{ds2} = G_{ds4}/2$. We can then deduce the output conductance of M_{4a}-M_{4b} $G_{ds4} = 146.751\ \frac{nA}{V}$ from which we get the length $L_4 = 340\ nm$ and width $W_4 = 4.17\ \mu m$.

This might not be long enough. To secure enough margin for the DC gain we set $L_4 = 1.30\ \mu m$ which, for the same W/L, gives a width $W = 15.94\ \mu m$. The cascode voltage gain is then given by $G_{ms4}/G_{ds4} = 279.9$. We can deduce the length of M_{1a}-M_{1b} and M_{2b}-M_{2c} accounting for the different currents flowing in M_{4a}-M_{4b}, M_{1a}-M_{1b} and M_{2b}-M_{2c} $L_1 = 9.58\ \mu m$ and $L_2 = 4.77\ \mu m$.

We can now finalize the sizing of M_{1a}-M_{1b} with $W_1 = 33.81\ \mu m$.

This is a very wide transistor. We will reduce it by reducing the lenghth to $L_1 = 3.00\ \mu m$. Keeping the same W/L we get the width $W_1 = 10.59\ \mu m$. We can re-calculate the theoretical DC gain $A_{dc} = 1.3\text{e+}04$ (estimation) or $A_{dc} = 82$ dB (estimation).

3.6 Sizing of M_{2a}-M_{2b}-M_{2c}

To size M_{2b}-M_{2c} we will choose its saturation voltage to $V_{SDsat2} = 200\ mV$ in order to maximize the output swing. This leads to an inversion coefficient $IC_2 = 10.948$, a specific current $I_{spec2} = 0.050\ \mu A$ and an aspect ratio $W_2/L_2 = 0.205$. Knowing $L_2 = 4.77\ \mu m$ we deduce $W_2 = 980\ nm$.

The contribution to the input-referred noise PSD of M_{2b}-M_{2c} can be quite large compared to that of the differential pair M_{1a}-M_{1b} because the current I_{b2} is at least two times I_{b1}. It is therefore hard to make G_{m2} much smaller than G_{m1} without drastically increasing IC and hence V_{DSsat} and loosing output swing. We can check the contribution of M_{2a}-M_{2b}-M_{2c} to the input-referred noise. If M_{2a}-M_{2b}-M_{2c} are perfectly matched, the noise produced by M_{2a} is canceled out at the output and can therefore be neglected. Since the noise is usually dominated by the flicker noise we can calculate the contribution of M_{2b}-M_{2c} to the input-referred flicker noise relative to the contribution of the differential pair M_{1a}-M_{1b}

$$\eta_{fl}|_{M_{2b}\text{-}M_{2c}} = \frac{\rho_p}{\rho_n} \left(\frac{G_{m2}}{G_{m1}}\right)^2 \frac{W_1\, L_1}{W_2\, L_2} \tag{3.19}$$

We get $\eta_{fl}|_{M_{2b}\text{-}M_{2c}} = 13.842$ which is very large.

We can increase the length and width by a factor 2 resulting in $W_2 = 1.96\ \mu m$ and $L_2 = 9.54\ \mu m$ which reduces $\eta_{fl}|_{M_{2b}\text{-}M_{2c}}$ to 3.531.

We can also check the contribution of M_{3a}-M_{3b} to the input-referred noise by calculating

$$\eta_{fl}|_{M_{3a}\text{-}M_{3b}} = \left(\frac{G_{m3}}{G_{m1}}\right)^2 \frac{W_1\, L_1}{W_3\, L_3} \tag{3.20}$$

which is equal to 11.344.

3.7 Sizing M_{7a}-M_{7b}

The size of M_{7a}-M_{7b} is conditioned by the minimum input common-mode voltage $V_{ic,min}$

$$V_{ic,min} = V_{GS1} + V_{DSsat7}. \tag{3.21}$$

The gate-to-source voltage V_{GS1} is given by

$$V_{GS1} = V_{T0n} + (n_{0n} - 1)\, V_{S1} + n_{0n}\, U_T\, (v_p - v_s) \tag{3.22}$$

Unfortunately, at this point we don't know the value of the source voltage V_{S1} of M_{1a}-M_{1b} (voltage at common-source node 1). In weak inversion $v_p - v_s \cong 0$ so that

$$V_{GS1} \cong V_{T0n} + (n_{0n} - 1)\, V_{S1} \tag{3.23}$$

Since we don't know V_{S1} we can approximate V_{GS1} by $V_{T0n} = 246\ mV$. If we set the minimum input common-mode voltage to $V_{ic,min} = 0.5\ V$, it results in a saturation voltage voltage for M_{7a}-M_{7b} equal to $V_{DSsat7} = 254\ mV$, which corresponds to an inversion coefficient $IC_7 = 20.1$. Having the IC and the current we can derive the specific current $I_{spec7} = 25\ nA$ and the aspect ratio $W_7/L_7 = 0.035$. Since the W/L is small, we need to set the width to $W_7 = W_{min} = 150\ nm$ resulting in a length $L_7 = 4.27\ \mu m$.

We now need to set the bias voltages V_{b1} and V_{b2}.

3.8 Bias voltages V_{b1} and V_{b2}

We still need to calculate the required bias voltages V_{b1} for M_{5a}-M_{5b} and V_{b2} for M_{4a}-M_{4b}.

We start calculating the maximum bias voltage $V_{b2,max}$ still keeping M_{2a}-M_{2b} in saturation

$$V_{b2,max} = V_{DD} - V_{SG4} - V_{SDsat2} \tag{3.24}$$

The saturation voltage of M_{2a}-M_{2b} is given by IC_2 as $V_{SDsat2} = 200\ mV$. The source-to-gate voltage V_{SG4} is given by

$$V_{SG4} = V_{T0p} + (n_{0p} - 1)\, V_{BS4} + n_{0p}\, U_T\, (v_p - v_s) \tag{3.25}$$

In weak inversion $v_p - v_s \cong 0$ and therefore

$$V_{SG4} \cong V_{T0p} + (n_{0p} - 1)\, V_{BS4} \tag{3.26}$$

However, at this point we don't know the source voltage of M_{4a}-M_{4b}. We can estimate that $V_{SG4} \cong V_{T0p} = 365\ mV$ which results in $V_{b2,max} = 635\ mV$.

We also want to make sure that the V_{SD} voltage across M_{4a}-M_{4b} is large enough not to degrade the output conductance. We can choose $V_{SD4} = 200\ mV$, which gives $V_{b2,min} = 404\ mV$.

The maximum value of V_{b1} is set by keeping a sufficient V_{DS} voltage for M_{5a}-M_{5b}

$$V_{b1,max} = V_{GS5} - V_{DS5} + V_{GS3} \tag{3.27}$$

where

$$V_{GS5} = V_{T0n} + (n_{0n} - 1)\, V_{SB5} + n_{0n}\, U_T\, (v_p - v_s) \cong V_{T0n} + (n_{0n} - 1)\, V_{SB5} \cong V_{T0n}. \tag{3.28}$$

If we choose $V_{DS5} = 200\ mV$, we get $V_{b1,max} = 615\ mV$.

The minimum V_{b1} voltage is set by keeping M_{3a}-M_{3b} in saturation

$$V_{b1,min} = V_{GS5} + V_{DSsat3} \tag{3.29}$$

resulting in $V_{b1,min} = 518\ mV$.

So the bias voltages V_{b1} and V_{b2} shuld satisfy the following inequalities

$0.518\ V \leq V_{b1} \leq 0.615\ V$ and

$0.404\ V \leq V_{b2} \leq 0.635\ V$

We finally choose $V_{b1} = 0.550\ V$ and $V_{b2} = 0.450\ V$.

The sizing process is now completed. The transistor sizes and bias are summarized below.

3.9 Summary

3.9.1 Specifications

The specifications are recalled in Table 3.5.

Table 3.5: OTA specifications.

Specification	Symbol	Value	Unit
Minimum DC gain	A_{dc}	70	dB
Minimum gain-bandwidth product	GBW	1	MHz

Table 3.5: OTA specifications.

Specification	Symbol	Value	Unit
Load capacitance	C_L	1	pF
Maximum input-referred random offset voltage	$V_{os,max}$	10	mV
Phase margin	PM	60	$^\circ$

3.9.2 Bias

The bias information are summarized in Table 3.6.

Table 3.6: OTA bias.

Bias voltage or current	Symbol	Value	Unit
Supply voltage	V_{DD}	1.20	V
Bias current	I_{b1}	250.00	nA
Bias current	I_{b2}	550.00	nA
Cascode bias voltage	V_{b1}	0.55	V
Cascode bias voltage	V_{b2}	0.45	V

3.9.3 Transistor information

The transistor sizes and large-signal variables are summarized in Table 3.7, whereas Table 3.8 gives the small-signal and thermal noise parameters. An Excel table is generated with more information (e.g. all the parasitic capacitances).

Table 3.7: Transistor size and bias information.

Transistor	W $[\mu m]$	L $[\mu m]$	I_D $[nA]$	I_{spec} $[nA]$	IC	$V_G - V_{T0}$ $[mV]$	V_{DSsat} $[mV]$
M1a	10.59	3.00	250	2500	0.1	-47	105
M1b	10.59	3.00	250	2500	0.1	-47	105
M2a	3.92	9.54	1100	101	10.9	142	200
M2b	1.96	9.54	550	50	10.9	142	200
M2c	1.96	9.54	550	50	10.9	142	200
M3a	0.15	8.35	300	13	23.6	217	272
M3b	0.15	8.35	300	13	23.6	217	272
M4a	15.94	1.30	300	2999	0.1	-46	105
M4b	15.94	1.30	300	2999	0.1	-46	105
M5a	8.47	2.00	300	3000	0.1	-47	105
M5b	8.47	2.00	300	3000	0.1	-47	105
M7a	0.15	4.27	500	25	20.1	199	254
M7b	0.15	4.27	500	25	20.1	199	254

Table 3.8: Transistor small-signal and thermal noise parameters.

Transistor	G_{spec} $[\mu A/V]$	G_{ms} $[\mu A/V]$	G_m $[\mu A/V]$	G_{ds} $[nA/V]$	γ_n
M1a	96.668	8.855	7.258	60.606	0.627
M1b	96.668	8.855	7.258	60.606	0.627
M2a	3.886	11.058	8.991	18.971	0.767

Table 3.8: Transistor small-signal and thermal noise parameters.

Transistor	G_{spec} $[\mu A/V]$	G_{ms} $[\mu A/V]$	G_m $[\mu A/V]$	G_{ds} $[nA/V]$	γ_n
M2b	1.943	5.529	4.495	9.485	0.767
M2c	1.943	5.529	4.495	9.485	0.767
M3a	0.492	2.155	1.767	26.130	0.776
M3b	0.492	2.155	1.767	26.130	0.776
M4a	115.955	10.625	8.638	37.968	0.632
M4b	115.955	10.625	8.638	37.968	0.632
M5a	115.974	10.625	8.709	109.091	0.627
M5b	115.974	10.625	8.709	109.091	0.627
M7a	0.962	3.858	3.162	85.161	0.773
M7b	0.962	3.858	3.162	85.161	0.773

4 OTA Characteristics

In this section, we check whether the specs are achieved.

4.1 Open-loop gain

We can calculate the various OTA features related to the open-loop transfer function, which are given in Table 4.1.

Table 4.1: OTA gain variables.

Symbol	Theoretical Value	Unit
A_{dc}	82.917	dB
G_{m1}	7.258	$\mu A/V$
G_{m2}	4.495	$\mu A/V$
G_{m3}	1.767	$\mu A/V$
G_{ms4}	10.625	$\mu A/V$
G_{ms5}	10.625	$\mu A/V$
G_{ds1}	60.606	nA/V
G_{ds2}	9.485	nA/V
G_{ds3}	26.130	nA/V
G_{ds4}	37.968	nA/V
G_{ds5}	109.091	nA/V
C_3	25.931	fF
C_4	96.248	fF
f_0	80.694	Hz
GBW	1.129	MHz
f_p	10.843	MHz
f_z	21.686	MHz
f_{p4}	17.570	MHz

The gain-bandwidth product from the specifications is repeated here

$GBW = 1.000$ MHz (from spec).

The estimate value assuming that all the non-dominant poles are much higher than the GBW is given by

$GBW_{est} = 1.129$ MHz (estimation).

The GBW accounting for the effect of the additional non-dominant poles is given by

$GBW_{the} = 1.122$ MHz (theory).

We see that there is only a small difference between GBW_{est} and GBW_{the}, which confirms that the non-dominant poles are sufficiently far from GBW as stated in Table 4.1.

We can now plot the gain response Using the variables given in Table 4.1. It is shown in Figure 4.1.

From Figure 4.1, we see that the GBW and DC gain are met.

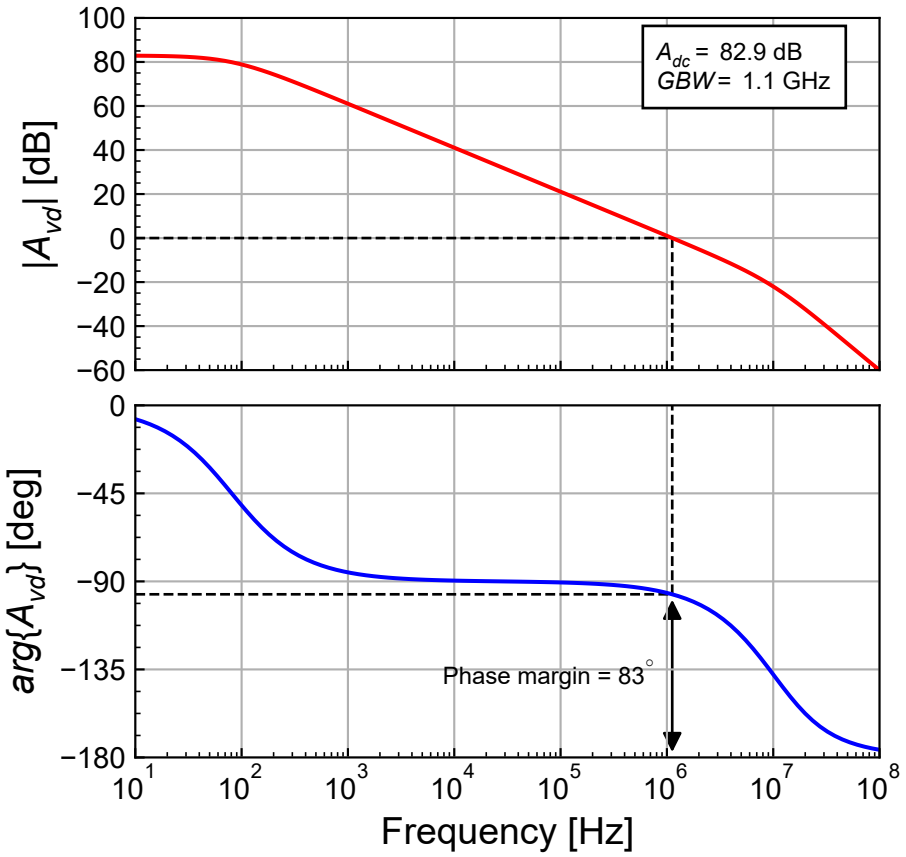

Figure 4.1: OTA theoretical transfer function.

We can now have a look at the input-refrred noise PSD.

4.2 Input-referred noise

We can now compute all the parameters needed for the calculation of the OTA thermal noise excess factor and its input-referred thermal noise PSD and resistance. They are given in Table 4.2.

Table 4.2: OTA thermal noise parameters.

Symbol	Theoretical Value	Unit
G_{m1}	7.258	$\mu A/V$
G_{m2}	4.495	$\mu A/V$
G_{m3}	1.767	$\mu A/V$
G_{m1}/G_{m2}	1.615	-
G_{m1}/G_{m3}	4.108	-
γ_{n1}	0.627	-
γ_{n2}	0.767	-
γ_{n3}	0.776	-
η_{th}	1.058	-
γ_{ota}	2.581	-
R_{nt}	355.675	$k\Omega$
S_{ninth}	5.9e-15	V^2/Hz
$\sqrt{S_{ninth}}$	76.783	$nV/\sqrt{Hz}$
$10 \cdot \log(S_{ninth})$	-142.295	$dBv/\sqrt{Hz}$

From Table 4.2, we see that the contribution of M_{2b}-M_{2c} and M_{3a}-M_{3b} is only 1.058 that of M_{1a}-M_{1b}. This leads to an OTA thermal noise excess factor $\gamma_{n,ota} = 2.581$, which is 2.1 times larger than that of the differential pair $2\,\gamma_{n1} = 1.254$.

We can now compute all the parameters needed for the calculation of the input-referred flicker noise and the corner frequency. They are given in Table 4.3.

Table 4.3: OTA flicker noise parameters.

Symbol	Theoretical Value	Unit
$(G_{m1}/G_{m2})^2$	2.6	-
$(G_{m1}/G_{m3})^2$	16.9	-
ρ_p/ρ_n	5.4	-
$\frac{W_1 \cdot L_1}{W_2 \cdot L_2}$	1.7	-
$\frac{W_1 \cdot L_1}{W_3 \cdot L_3}$	29.6	-
η_{fl}	5.284	-
$\sqrt{S_{ninfl}(1\,Hz)}$	7.5	$\mu V/\sqrt{Hz}$
$10 \cdot \log(S_{ninfl}(1\,Hz))$	-102.5	$dBv/\sqrt{Hz}$
f_k	9	kHz

As desired, from Table 4.3, we see that the contribution of M_{2b}-M_{2c} and M_{3a}-M_{3b} to the flicker noise is only 5.284 that of M_{1a}-M_{1b}. It is sligthly lower than the value obtained in the design phse simply because the calculations use the effective width and length.

We can plot the input-reffered noise which is shown in Figure 4.2.

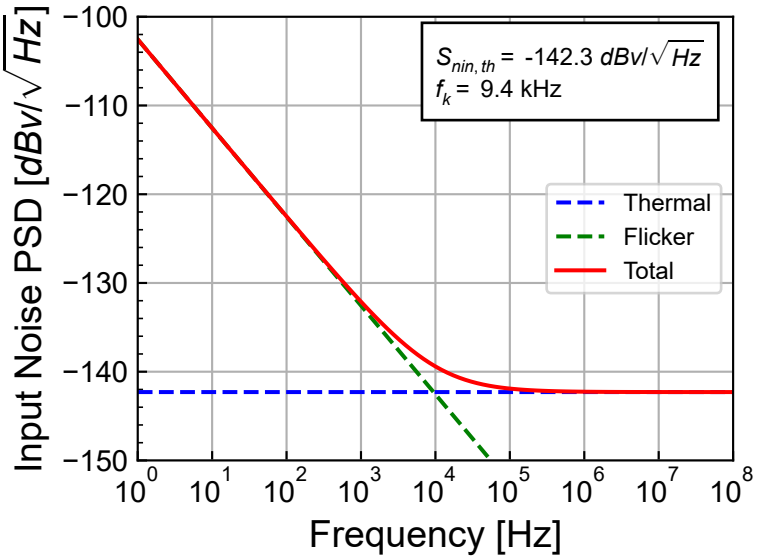

Figure 4.2: OTA theoretical input-referred noise PSD.

4.3 Input-referred offset

The variance of the input-referred offset voltage is given by (2.32), which is repeated below

$$\sigma^2_{V_{os}} = \sigma^2_{V_T} + \sigma^2_{\beta}. \tag{4.1}$$

$\sigma^2_{V_T}$ is the V_T-mismatch given by

$$\sigma^2_{V_T} = \sigma^2_{V_{T1}} \cdot (1 + \xi_{V_T}) \tag{4.2}$$

where

$$\xi_{V_T} = \left(\frac{G_{m2}}{G_{m1}}\right)^2 \left(\frac{A_{V_{Tp}}}{A_{V_{Tp}}}\right)^2 \frac{W_1\,L_1}{W_2\,L_2} + \left(\frac{G_{m3}}{G_{m1}}\right)^2 \frac{W_1\,L_1}{W_3\,L_3} \tag{4.3}$$

represents the V_T-mismatch contributions to the input-referred offset of the current source M_{2b}-M_{2c} and current mirror M_{3a}-M_{3b} relative to that of the differential pair.

σ_β^2 is the β-mismatch given by

$$\sigma_\beta^2 = \left(\frac{I_{b1}}{G_{m1}}\right)^2 \cdot \sigma_{\beta_1}^2 \cdot (1 + \xi_\beta) \tag{4.4}$$

where

$$\xi_\beta = \left(\frac{I_{b2}}{I_{b1}}\right)^2 \left(\frac{A_{\beta_p}}{A_{\beta_n}}\right)^2 \frac{W_1\,L_1}{W_2\,L_2} + \left(\frac{I_{b3}}{I_{b1}}\right)^2 \frac{W_1\,L_1}{W_2\,L_2}. \tag{4.5}$$

The parameter for calculating the standard deviation of the input-referred offset voltage are given in Table 4.4.

Table 4.4: OTA input-referred offset parameters.

Symbol	Theoretical Value	Unit
σ_{VT1}	0.887	mV
σ_{VT2}	1.156	mV
σ_{VT3}	4.468	mV
$\sigma_{\beta 1}$	0.177	%
$\sigma_{\beta 2}$	0.231	%
$\sigma_{\beta 3}$	0.894	%
$(G_{m2}/G_{m1})^2$	0.384	-
$(G_{m3}/G_{m1})^2$	0.059	-
$W_1\,L_1/(W_2\,L_2)$	1.699	-
$W_1\,L_1/(W_3\,L_3)$	25.365	-
$(I_{b2}/I_{b1})^2$	4.840	-
$(I_{b3}/I_{b1})^2$	1.440	-
ξ_{VT}	2.155	-
ξ_β	44.750	-
$\sigma_{V_T}^2$	2.482	mV^2
σ_{V_T}	1.576	mV
σ_β^2	0.171	mV^2
σ_β	0.413	mV
σ_{Vos}	1.629	mV

From Table 4.4 and contrary to the other OTA, we see that in the case of the folded cascode OTA the β-mismatch cannot be neglected. This is mostly because of the $W_1\,L_1/(W_3\,L_3)$ ratio equal to 25.365.

4.4 Current and power consumption

The total current consumption, ignoring the current drawn by M_{2a} and M_{7a}, is $I_{tot} = 2I_{b2} = 1.1\ \mu A$ and the power consumption is $P = 1.320\ \mu W$.

Assuming the input differential pait M_{1a}-M_{1b} is biased in deep weak inversion, the minimum bias current $I_{b1,min}$ is directly related to the gain-bandwidth product GBW according to

$$I_{b1,min} \cong nU_T \cdot C_L \cdot \omega_u = nU_T \cdot C_L \cdot 2\pi\,GBW, \tag{4.6}$$

which is equal to 198 nA. The minimum total current consumption assuming $\alpha \triangleq I_{b3}/I_{b1} = 1.2$ can then be estimated as $I_{tot,min} \cong 2I_{b2,min} = 2(I_{b1,min} + I_{b3,min}) = 2I_{b1,min}\,(1+\alpha) = 872\ nA$. The actual current consumption accounting for some margin taken on the GBW is therefore 26% higher than the minimum.

The current consumption of the folded cascode OTA is 2.2 times larger than that of the telescopic OTA for the same specifications.

> **Note**
>
> Note that the telescopic OTA is the differential OTA that has the minimum power consumption for similar gain-bandwidth product, DC gain and phase margin specifications. Of course it has the smallest input and output voltage swings.

The above design will now be checked against simulations.

5 Simulation results from ngspice

The theoretical results can be validated with the results obtained from simulations performed with ngspice. In order to run the simulations you need to have ngspice correctly installed. Please refer to the installation instructions.

> **Note**
>
> The simulations are performed with the PSP 103.6 compact model [4] using the parameters from the IHP open source PDK [2]. For ngspice, we use the Verilog-A implementation given in the IHP package [2] and compiled the OSDI file with OpenVAF [7] to run with ngspice [8]. In addition to the PSP user manual [4] a documentation of PSP and other MOSFET compact models and their parameter extraction can be found in [9].

5.1 Operating point

We first write the parameter file for this specific design for running the ngspice simulations. Before running the AC and NOISE simulations, we first need to check the quiescent voltages and currents and the operating points of all transistors by running a .OP simulation.

Table 5.1: OTA node voltages with the OTA in open-loop without offset correction.

Node	Voltage
vdd	1.2
vb1	0.55
vb2	0.45
inp	0.6
inn	0.6
out	0.528168
ic	0.6
id	0
1	0.421646
2	0.607732
3	0.51668
4	0.823598
5	0.82362
6	0.352809
7	0.353196
8	0.487271

We can extract the OTA quiescent node voltages from the ngspice .ic file. They are presented in Table 5.1. We see that the simulated quiescent output voltage is $V_{outq} = 528\ mV$. This is fine and the OTA is biased in the high gain region. Similarly to the telescopic OTA, the quiescent output voltage is actually set by the V_{GS} voltage of M_{3a}. Indeed, if perfect matching is assumed (which is the case in the simulation), since M_{5a}-M_{5b} have the same drain current and share the same gate voltage, they

have the same V_{GS} voltage and also have the same drain voltage. We therefore don't need to extract any offset voltage at this point and can proceed with the simulation of the large-signal characteristic.

The operating point information for all transistors are extracted from the ngspice .op file. The data is split into the large-signal operating informations in Table 5.2, the small-signal operating point informations in Table 5.3 and the noise operating point informations in Table 5.4.

Table 5.2: Operating point information extracted from ngspice .op file for each transistor.

Transistor	I_D $[nA]$	V_{GS} $[mV]$	V_{DS} $[mV]$	V_{SB} $[mV]$	$V_{GS} - V_T$ $[mV]$	V_{Dsat} $[mV]$
M1a	247.977	178	402	422	-51	109
M1b	247.981	178	402	422	-51	109
M2a	1099.885	592	592	-0	250	232
M2b	574.355	592	376	-0	251	233
M2c	574.354	592	376	-0	251	233
M3a	326.339	517	353	0	355	300
M3b	326.357	517	353	0	355	300
M4a	326.358	374	307	376	-37	106
M4b	326.353	374	295	376	-37	106
M5a	326.334	197	164	353	-42	110
M5b	326.352	197	175	353	-43	110
M7a	499.992	487	487	0	313	272
M7b	495.960	487	422	0	313	272

Table 5.3: PSP small-signal operating point information extracted from ngspice .op file for each transistor.

Transistor	G_m $[\mu A/V]$	G_{mb} $[\mu A/V]$	G_{ds} $[nA/V]$
M1a	7.529	0.773	202.371
M1b	7.529	0.773	202.369
M2a	8.218	1.626	17.100
M2b	4.273	0.840	21.518
M2c	4.273	0.840	21.522
M3a	1.880	0.261	47.519
M3b	1.881	0.261	47.338
M4a	9.512	1.524	21.115
M4b	9.508	1.523	21.999
M5a	9.711	1.067	378.445
M5b	9.715	1.065	376.234
M7a	3.256	0.447	58.346
M7b	3.232	0.444	64.920

Table 5.4: PSP noise operating point information extracted from ngspice .op file for each transistor.

Transistor	$S_{ID,th}$ $[A^2/Hz]$	$S_{ID,fl}$ at 1Hz $[A^2/Hz]$
M1a	7.501e-26	2.857e-22
M1b	7.501e-26	2.857e-22
M2a	1.995e-25	1.481e-21
M2b	1.039e-25	8.116e-22
M2c	1.039e-25	8.116e-22
M3a	2.397e-26	7.247e-22

Table 5.4: PSP noise operating point information extracted from ngspice .op file for each transistor.

Transistor	$S_{ID,th}$ $[A^2/Hz]$	$S_{ID,fl}$ at 1Hz $[A^2/Hz]$
M3b	2.397e-26	7.249e-22
M4a	1.837e-25	2.308e-21
M4b	1.837e-25	2.309e-21
M5a	9.724e-26	1.080e-21
M5b	9.728e-26	1.068e-21
M7a	4.098e-26	4.086e-21
M7b	4.066e-26	4.039e-21

Table 5.5: sEKV parameters calculated from the values extracted from the simulation to compare with Table 3.7.

Transistor	W_{eff} $[\mu m]$	L_{eff} $[\mu m]$	W_{eff}/L_{eff}	I_{spec} $[\mu A]$	IC
M1a	10.610	2.941	3.607	2.555	0.097
M1b	10.610	2.941	3.607	2.555	0.097
M2a	3.890	9.586	0.406	0.099	11.081
M2b	1.930	9.586	0.201	0.049	11.663
M2c	1.930	9.586	0.201	0.049	11.663
M3a	0.170	8.291	0.021	0.015	22.471
M3b	0.170	8.291	0.021	0.015	22.472
M4a	15.910	1.344	11.839	2.896	0.113
M4b	15.910	1.344	11.839	2.896	0.113
M5a	8.490	1.941	4.374	3.098	0.105
M5b	8.490	1.941	4.374	3.098	0.105
M7a	0.170	4.211	0.040	0.029	17.486
M7b	0.170	4.211	0.040	0.029	17.345

Table 5.6: sEKV small-signal parameters calculated from the values extracted from the simulation to compare with Table 3.8.

Transistor	G_{spec} $[\mu A/V]$	n	G_m $[\mu A/V]$	G_{ms} $[\mu A/V]$	G_{ds} $[nA/V]$
M1a	98.788	1.103	7.529	8.302	202.371
M1b	98.788	1.103	7.529	8.302	202.369
M2a	3.838	1.198	8.218	9.844	17.100
M2b	1.904	1.197	4.273	5.113	21.518
M2c	1.904	1.197	4.273	5.113	21.522
M3a	0.561	1.139	1.880	2.142	47.519
M3b	0.561	1.139	1.881	2.142	47.338
M4a	111.958	1.160	9.512	11.036	21.115
M4b	111.958	1.160	9.508	11.031	21.999
M5a	119.772	1.110	9.711	10.778	378.445
M5b	119.772	1.110	9.715	10.780	376.234
M7a	1.105	1.137	3.256	3.703	58.346
M7b	1.105	1.137	3.232	3.676	64.920

Table 5.7: sEKV noise parameters calculated from the values extracted from the simulation to compare with Table 4.2.

Transistor	$\sqrt{S_{nin,th}}$ $[nV/\sqrt{Hz}]$	$R_{nin,th}$ $[k\Omega]$	γ_n $[-]$	$\sqrt{S_{nin,fl}}$ $[nV/\sqrt{Hz}]$
M1a	36.378	79.834	0.601	2244.920
M1b	36.377	79.833	0.601	2244.902
M2a	54.354	178.229	1.465	4683.598
M2b	75.427	343.223	1.466	6667.880
M2c	75.427	343.224	1.466	6667.882
M3a	82.342	409.033	0.769	14316.833
M3b	82.333	408.942	0.769	14316.983
M4a	45.055	122.461	1.165	5050.903
M4b	45.073	122.561	1.165	5054.063
M5a	32.111	62.206	0.604	3383.623
M5b	32.106	62.186	0.604	3363.983
M7a	62.167	233.152	0.759	19630.733
M7b	62.384	234.783	0.759	19663.094

Table 5.8: Bias voltages and operating regions extracted from ngspice for each transistor.

Trans.	Type	Funct.	V_G $[V]$	V_S $[V]$	V_D $[V]$	V_{DS} $[mV]$	V_{DSsat} $[mV]$	Reg.	Sat.
M1a	n	DP	0.600	0.422	0.824	402	109	WI	sat
M1b	n	DP	0.600	0.422	0.824	402	109	WI	sat
M2a	p	CM	0.592	0.000	0.592	592	232	SI	sat
M2b	p	CM	0.592	0.000	0.376	376	233	SI	sat
M2c	p	CM	0.592	0.000	0.376	376	233	SI	sat
M3a	n	CM	0.517	0.000	0.353	353	300	SI	sat
M3b	n	CM	0.517	0.000	0.353	353	300	SI	sat
M4a	p	CA	0.750	0.376	0.683	307	106	MI	sat
M4b	p	CA	0.750	0.376	0.672	295	106	MI	sat
M5a	n	CA	0.550	0.353	0.517	164	110	MI	sat
M5b	n	CA	0.550	0.353	0.528	175	110	MI	sat
M7a	n	CM	0.487	0.000	0.487	487	272	SI	sat
M7b	n	CM	0.487	0.000	0.422	422	272	SI	sat

From Table 5.1, we see that all transistors have a sufficiently large V_{DS} voltage and are therefore biased in saturation. Additionnally we see that all the saturation voltages are summing up to 749 mV leaving an output voltage swing equal to $V_{out,swing} = 451\ mV$.

From Table 5.3, we see that the output conductance of M_{5a}-M_{5b} is very large (much larger than the theoretical estimation) despite we have chosen a very long transistor $L_5 = 2.00\ \mu m$. This will impact the small-signal DC gain as shown below.

The data extracted from the operating point simulation can be translated into sEKV parameters. Table 5.5 presents the effective width and length, the specific current and the resulting inversion coefficient which can be compared to the values resulting from the design given in Table 3.7. We observe that the values are close.

Table 5.6 presents the sEKV parameters including the specific conductance G_{spec}, slope factor n, gate transconductance G_m, source transconductance G_{ms} and output conductance G_{ds}. They can be compared to the results of the design presented in Table 3.8. Again, we see that the simulated values of the transconductances are close to the theoretical estimation. However, the simulated values of the

output conductances are much larger than the estimated values, explaining why the estimated DC gain is much larger than the simulated value.

Table 5.7 presents the sEKV noise parameters. We can observe that the thermal noise excess factors of pMOS transistors are larger than the estimated values shown in Table 4.2.

Finally we can check whether all transistors are biased in saturation. From Table 5.8, we see that all transistors are biased in saturation. The operating points look fine. We can now proceed with the simulation of the open-loop large-signal transfer characteristic.

5.2 Large-signal differential transfer characteristic

We now simulate the DC differential transfer characteristic. The simulation of the large-signal input-output characteristic is presented in Figure 5.1.

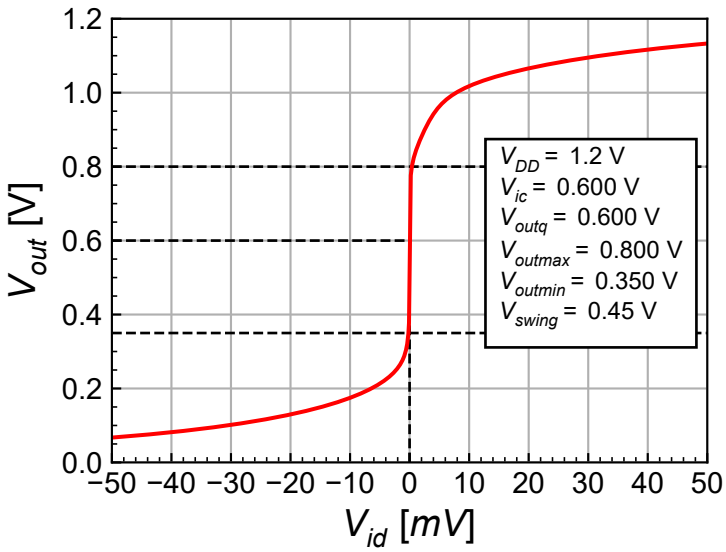

Figure 5.1: Simulated large-signal input-output characteristic.

From Figure 5.1, we see that the output swing is about $V_{out,swing} = 450\ mV$, which corresponds to the above estimation. We can now zoom into the high gain region in order to extract the offset voltage that is needed to bring the output voltage back to $V_{outq} = 0.600\ V$. The simulation results are presented in Figure 5.2.

We can now save the extracted offset voltage $V_{os} = 19.992\ \mu V$ that is required to bring the output voltage to $V_{outq} = 0.600\ V$ and that will be used for the following .AC and .NOISE simulations.

5.3 Open-loop gain

After having checked the operating point information and making sure that the OTA output is not saturated by extracting the required offset voltage for bringing the output operating point to the desired value $V_{outq} = 0.600\ V$, we can now perform the AC simulation.

From Figure 5.3, we see that the simulated gain-bandwidth product $GBW = 1.141\ MHz$ is about equal to the theoretical estimation 1.122 MHz and slightly above target 1.000 MHz. The simulated DC gain $A_{dc} = 70.962\ dB$ is way lower than the estimated DC gain 82.917 dB but still above the specifications 70 dB. This low DC gain is mainly due to the output conductance of M_{5b} which is very high despite the fact that the transistor length has been significantly increased and that the V_{DS} is

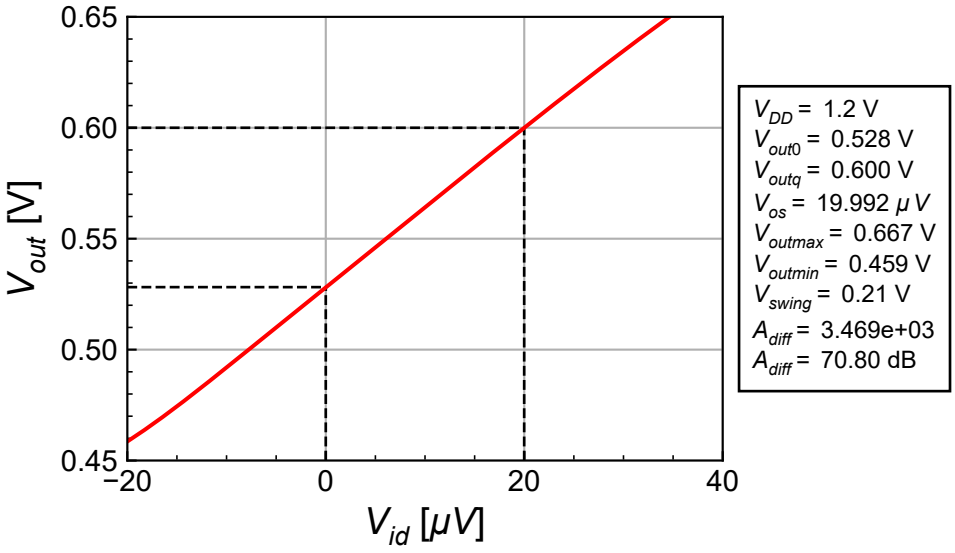

Figure 5.2: Zoom of the simulated large-signal input-output characteristic in the gain region.

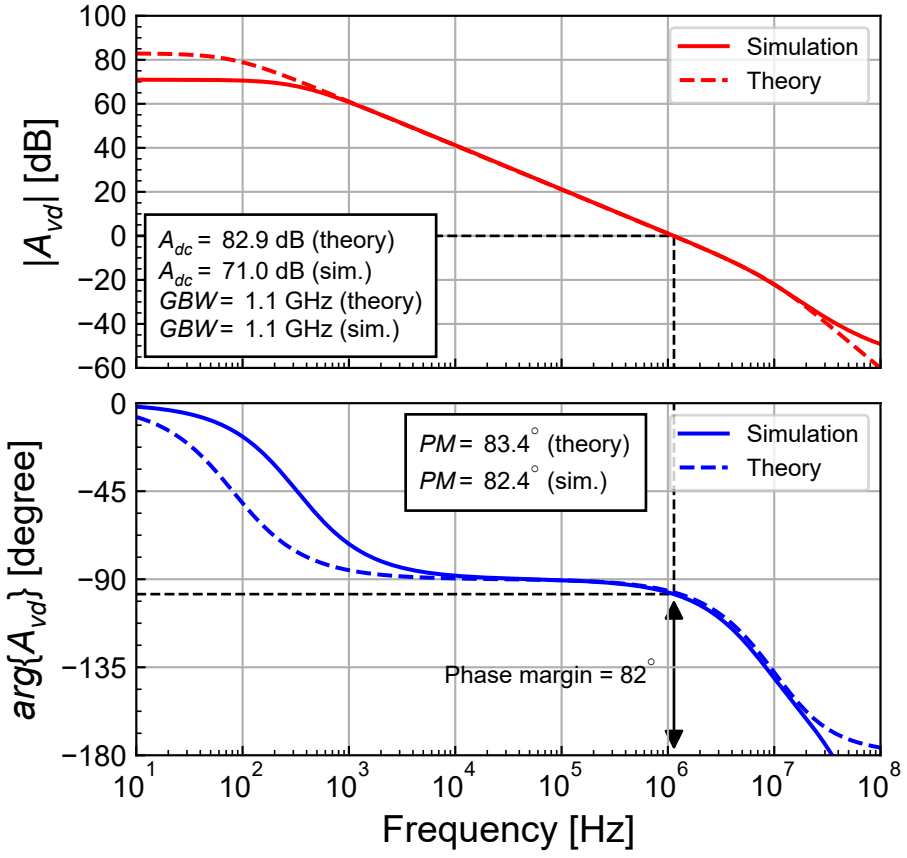

Figure 5.3: Simulated gain response compared to theoretical estimation.

large enough. This is obviously due to the poor output conductance of the nMOS transistors and the low voltage constraint.

5.4 Input-referred noise

The simulated input-referred noise PSD is presented in Figure 5.4 and compared to the theoretical estimation.

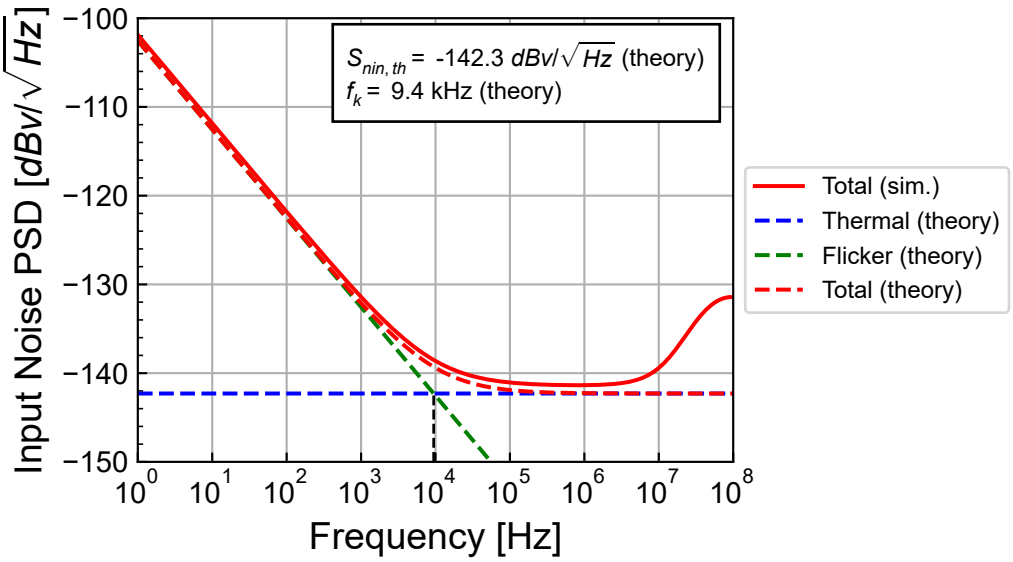

Figure 5.4: Simulated input-referred noise PSD compared to theoretical estimation.

From Figure 5.4, we see that the simulated input-referred noise PSD is very close to the theoretical prediction. The simulated flicker noise is exactly equal to the theoretical estimation, while the simulated white noise is slightly higher. We can have a closer look at the contributions of the various transistors to the input-referred white noise PSD.

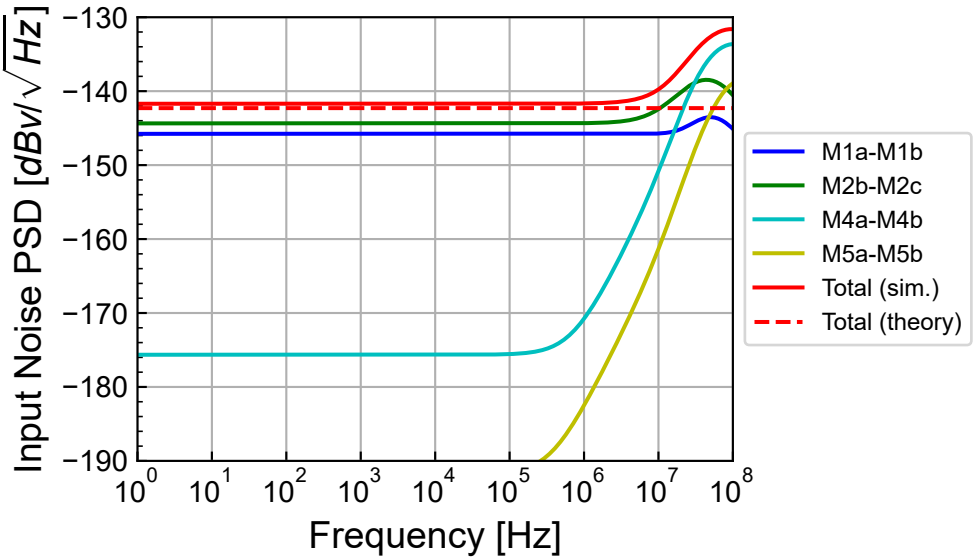

Figure 5.5: Breakdown of the contributions to the simulated input-referred white noise PSD.

The contributions of M_{1a}-M_{1b}, M_{2b}-M_{2c}, M_{4a}-M_{4b} and M_{5a}-M_{5b} to the input-referred white noise PSD are detailed in Figure 5.5 and compared to the theoretical white noise. Note that the contribution

of M_{3a}-M_{3b} to the input-referred white noise PSD could not be extracted because it is very small. We can observe that the current mirror M_{2b}-M_{2c} is contributing a bit more than the differential pair M_{1a}-M_{1b}. This means that the simulated value of η_{th} is larger than the theoretical value $\eta_{th} = 1.058$ given in Table 4.2. This is probably due to the fact that the simulated thermal noise excess factors of pMOS transistors shown in Table 5.7 are larger than the predicted values given in Table 4.2. As expected, the contributions of the cascode transistors M_{4a}-M_{4b} is 30dB lower and that of M_{5a}-M_{5b} is about 50 dB lower, conforming that they can indeed be neglected. The total simulated white noise is slightly higher than the theoretical estimation, which is acceptable. This results in an OTA thermal noise excess factor $\gamma_{n,ota} = 3.066$ that is slightly larger than the predicted value $\gamma_{n,ota} = 2.581$. This is a rather high value of the thermal noise excess factor compared to other OTAs. Remember that this is due to the small gain $G_{m1}/G_{m2} = 1.762$ because M_{2b}-M_{2c} have a drain current 2.2 times larger than M_{1a}-M_{1b}. Assuming that M_{1a}-M_{1b} are biased in weak inversion and M_{2b}-M_{2c} in strong inversion, the G_m ratio can then be written as

$$\frac{G_{m1}}{G_{m2}} \cong \frac{I_{b1}}{n_{0n}\,U_T} \cdot \frac{n_{0n}\,V_{P2}}{2\,I_{b2}} = \frac{I_{b1}}{I_{b2}} \cdot \frac{V_{G2} - V_{T0n}}{2\,n_{0n}\,U_T}. \tag{5.1}$$

Now

$$\frac{I_{b2}}{I_{b1}} = \frac{I_{b1} + I_{b3}}{I_{b1}} = 1 + \alpha \tag{5.2}$$

and hence

$$\frac{G_{m1}}{G_{m2}} = \frac{1}{1+\alpha} \cdot \frac{V_{G2} - V_{T0n}}{2\,n_{0n}\,U_T}. \tag{5.3}$$

The G_{m1}/G_{m2} gain can be increased by increasing the overdrive voltage $V_{G2} - V_{T0n}$ or the inversion coefficient at the cost of a higher saturation voltage and hence a lower output voltage swing.

Figure 5.6 presents the breakdown of the contributions of the various transistors to the input-referred flicker noise. We see that it is dominated by the contributions of the current mirrors M_{2b}-M_{2c} and M_{3a}-M_{3b} which are larger than the contribution of the differential pair M_{1a}-M_{1b}. This is consistent with the value of $\eta_{fl} = 5.284$ given in Table 4.3.

The breakdown of the contributions of the various transistors to the total input-referred noise is presented in Figure 5.7. We can observe that the simulation is close to the theoretical estimation.

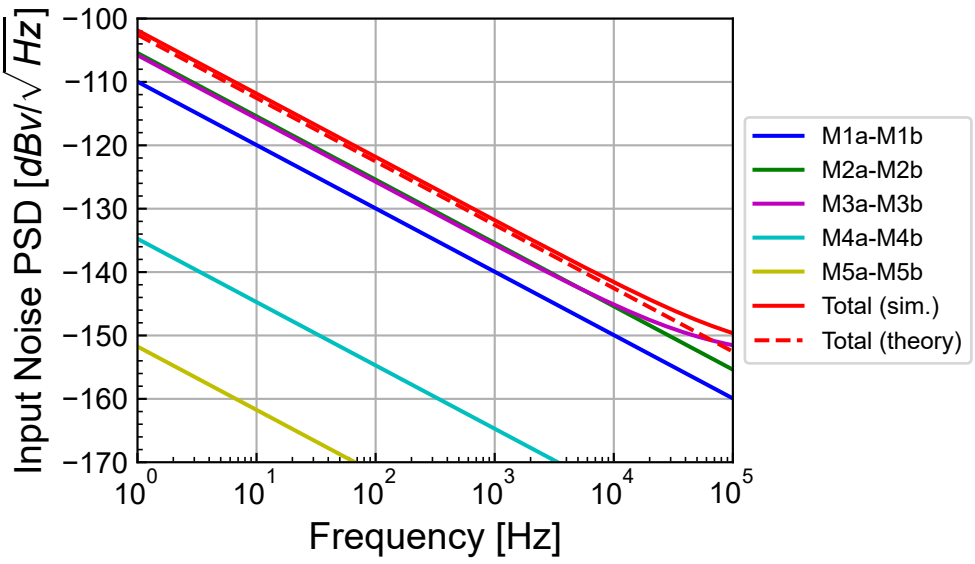

Figure 5.6: Breakdown of the contributions to the simulated input-referred flicker noise PSD.

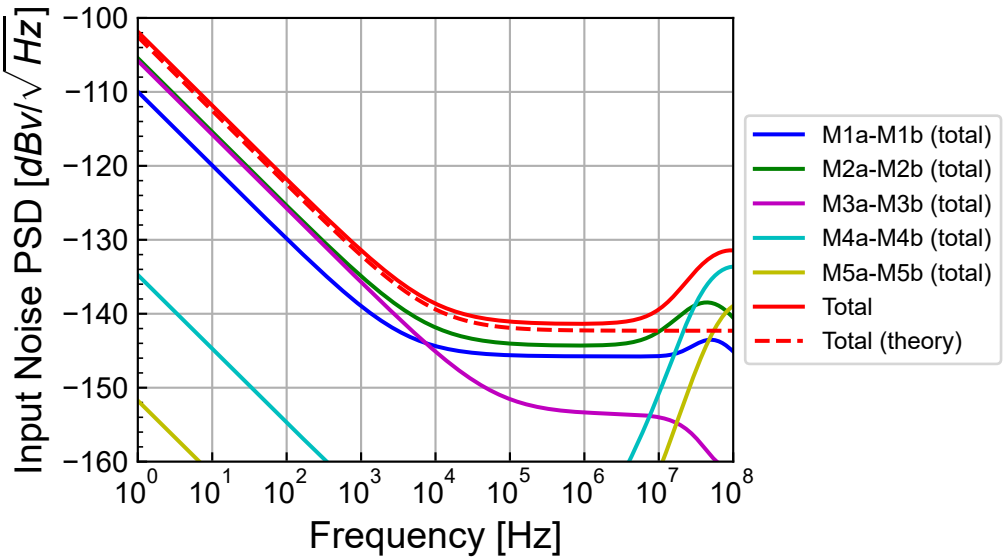

Figure 5.7: Breakdown of the contributions to the simulated input-referred noise PSD.

5.5 Input common-mode voltage range

We can check the input common-mode voltage range by connecting the OTA as a voltage follower and sweeping the positive input. As shown in Figure 5.8, the output follows the input voltage from $V_{in} = 0.22\,V$ to about $V_{in} = 1\,V$ resulting in an input common-mode voltage range of about $0.78\,V$.

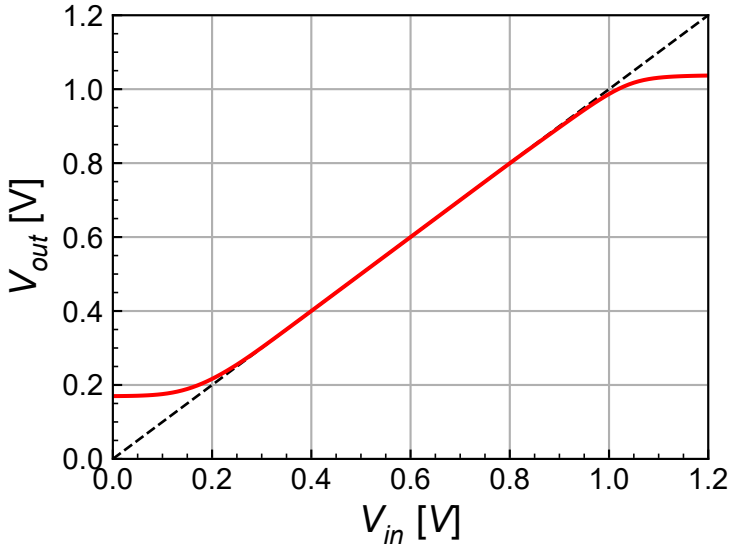

Figure 5.8: Simulated input common-mode voltage range.

5.6 Step-response

In this section we will check the step response of the OTA operating as a voltage follower as shown in Figure 5.9 with its output connected to the negative input and with the same load capacitance $C_L = 1$ pF.

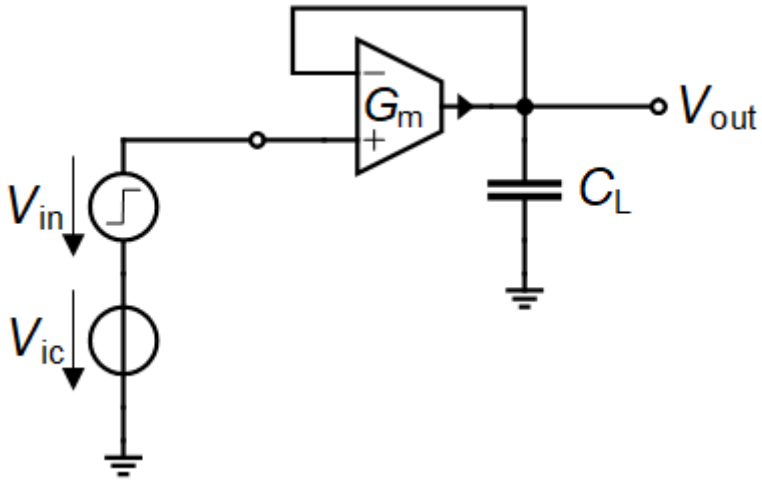

Figure 5.9: Schematic of the OTA connected as a voltage follower.

5.6.1 Small-step

According to the input common-mode voltage range established above, we will set the input common-mode voltage to $V_{ic} = 0.600\ V$ to make sure that the OTA is in the high gain region. We start by imposing a small step $\Delta V_{in} = 10\ mV$ on top of a common mode voltage $V_{ic} = 0.600\ V$. The simulation results are shown in Figure 5.10 where $\Delta V_{in}(t) \triangleq V_{in+}(t) - V_{ic}$ and $\Delta V_{out}(t) \triangleq V_{out}(t) - V_{outq}$ with $V_{outq} \cong V_{ic}$ is the quiescent output voltage. ΔV_{in} and ΔV_{out} are compared to the response of a single pole circuit having a cut-off frequency equal to the GBW. From Figure 5.10, we see that the simulation result is very close to the first-order response.

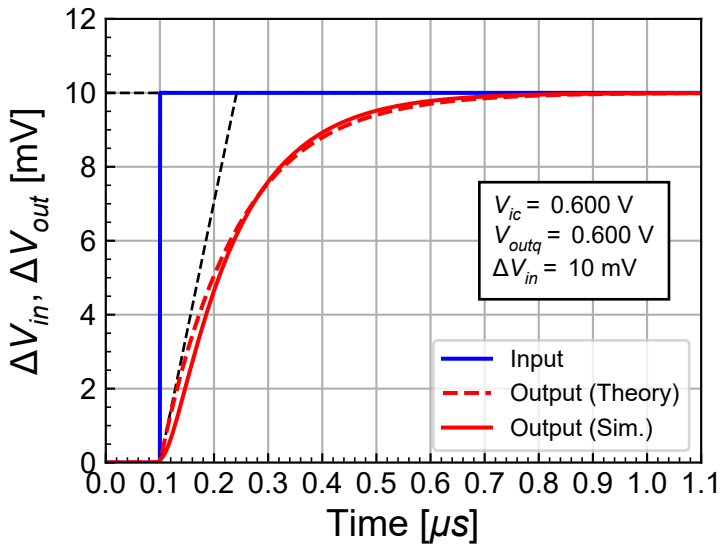

Figure 5.10: Step response of the OTA connected as a voltage follower for a small input step.

5.6.2 Large step

Since we now impose a larger step $\Delta V_{in} = 300\ mV$, we need to lower the input common-mode voltage to $V_{ic} = 400\ mV$, to make sure that after the step the OTA remains in the high gain region and correctly settles to the right voltage. The simulation results are shown in Figure 5.11 where $\Delta V_{in}(t) \triangleq V_{in+}(t) - V_{ic}$ and $\Delta V_{out}(t) \triangleq V_{out}(t) - V_{outq}$ with $V_{outq} \cong V_{ic}$ is the quiescent output voltage. ΔV_{in} and ΔV_{out} are compared to the response of a single pole circuit having a cut-off frequency equal to the GBW. We now observe the effect of slew-rate which increases the settling time.

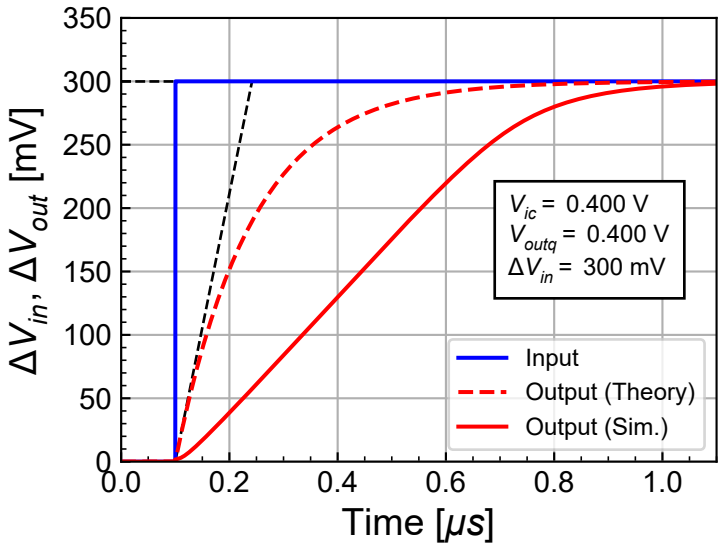

Figure 5.11: Step response of the OTA connected as a voltage follower for a large input step highlighting the slew-rate effect.

5.7 Current and power consumption

The total current consumption, ignoring the current drawn by M_{2a} and M_{7a}, is $I_{tot} = 2I_{b2} = 1.1\ \mu A$ and the power consumption is $P = 1.320\ \mu W$.

Assuming the input differential pait M_{1a}-M_{1b} is biased in deep weak inversion, the minimum bias current $I_{b1,min}$ is directly related to the gain-bandwidth product GBW according to

$$I_{b1,min} \cong nU_T \cdot C_L \cdot \omega_u = nU_T \cdot C_L \cdot 2\pi\, GBW, \tag{5.4}$$

which is equal to 198 nA. The minimum total current consumption assuming $\alpha \triangleq I_{b3}/I_{b1} = 1.2$ can then be estimated as $I_{tot,min} \cong 2I_{b2,min} = 2(I_{b1,min} + I_{b3,min}) = 2I_{b1,min}\,(1+\alpha) = 872\ nA$. The actual current consumption accounting for some margin taken on the GBW is therefore 26% higher than the minimum.

The current consumption of the folded cascode OTA is 2.2 times larger than that of the telescopic OTA for the same specifications.

> **Note**
>
> Note that the telescopic OTA is the differential OTA that has the minimum power consumption for similar gain-bandwidth product, DC gain and phase margin specifications. Of course it has the smallest input and output voltage swings.

6 Conclusion

This notebook presented the detailed analysis, design and verification by simulation of the folded cascode OTA [1]. The analysis allowed to derive the design equations to achieve the target specifications. The OTA has been designed for specifications on the gain-bandwidth product and DC gain in the IHP SG13G2 130nm BiCMOS technology [2] using the inversion coefficient approach with the sEKV model and parameters [10] [11] [12]. The design has been validated by simulations with the ngspice simulator [3] using the PSP compact model [4] and the parameters provided by the open source IHP [2].

The simulations have shown that the target gain-bandwidth product GBW was achieved. The simulated DC gain is slightly less than the specification, despite the increase of the nMOS transistor length in the design phase. The estimated DC gain was $A_{dc} \cong 82.9\ dB$, way above the simulated value of $A_{dc} \cong 71.0\ dB$ which is slightly above the $A_{dc} = 70\ dB$ specification. The noise simulations on the other hand match the theoretical values very well prooving the negligible contributions of the cascode transistors and showing that the corner frequency was slightly below the specification. Because of the high current in the current sources M_{2a}-M_{2b}, it is difficult to minimize their contribution to the input-referred white noise resulting in a degraded thermal noise excess factor $\gamma_{n,ota} = 3.066$.

The folded cascode OTA is easier to design than the tlescopic OTA, because of the relaxed voltage constraints. It can reach higher DC gain and output voltage swing at the cost of more noise.

References

[1] T. C. Choi, R. T. Kaneshiro, R. W. Brodersen, P. R. Gray, W. B. Jett, and M. Wilcox, "High-frequency CMOS switched-capacitor filters for communications application," *IEEE Journal of Solid-State Circuits*, vol. 18, no. 6, pp. 652–664, 1983, doi: 10.1109/JSSC.1983.1052015.

[2] IHP, "IHP SG13G2 Open Source PDK." https://github.com/IHP-GmbH/IHP-Open-PDK, 2025.

[3] Holger Vogt, Giles Atkinson, Paolo Nenzi, "Ngspice User's Manual Version 43." https://ngspice.sourceforge.io/docs/ngspice-43-manual.pdf, 2024.

[4] G.D.J. Smit, A.J. Scholten, D.B.M. Klaassen, O. Rozeau, S. Martinie, T. Poiroux and J.C. Barbé, "PSP 103.6 - The PSP model is a joint development of CEA-Leti and NXP Semiconductors." https://www.cea.fr/cea-tech/leti/pspsupport/Documents/psp103p6_summary.pdf, 2017.

[5] F. Krummenacher, E. Vittoz, and M. Degrauwe, "Class AB CMOS amplifier micropower SC filters," *Electronics Letters*, vol. 17, no. 13, pp. 433–435, 1981.

[6] M. G. Degrauwe, J. Rijmenants, E. A. Vittoz, and H. J. D. Man, "Adaptive biasing CMOS amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 17, no. 3, pp. 522–528, 1982.

[7] SemiMod GmbH, "Open VAF." https://openvaf.semimod.de/docs/getting-started/introduction/, 2025.

[8] Dietmar Warning, "Verilog-A Models for Circuit Simulation." https://github.com/dwarning/VA-Models, 2024.

[9] Agilent Technologies, "IC-CAP 2008, Nonlinear Device Models - Volume 1." https://people.ece.ubc.ca/robertor/Links_files/Files/ICCAP-2008-doc/pdf/icmdl.pdf, 2008.

[10] C. C. Enz and E. A. Vittoz, *Charge-Based MOS Transistor Modeling - The EKV Model for Low-Power and RF IC Design*, 1st ed. John Wiley, 2006.

[11] C. Enz, F. Chicco, and A. Pezzotta, "Nanoscale MOSFET Modeling: Part 1: The Simplified EKV Model for the Design of Low-Power Analog Circuits," *IEEE Solid-State Circuits Magazine*, vol. 9, no. 3, pp. 26–35, 2017.

[12] C. Enz, F. Chicco, and A. Pezzotta, "Nanoscale MOSFET Modeling: Part 2: Using the Inversion Coefficient as the Primary Design Parameter," *IEEE Solid-State Circuits Magazine*, vol. 9, no. 4, pp. 73–81, 2017.