Francesco Sisti, BSc

Mobility-based oscillator with temperature compensation in SPT9 technology

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Electrical Engineering

submitted to

Graz University of Technology

Supervisor

TU Graz:
Dipl.-Ing. Dr. Peter Söser

Institute of Electronics
in cooperation with Infineon Technologies Austria AG, Villach

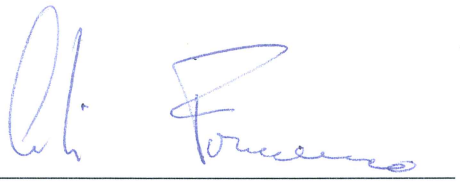
Infineon Technologies AG:
Dipl. Ing. David Åström

Graz, February 2018

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

14.02.2018
Date


Signature

Mobilitätsbasierender Relaxations-Oszillator mit Temperaturkompensation in SPT9

Oszillatoren finden immer Verwendung nicht nur in der Digital Domäne, sondern auch in der analogen Schaltungsentwicklung. Vollintegrierte anstatt Quartz-basierte Zeitreferenzen werden immer wichtiger in Bezug auf Reduzierung der Produktionskosten.

Basierend auf einer Veröffentlichung eines stromgesteuerten Relaxationsoszillators mit Ausnützung der Mobilität von Silizium [SBM+09] soll mit dieser Arbeit eine mögliche Implementierung in der SPT9 Technologie erarbeitet und das Verhalten des Oszillators in der verwendeten Technologie untersucht werden. Die zentrale Komponente dieser Zeitreferenz bildet die Referenzstromgenerierung und die Verwendung eines MOS Transistors als Referenzkapazität. Desweiteren wird das Temperaturverhalten der Frequenz untersucht und eine mögliche Kompensation präsentiert.

Schlüsselwörter: Oszillator, Zeitreferenz, MOS-Kapazität, Mobilität, BCD-Technologie

Mobility-based oscillator with temperature compensation in SPT9 technology

The need of oscillators for various applications is constantly increasing, not only in digital but also in analog domain for different reasons like error compensation with chopping or auto-zeroing, analog-to-digital conversion, network communication and many more.

Based on a concept study of a current controlled relaxation oscillator with the carrier moblility as reference [SBM+09], this thesis investigates in a possible implementation in the SPT9 technology and the circuit's behavior in this technology. The core part of this time reference is the generation of the reference current and the utilization of a MOS transistor as the reference capacitance. In addition, the frequency behavior of temperature is considered and a possible temperature compensation concept is presented.

Keywords: oscillator, time-reference, MOS-capacitor, mobility, BCD-technology

Danksagungen

An dieser Stelle möchte ich mich bei allen Personen bedanken, die zum erfolgreichen Abschluss dieser Arbeit beigetragen haben.

Besonders bedanken möchte ich mich bei bei folgenden Personen:

- meinem Betreuer David, der mich während der Durchführung meiner Arbeit stets tatkräftig unterstützt hat. Ich konnte aus den zahlreichen Diskussionen viel lernen.

- meinem Betreuer und Mentor seitens des Institutes für Elektronik der Technischen Universität Graz, Dr. Peter Söser für die unkomplizierte Betreuung und die Korrektur der Arbeit.

- Bei David, Daniel und Andreas für die vielen nützlichen "'Inputs"' und Diskussionen über diverse Themen.

- Hannes, für die praktische Einführung in die Welt der Layout-Erstellung.

- Arnaud, Uros, Niranjan und Marko vertretend für alle Kollegen der Dienststelle genannt. Danke an euch für die tolle Zusammenarbeit und das angenehme Arbeitsklima.

- Michaela, für den Rückhalt, den du mir gegeben hast und die motivierenden Gespräche und aufbauenden Worte während dieser Zeit.

Abschließend danke ich meinen Eltern für die großartige Unterstützung auf meinem bisherigen Lebensweg, seit der ersten Stunde an. Ohne euch würde ich nicht an jener Stelle stehen, an der ich heute angelangt bin.

Contents

List of Figures

1 Motivation

Nowadays clocking is present in almost every integrated circuit. The need of oscillators for various applications is constantly increasing, not only in digital but also in analog domain for different reasons like error compensation with chopping or auto-zeroing, analog-to-digital conversion, network communication and many more.

Since the first observation of electronic oscillations in the 19th century many different concepts of time references have evolved with each having different priorizations, whether it be accuracy, power dissipation, level of integration, jitter, or similiar.

Of course accuracy is an important consideration and an elementary criteria together with the power dissipation of the oscillator. It is commonly distinguished between low power consuming clock-generating circuits as part of a central function unit being responsible for power-down or sleep modes, in which the circuit's functionality like accuracy and speed is loosened but a very low current consumption is demanded since it is always on and in contrast, high accuracy oscillators for time sensitive and critical applications.
The reference value for accurate oscillators is still the quartz driven oscillator. Publications with new or improved concepts of time references compare their new approach with results achieved by a circuit with a quartz as the time-defining source. Thus, the quartz is an external component, which results in higher production cost due to processing of the quartz blank and the additional production step for combining the blank together with the electronic circuit in one package. This drawback drives development of fully integrated circuits forward.
A possible concept of an integrated oscillator is presented in [SBM+09] [SBM+11]. It is based on a principle of a current-controlled relaxation oscillator charging or discharging a capacitor, whereby the discharging current is referenced to the electron mobility. Altough the temperature dependence of the charge carrier mobility is large, it is well

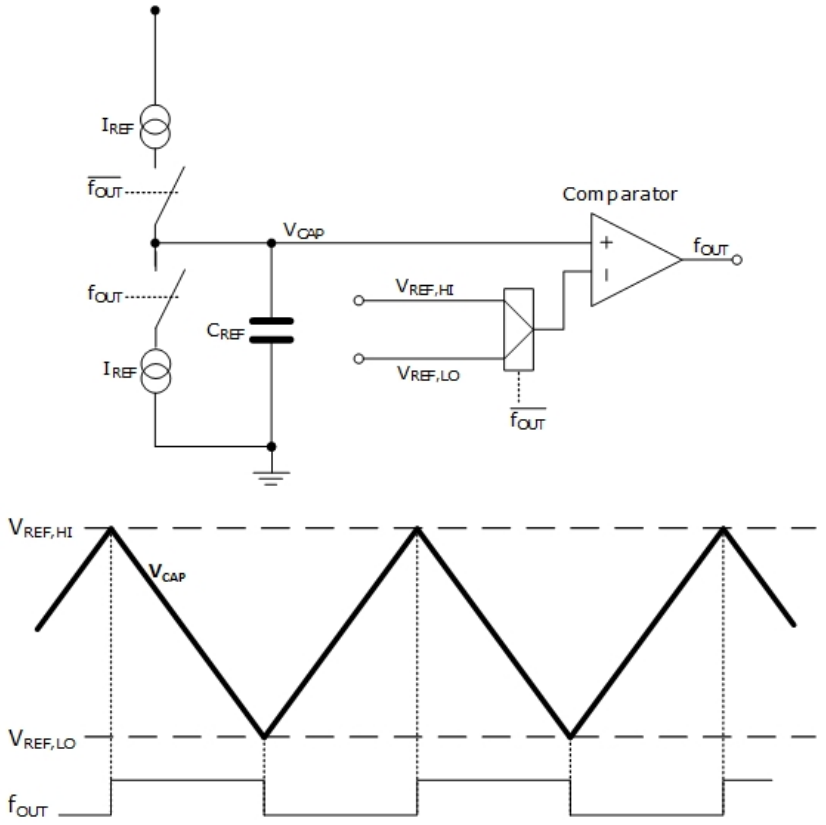


Figure 1.1: Principle of a current controlled relaxation oscillator

defined and can be compensated for. Furthermore the process variation has less impact on the mobility than on other parameters, such as polysilicon resistance for a pure RC oscillator, which does not only depend on the material, but is influenced on other not well defined parameters, or oxide capacitance in their 65nm CMOS technology.

The goal of the presented study is to investigate in the possibility of implementing similiar approaches for the SPT9 technology and if it can compete with already deployed time references.

1.0.1 Smart Power Technology

The circuit, which is introduced in this work, is implemented in a 130nm BCD (Bipolar-CMOS-DMOS) technology.

In the automotive sector, the BCD technology combines low-voltage and high-voltage circuit applications on a single chip, which is a so-called System on Chip (SoC). Every semiconductor company describes this technology differently and Infineon calls it Smart Power Technology (SPT).

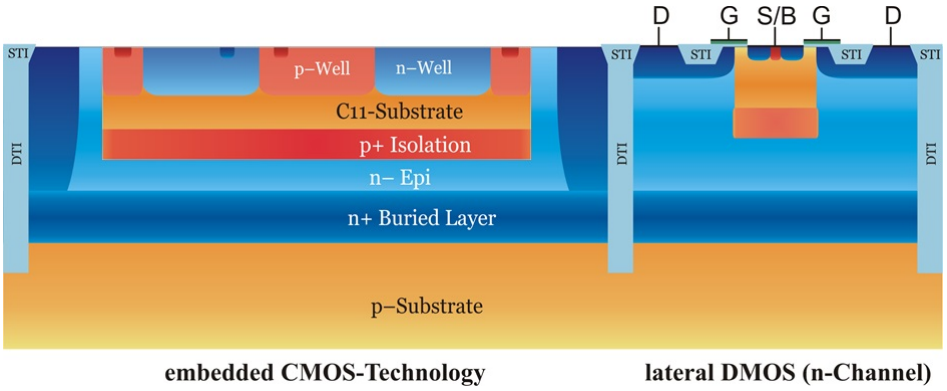


Figure 1.2: Cross section of a BCD technology [Wap07]

The basic structure of the used Smart Power Technology is depicted in figure 1.2. The actual deep-sub-micron core is implanted in a separate well, which forms the original substrate of the technology and is therefore named as "pseudo substrate". The enclosed p-well is surrounded by an n-epitaxie. The highly doped buried layer isolates the "pseudo-substrate" from the actual p-substrate of the wafer.

The horizontal isolation of the n-epitaxial wells, in the cross section depicted in figure 1.2 to separate the low-voltage from the high-voltage transistors, is guaranteed by deep trench isolations by means of etching a deep ditch of several μm and filling it with silicon dioxide. The advantage of this methodology is the space savings and better physical isolation compared to junction isolation. A principle challenge in BCD technologies is the suppression of parasitic structures.

On the basis of using the low-voltage domain with 1.5V nominal supply, just the pure CMOS part has to be taken into account. The technology provides a high integration density for low voltage transistors with a minimum channel length of 130nm. In general, the direction of development seeks to have more and more circuit concepts being mainly based on low-voltage transistors in order to remain competitive.

1.0.2 Assignment of Task

The goal of this work is a detailed analysis of a mobility-based reference circuit with the current reference generation as the core of this new concept in the SPT9 technology. In addition, a testchip circuit will be implemented to prove results based on simulations with measurements on silicon.
Due to the temperature dependence of the carrier mobility, which is the main influence on the oscillator's frequency, the temperature compensation is considered and a propsal for a compensation circuit is presented.

This technology is the basis for circuitry, which is employed in the automotive sector. Thus, relevenat characteristics for this work is

- temperature range $-40°C < \mathrm{T} < +175°C$
- supply voltage $V_{DD1v5} = 1.5\mathrm{V} \pm 10\%$

Most other papers regarding time reference circuits do not cover this temperature range. They are limited to the industrial, commercial or military temperature specificationo of operation. Hence, it is difficult to compare their achievements with this work.

2 Current Reference

2.1 Background

Numerous methodologies how to generate a reference current are already established and used today. The one observed and applied in this thesis is explained in many different books, papers and publications. One popular is [Raz01]. The basic concept is depicted in figure 2.1.

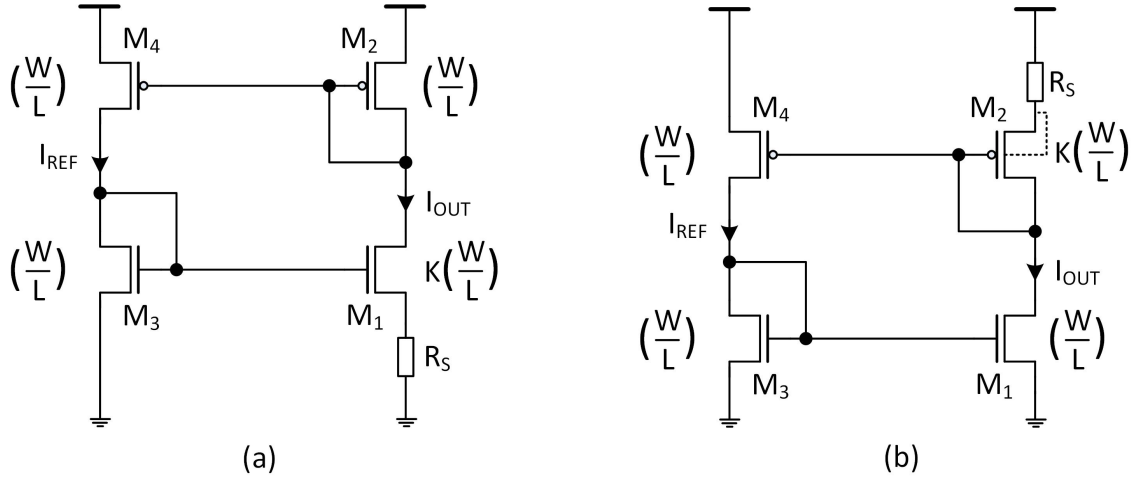


Figure 2.1: Basic concept of the current reference

This circuit generates a current I_{OUT} which is quite insensitive to V_{DD}. In figure 2.1 (a) assuming all transistors operate in saturation, the pmos current mirror which consists of M_2 and M_4 ensure $I_{OUT} = I_{REF}$. The actual current I_{OUT} is defined by

$$I_{OUT} = \frac{2}{\mu_n C_{OX} \left(\frac{W_1}{L_1}\right)_N} \frac{1}{R_S^2} \left(1 - \frac{1}{\sqrt{K}}\right)^2 \tag{2.1}$$

where

$$K = \frac{(W/L)_3}{(W/L)_1} \Big/ \frac{(W/L)_4}{(W/L)_2}$$

As shown in this equation the current is independent of the supply voltage and furthermore it leads to the assumption, that the current is independent of the threshold voltage on the basis of $V_{TH1} = V_{TH3}$, but as can be seen in figure 2.1 (a) the threshold voltages of the n-channel transistors are different due to the resistor at the source of M_1 and the resulting body effect and mismatching, which causes a deviation in the reference current. One possibility to circumvent this source of error is depicted in figure 2.1 (b) where the resistor is placed on top of the upper current mirror and the source and bulk of each p-channel transistor are tied. The same effect is achieved, if in figure 2.1 (a) M_1 and M_3 can be placed in isolated wells and source and bulk tied. This depends on the availability of the isolated wells in the technology.

A further improved circuit based on the same principle is described in [SOES88]. In this concept the resistor and the voltage drop over it is replaced by an ideal voltage source. In figure 2.2 (a) the voltage source is placed below the diode-connected n-channel transistor, which leads to a slightly different but similar current reference generation.

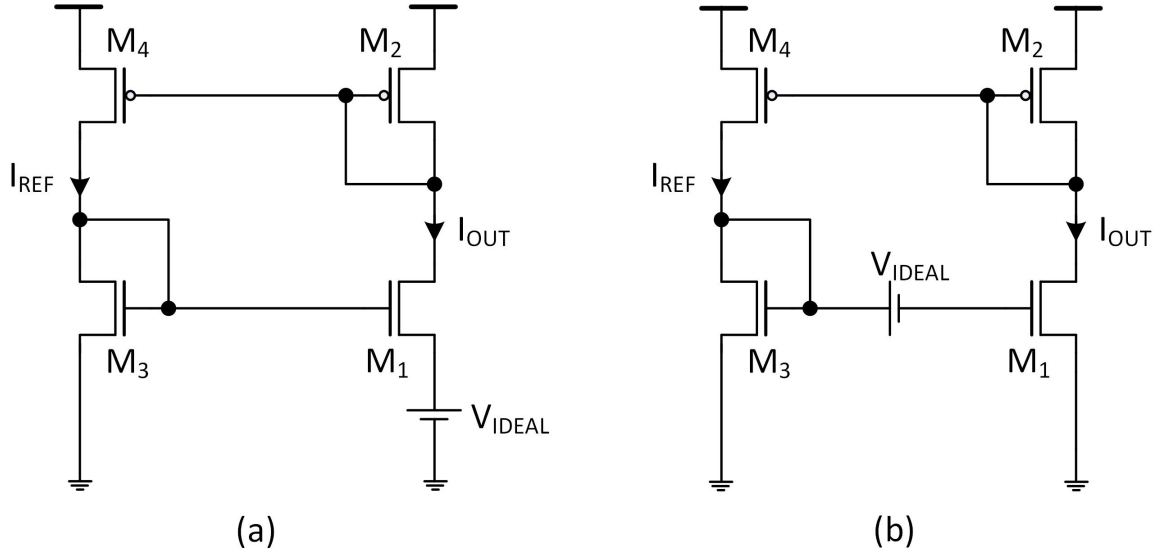


Figure 2.2: Principle to avoid the back gate effect

Taken the body effect into account and $V_{TH1} \neq V_{TH3}$ the current through M_1 is given by

$$I_{OUT} = \frac{\mu_n C_{OX}}{2} \frac{W_1}{L_1} \frac{(V_{IDEAL} + V_{TH2} - V_{TH1})^2}{\left(1 - \sqrt{K}\right)^2} \tag{2.2}$$

It becomes clear again, that supply voltage has no influence to the generated current.

The bulk effect's influence is eliminated in figure 2.1 (b) where the voltage source is placed floating at the gate of M_3. Neglecting local mismatch between the p- or the n-channel transistors neither the upper current mirror nor the lower current mirror introduce any error due to back-gate effect.

The fact that the transistors in each branch operate with different drain-source voltages, an error can be introduced due to the channel-length modulation. Channel-length modulation is a shortening of the length of the inverted channel region, what results in an increase in current

$$\frac{i_{OUT}}{i_{REF}} = \frac{1 + \lambda v_{DS2}}{1 + \lambda v_{DS4}}$$

with the channel-length modulation parameter

$$\lambda = \frac{\Delta L}{V_E L}$$

which depends on the early voltage V_E and the transistor's length L. This negative effect is more pronounced in short channel devices. Hence, relatively long channels has to be used to minimize the influence from $\Delta L/L$.

2.2 Actual Concept

Another improvement of this circuit is presented in [SBM+09] and [SBM+11], of which the current reference is depicted in figure 2.3.

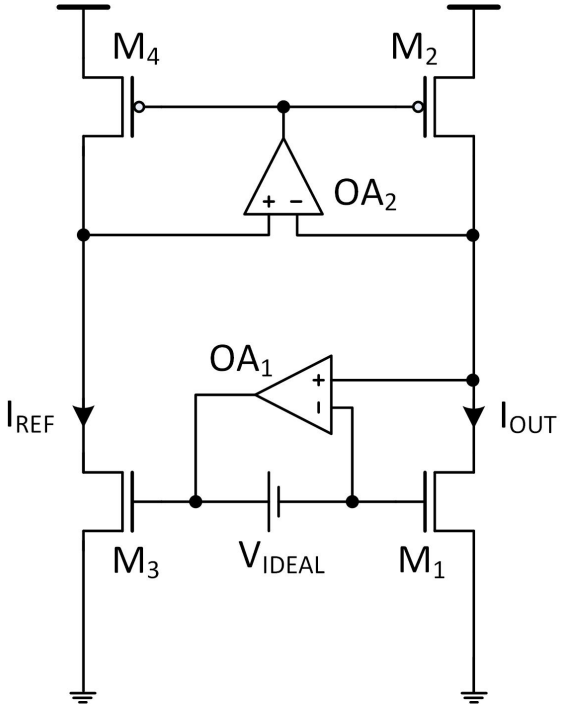


Figure 2.3: Basis for the concept study

In contrast to the previous current reference, this circuit uses an amplifier to operate the transistor M_1 as a diode and a second amplifier to bias the p-channel current mirror on top. Both of the mirroring transistors operate with the same drain-source-voltage and have the same gate-source voltage, which leads to a very good matching in case of channel-length modulation and threshold voltage as sources of error.

The principle of this circuit is used as basis in the present thesis to implement the current reference.

The starting point of the current's derivation in this concept starts from the following basic equation for the current

$$I_{D1} = \frac{\beta_{EFF1}}{2}(V_{GS1} - V_{TH1})^2 \tag{2.3}$$

The same formula used and rewritten to express the gate-source voltage for M_3 is

$$V_{GS3} = \sqrt{\frac{2I_{D3}}{\beta_{EFF3}}} + V_{TH3} \tag{2.4}$$

With Kirchhoff's loop rule V_{GS1} can be expressed as

$$\begin{aligned} V_{GS1} &= V_{GS3} - V_{IDEAL} \\ &= \sqrt{\frac{2I_{D3}}{\beta_{EFF3}}} + V_{TH3} - V_{IDEAL} \end{aligned}$$

and equation 2.4 is rewritten.

$$I_{D1} = \frac{\beta_{EFF1}}{2}\left(\sqrt{\frac{2I_{D3}}{\beta_{EFF3}}} + V_{TH3} - V_{IDEAL} - V_{TH1}\right)^2$$

Under assumption that V_{TH1}=V_{TH3} it can be simplified by

$$I_{D1} = \frac{\beta_{EFF1}}{2}\left(\sqrt{\frac{2I_{D3}}{\beta_{EFF3}}} - V_{IDEAL}\right)^2$$

The p-channel transistors forming the current mirror have the same length and width but differ only from the multiplicity factor. Therefore I_{D3} can be expressed as

$$\begin{aligned} I_{D3} &= \frac{\frac{W_4}{L_4}}{\frac{W_2}{L_2}} I_{D1} = \frac{n\frac{W_2}{L_2}}{\frac{W_2}{L_2}} I_{D1} \\ I_{D3} &= nI_{D1} \end{aligned}$$

and

$$I_{D1} = \frac{\beta_{EFF1}}{2}\left(\sqrt{\frac{2nI_{D1}}{\beta_{EFF3}}} - V_{IDEAL}\right)^2$$

$$\begin{aligned}\sqrt{I_{D1}} &= \sqrt{\frac{\beta_{EFF1}}{2}}\left(\sqrt{\frac{2nI_{D1}}{\beta_{EFF3}}} - V_{IDEAL}\right) \\ &= \sqrt{\frac{\beta_{EFF1}}{\beta_{EFF3}}}\sqrt{nI_{D1}} - \sqrt{\frac{\beta_{EFF1}}{2}}V_{IDEAL} \\ &= \sqrt{\frac{n}{m}}\sqrt{I_{D1}} - \sqrt{\frac{\beta_{EFF1}}{2}}V_{IDEAL}\end{aligned}$$

$$\sqrt{I_{D1}}\left(\sqrt{\frac{n}{m}} - 1\right) = \sqrt{\frac{\beta_{EFF1}}{2}}V_{IDEAL}$$

$$\sqrt{I_{D1}} = \frac{\sqrt{\frac{\beta_{EFF1}}{2}}V_{IDEAL}}{\left(\sqrt{\frac{n}{m}} - 1\right)}$$

$$I_{D1} = \frac{\frac{\mu_n C_{OX}}{2}\frac{W_1}{L_1}V_{IDEAL}^2}{\left(\sqrt{\frac{n}{m}} - 1\right)^2} \tag{2.5}$$

As shown in equation 2.7, the current is dependent on process of the oxide capacitance C_{OX} and the mobility μ_N, which is covered and discussed in detail.

2.3 Carrier Mobility

Mobility is closely linked to the carrier's average velocity, called drift velocity ν_d. For better understanding a piece of semiconductor is considered with no external field applied to it. Despite of the random motion of the holes and electrons in all directions due to thermal effects, they are canceled out on the average, which means there is no current flow. As soon as an electric field is applied to the semiconductor, a force is exerted on the charges and the movement along the field lines generate an electric current flow. This movement is called drift. The way of the charged particles through the silicon is not straight, but they interact with their environment, what is also called

scattering. In case of low electric fields, the proportionality between the applied electric field and the velocity, by which an electron travels through the silicon is called mobility. [WA12]

$$\nu_d = \mu_B \vec{\boldsymbol{E}}$$

For this simplified expression μ_B represents the mobility characterizing the bulk of the semiconductor, because it is only or mainly valid for low electric fields.

Figure 2.4 shows the mobility versus the doping concentration at room temperature. As can be seen the mobility of electrons is two to three times the mobility of holes, which explains the difference between n-channel and p-channel transistors regarding the higher drift velocity of electrons and the resulting faster n-channel devices.

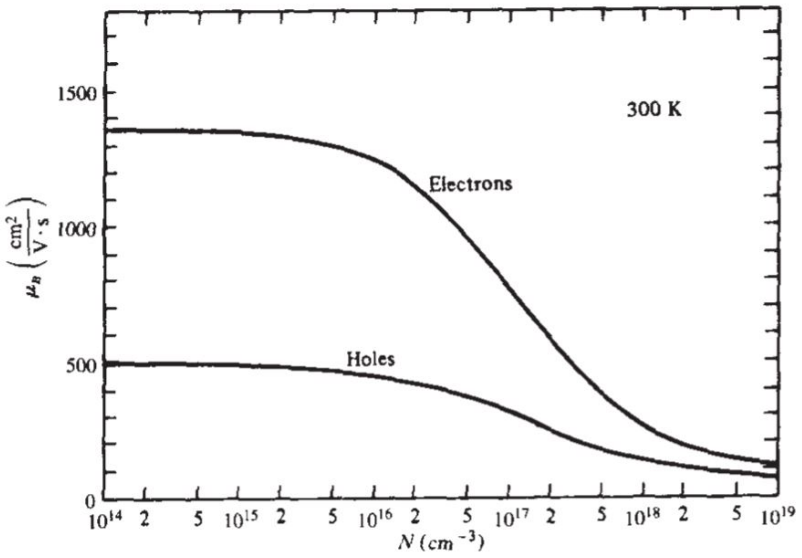


Figure 2.4: Electron and hole bulk mobility μ_B in silicon at 300K versus doping concentration [Tsi]

The bulk mobility is not the only scattering mechanism, which influences the drift velocity, but the effective mobility is defined by the following four main scattering parameters:

$$\frac{1}{\mu} \propto \frac{1}{\mu_{ac}} + \frac{1}{\mu_b} + \frac{1}{\mu_{sr}} + \frac{1}{\mu_{int}}$$

where μ_{ac} is the carrier mobility limited by scattering with surface acoustic phonons due to the energy of lattice vibrations, μ_b the bulk charge Coulombic scattering having its root cause in the electrically charged ionized impurity atoms and previously written as μ_B, μ_{sr} the mobility limited by the surface roughness scattering and μ_{int} the interface charge Coulombic scattering due to interface trapped charges and to charges trapped within the oxide. All these effects tend to lower the surface mobility, which describes the mobility of electrons in the inversion layer, to values smaller than the bulk mobility considered before. Moreover, each of them dominates in different ranges of transversale field values and are influenced by temperature, doping concentration, surface potential, terminal voltages and many more. Coming to the temperature be-

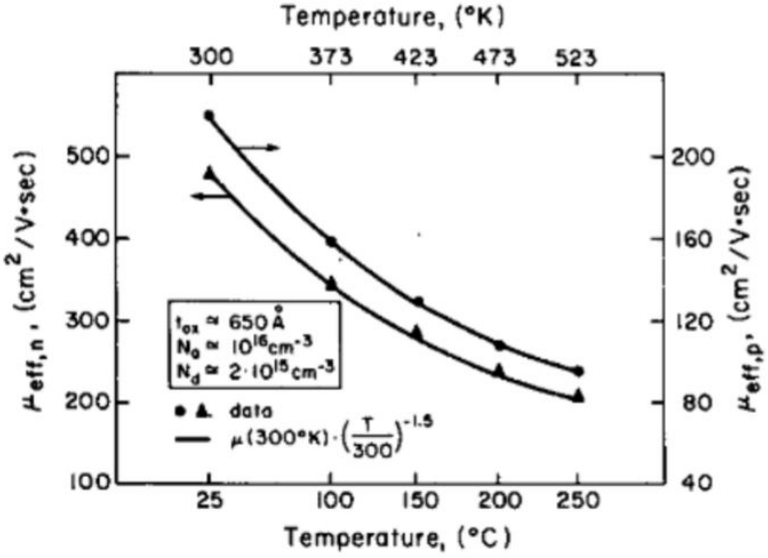


Figure 2.5: effective mobility μ_n for an n-channel transistor over temperature [URAC09]

havior, the effective mobility is known to decrease with temperature. An often used approximation is

$$\mu(T) = \mu(T_{rt}) \left(\frac{T}{T_{rt}} \right)^{-k_2}$$

with the absolute temperature T, the room absolute temperature T_{rt} and the constant factor k_2 with typical values between 1.2 and 2.5. This leads to expect a decreasing exponential function of the current with increasing temperature.

2.4 Oxide Capacitance

The oxide capacitance C_{OX} is the actual capacitance calculated out of the gate area multiplied with the oxide capacitance per unit area given by

$$C'_{OX} = \frac{\epsilon_{OX}}{t_{OX}}[\frac{F}{\mu m}]$$

where ϵ_{OX} is the insulator's permittivity and t_{OX} its thickness. It can be seen as a plate-plate capacitor, which are usually temperature stable. For better understanding a the unit capacitance of a transistor with an area of $1\mu m^2$ is depicted. Apparently, the oxide capacitance remains stable over temperature.

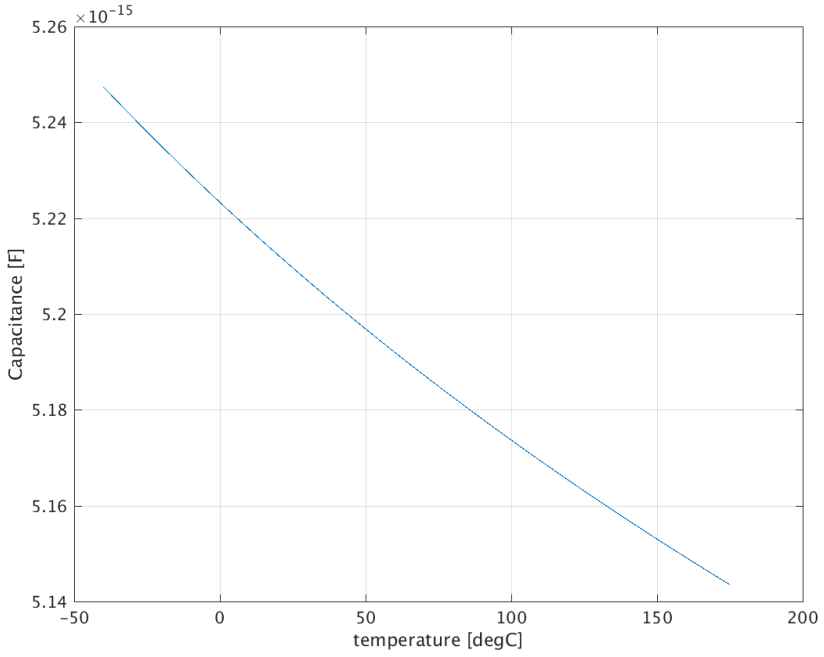


Figure 2.6: unit capacitance ($1\mu m \cdot 1\mu m$) over temperature

2.5 Design Considerations

The important detail for designing the reference is the temperature behavior of these process variables.

The automotive standard prescribes, inter alia, a wide temperature range from -40°C to 175°C, in which the circuit has to perform correctly. Therefore, all the circuits have to be designed in a way to guarantee the functionality for this specification. Regarding the current reference circuit, the temperature dependence of all nodes has to be taken into account to ensure all the transistors are still in the desired region.

Limiting introduced errors due to mismatch and static offset in the current mirror, it is important to operate both p-channel transistors in saturation. Hence, the gate-source voltage always needs to be

$$V_{GS,p} = V_{DS,p} + V_{TH,p}$$

whereby the drain-source voltage can be seen as equal for both p-channel transistors due to the implementation of the amplifier. This voltage can also be expressed as

$$V_{DS,p} = V_{DD} - V_{DS,n}$$

because the potential on this node is only defined by the diode-connected n-channel transistor M_1. Both the equation 2.3 and 2.7 are still valid for this device, from which the expected voltage level can be easy determined by equating them.

$$\begin{aligned} \frac{\mu_n C_{OX}}{2}\frac{W_1}{L_1}(V_{GS1} - V_{TH1})^2 &= \frac{\frac{\mu_n C_{OX}}{2}\frac{W_1}{L_1}V_{IDEAL}^2}{\left(\sqrt{\frac{n}{m}} - 1\right)^2} \\ (V_{GS1} - V_{TH1})^2 &= \frac{V_{IDEAL}^2}{\left(\sqrt{\frac{n}{m}} - 1\right)^2} \end{aligned} \tag{2.6}$$

A common gate-source extrapolated threshold voltage definition is

$$V_{TH}(T) = V_{FB} + \phi_0 + \gamma\sqrt{\gamma_0 + V_{SB}}$$

with the flatband voltage V_{FB}, which separates the accumulation regime from the depletion regimes and the surface potential ϕ_0 which sometimes is also written as $2\phi_0$

in literature. If the bulk and source are tied (V_{SB}=0) and $V_{FB} + \phi_0 = V_T(T_{rt})$ the threshold voltage can be simplified and expressed as [Tsi]

$$V_{TH}(T) = V_T(T_{rt}) - k_1(T - T_{rt})$$

where T is the absolute temperature, T_{rt} room absolute temperature and k_1 the temperature gradient, which is usually between 0.5mV/K and 3mV/K depending on the doping concentration, the oxide thickness and the biasing of the source-bulk voltage V_{SB}. V_{TH} is found to exhibit an almost straight-line decrease with temperature.

From a simulation with a DC sweep over temperature and plotting the operating point of a transistor and its threshold voltage confirms the expected behavior of the threshold voltage as can be seen in figure 2.7 . Also the linearity of the gradient of about 0.9mV/K can be seen out of the derivative in the plot.

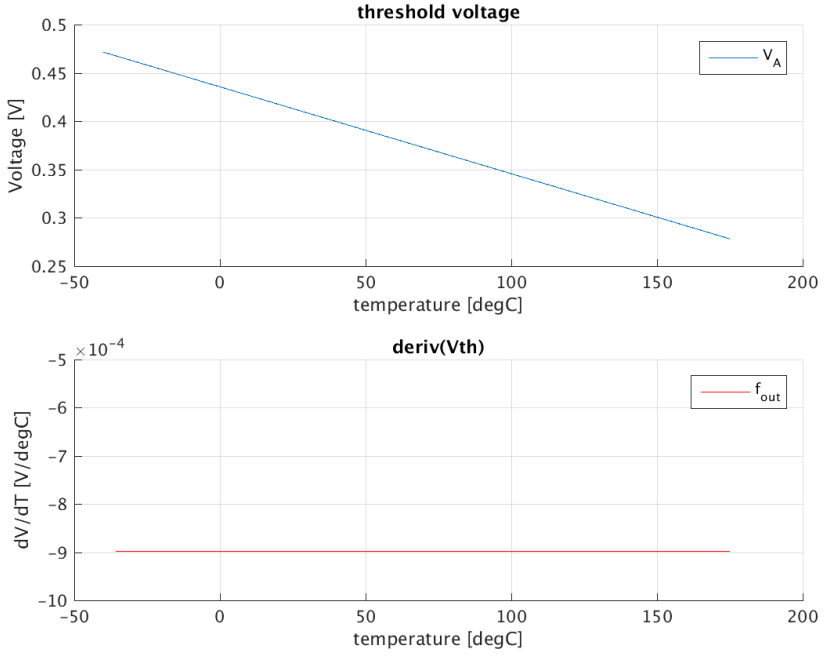


Figure 2.7: Vth temperature dependence for a 1um/4um n-mos transistor and its derivative

The right expression in equation 2.8 is constant over temperature and having a threshold voltage which is commonly decreasing linearly with increasing temperature, the

gate-source voltage has to decrease with the same behavior to remain the left expression constant as well.
Another important aspect is the current flowing through the current mirror. As defined in equation 2.7 the current depends on the mobility of the transistors M_1 or rather M_2. As a conclusion, at low temperature both the highest current is expected, which demands higher biasing voltage of the current mirror, and the highest voltage level at the diode-connected transistor M_1, or, respectively, the lowest drain-source voltage at the current mirror. This means, this operating point is the worst for the p-channel transistors in terms of remaining still in saturation region.
The next action for an initial design is the ideal voltage source and its influence. Regarding the transistor M_3, this voltage cannot be chosen arbitrary in respect to keep M_3 in saturation as well as M_1. It increases the gate-source voltage of M_3 compared to that of M_1, though, both transistors have the same drain-source voltage due to OA_2. With the assumptions of

$$\begin{aligned} V_{GS1} &= V_{DS1} \\ V_{DS1} &= V_{DS3} \\ V_{GS3} &= V_{GS1} + V_{IDEAL} = V_{DS1} + V_{IDEAL} = V_{DS3} + V_{IDEAL} \end{aligned}$$

the saturation condition of M_3 can be rewritten as

$$\begin{aligned} V_{GS3} - V_{TH3} &\leq V_{DS3} \\ V_{DS3} + V_{IDEAL} - V_{TH3} &\leq V_{DS3} \\ V_{IDEAL} &\leq V_{TH3} \end{aligned}$$

The ideal voltage source must not be higher than the threshold voltage of the transistor. Especially at hot temperature, where the threshold is expected to be the lowest, this restriction has to be taken into consideration. Though, depending the way in which this voltage drop is implemented and how stable it is over process and temperature, deviations on this voltage lead to a higher error on the current if the absolute value is very small. Thus, the tradeoff between remaining in saturation and not suffering from this drawback has to be taken.
Taking the two amplifiers into consideration, the common mode input voltage range of both the upper one, which provides the gate voltage of the current mirror's p-channel transistors, and the lower one, which aims to work the transistor M_1 as a diode, ideally

is the drain-source voltage of the diode-connected M_1 which decreases with increasing temperature. Hence, ensuring the input transistors always to be in strong inversion is necessary. OA_1's output voltage is $V_{OUT} = V_{IN,CM} + V_{IDEAL}$ and always ensures this requirement, though, OA_2 has to be dimensioned in the way the threshold voltage of the input transistors does not exceed the threshold voltage of the p-channel transistors M_2 and M_4, respectively.

$$\begin{aligned} V_{G5} &= V_{DD} - V_{DS2} \\ V_{D5} &= V_{DD} - V_{GS2} \end{aligned}$$

$$\begin{aligned} V_{G5} - V_{TH5} &\leq V_{D5} \\ V_{DD} - V_{DS2} - V_{TH5} &\leq V_{DD} - V_{GS2} \\ -V_{DS2} - V_{TH5} &\leq -V_{GS2} \\ V_{DS2} &\leq V_{GS2} - V_{TH5} \end{aligned}$$

To summarize the last paragraph:

- Designing the current mirror, the worst condition is expected to be at low temperature due to the low drain-source voltage of the p-channel transistor and the high drain-source voltage on the n-channel transistor, respectively, and also because of the highest current, which has to be provided and therefore a high gate-source voltage on the mirror.

- The voltage source shall not exceed the threshold voltage of the n-channel transistors in order not come into linear region, but should still be high enough to minimize deviation influences.

2.6 Voltage Source

As depicted in figure 2.8, the ideal voltage source is implemented as follows: a current source generates a voltage drop over a resistor, which is the voltage drop between the two gate-source voltages, whereby this current is provided by the amplifier OA_1. The difficulty of this concept is the unsymmetrical structure of OA_1 and the constant output current.

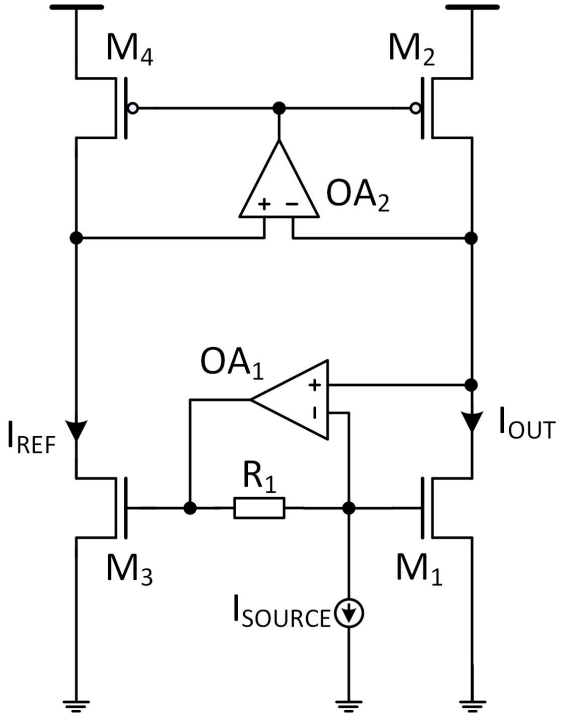


Figure 2.8: Realization of the voltage drop between the gates of M_1 and M_3

Thus, a slightly adopted version is taken into account and presented in the following chapter, in which the implementation is explained. The voltage drop over the resistor is implemented by two current sources and the transistor M_1 is directly diode-connected. Important to be noted is that these two current source have to match very precisely. The deviation between them is directly subtracted by or added on the current through M_1.

2.7 Implementation

2.7.1 N-channel Transistors

As previously defined, the current is generated with

$$I_{D1} = \frac{\mu_n C_{OX}}{2} \frac{W_1}{L_1} \frac{V_{IDEAL}^2}{\left(\sqrt{\frac{n}{m}} - 1\right)^2} \tag{2.7}$$

With a short width and a long length of the transistors the current is kept small, on one hand, to prevent unnecessary power consumption, and on the other hand to reduce the demanding circuitry of mirroring the generated to a very low discharging current to a minimum.

For the time being the multiplication factor between the two transistors is set

$$\frac{W_1}{L_1} = \frac{W_3}{L_3} << 1$$

In the low voltage domain of 1.5V nominal supply voltage, different n-channel transistors are available, whereby a crucial factor for the right implementation is the voltage difference between the two gates. Knowing that in the transistor M_1 the current is

$$I_1 = \frac{\beta_{EFF1}}{2} \left(V_{GS1} - V_{TH1}\right)^2$$

the gate-source voltage is

$$V_{GS1} = \sqrt{\frac{2I_1}{\beta_{EFF1}}} + V_{TH1}$$

Depending on the transistor's dimension, the voltage at this node can come close to the threshold voltage. As depicted in figure 2.8 the amplifier keeps both drain voltages on the same level ($V_X = V_Y$), thus, in order to remain the the transistor M_3 in saturation as well, its gate must not be higher than the gate voltage plus an additional threshold voltage.

$$\begin{aligned} V_{GS3} - V_{TH3} &\leq V_Y \\ V_X + V_{IDEAL} - V_{TH3} &\leq V_X \\ V_{IDEAL} &\leq V_{TH3} \end{aligned}$$

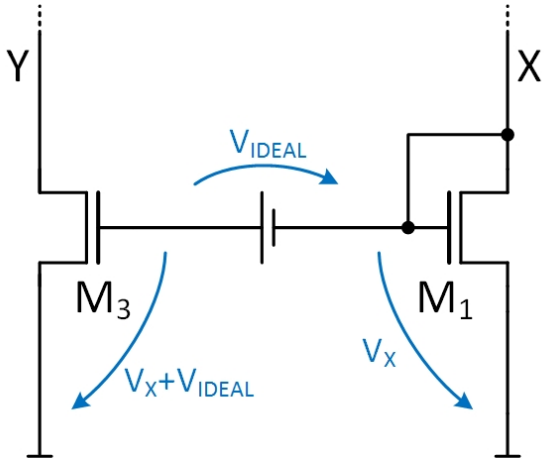


Figure 2.9: N-channel transistors of the current reference

Typically, a standard CMOS transistor with thin gate oxide and no further technological modifications like halo pockets or lightly doped drain for low voltage is thought for such circuitry in terms of matching. The drawback for this purpose is its low threshold voltage, especially fast process corners and high temperature lead to threshold levels below 100mV, what limits the operating range of the voltage drop between the gates. Furthermore the threshold voltage sets the level on the drain of M_1 and M_3 which defines the gate-source voltage of the input transistor in the amplifier, hence, it determines whether they are in weak or strong inversion.
A possibility to raise the threshold level is the implementation with transistors which are intended for the 2.5V supply domain. To ensure not only higher drain-source voltage but also gate strength, the gate oxide thickness is doubled compared to the 1.5V transistor.

2.7.2 Current Mirror

The current is expected to strongly change with temperature. Over the whole temperature range the current is provided from the current mirror, wherefore due to the voltage changes depicted in figure 2.10 two main critical conditions arise:

- The potential at node X and Y, respectively, increases to its maximum value at cold temperature and the drain-source voltage at the p-channel transistors reaches its minimum level. However, the current is the highest at low temperature as well what involves a high gate-source voltage to ensure the demanded current is

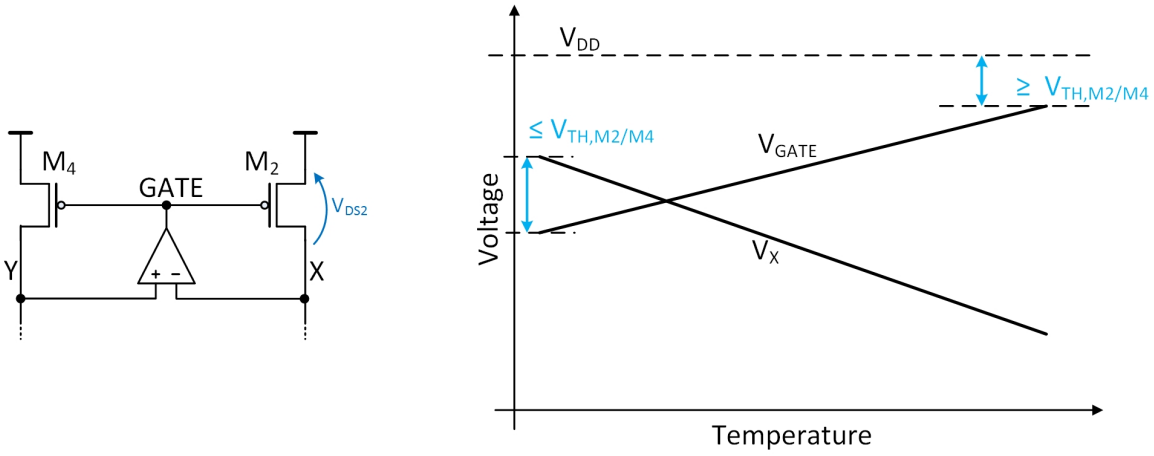


Figure 2.10: Amplifier biased p-channel current mirror and expected voltage over temperature

provided. Incorrect dimensions with a low W/L ratio of the transistors can result in high overdrive voltage for the demanded current and therefore could enter the triode region, what shall be avoided.

- The other critical condition is at high temperature, where due to low current and low threshold level the gate-source voltage is the lowest. The gate potential at the current mirror is the highest, which is biased by the amplifier. Thus, the amplifier must be able to bias the current mirror also in this corner condition. In addition, it is desired not to go into weak or moderate inversion in order to reduce deviations due to local mismatch of the transistors, which is more pronounced in this operating region, to a minimum.

Both M_2 and M_4 must be sized such that they are always in saturation and still can be biased correctly by the amplifier. The ratio between the two transistors is chosen with 2:1, which results in double the current in the left branch that in the right branch through the diode-connected transistor.

With the amplifier the drain-source voltage of both transistors are exactly the same except the amplifier's input offset voltage due to the finite gain. In contrast to a normal current mirror with the diode-connected master transistor and the different drain-source voltages, the accuracy between the two currents is increased in terms of channel length modulation. Although this definition is good for previous technologies and lambda λ is not constant in sub-micron technologies the following term gives a rough estimation about the influence of channel length modulation

$$\frac{i_{OUT}}{i_{REF}} = \frac{1 + \lambda v_{DS2}}{1 + \lambda v_{DS4}}$$

assuming that for both transistors the channel length modulation parameter is the same($L_4 = L_2$).

2.7.3 Amplifier

For a first implementation, a simple operational transconductance amplifier (OTA) is used to bias both the current mirror and the diode-connected transistor. With the whole circuitry and the temperature influencing voltages and currents, the first requirements to the OTA are both the input common mode range as well as the output voltage swing. This decision to use the same OTA for both amplifiers is made due to the fact, that both amplifiers have to deal with the same common-mode input voltage and similar output voltage swings. Slightly adoptions regarding stability is solved by different bias currents for the OTAs.

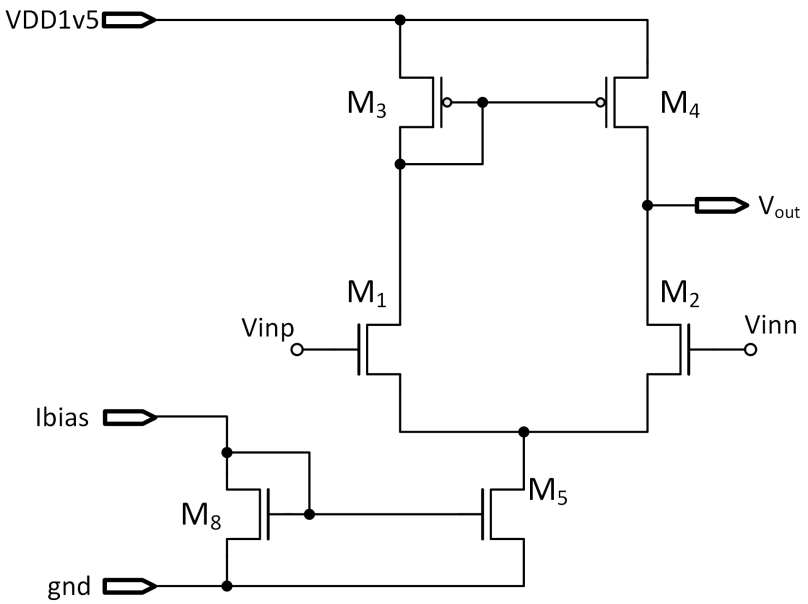


Figure 2.11: single stage operational transconductance amplifier

The design and sizing is determined by the same operating points as for the current mirroring transistors M_2 and M_4. Especially at hot temperature, the expected output voltage to bias the reference's current mirror is high (figure 2.10). For this operating

temperature, M_4 must still be in saturation in order not to cause mismatch in the bias current for the input transistors.
For both the common mode voltage range and the demanded output voltage, the design of the a amplifiers must meet the requirements over the whole temperature range.

2.7.4 Ideal voltage source

For the voltage drop between the two gates of M_1 and M_3, respectively, a resistor is placed and biased by a defined current. A possibility to bias the resistor is to provide the required current by the amplifier, which diode-connects the transistor M_1 and sink the current by an additional current mirror placed at the gate of M_1 as depicted in figure 2.12 (a). This method has two drawbacks, wherefore it is not used for the oscillator. On the one hand, the amplifier has be designed asymmetrical to provide the static bias current and the p-channel current mirror's ratio factor of the OTA together with the bias current of the OTA itself must match with the n-channel current mirror shown in figure 2.12 (a). Mismatch of these currents introduce additional offset in the OTA, what is not wanted. The second reason, why it is not used for the design, is the difficulty to change the bias current if the voltage drop has to be trimmed. Although it is possible to trim the voltage drop in defined steps by changing the current mirrors' ratios with additional switched transistors and digital trim signals, it is decided to use an alternative way to bias the resistor.

The generation of this defined current is intended to be variable for a better testability. Thus, a dedicated circuit, which is normally not included in the oscillator, is designed for the testchip. The amplifier does not have to provide the bias current anymore, but an additional p-mos current mirror is attached to the positive port of the resistor, like depicted in 2.12 (b). To match the currents, the n-channel mirror is also biased by the p-channel current mirror and both the driving and sinking mirror are implemented as cascoded current mirror.
The generation of the bias current is designed with a simple internal amplifier, which senses a voltage applied from external to the pin *Vbias* and keeps the same voltage drop over a resistor, by which the bias current is defined. This current then is mirrored to the resistor between the two gates of M_1 and M_3. The final solution, which is implemented, is depicted in figure 2.13.

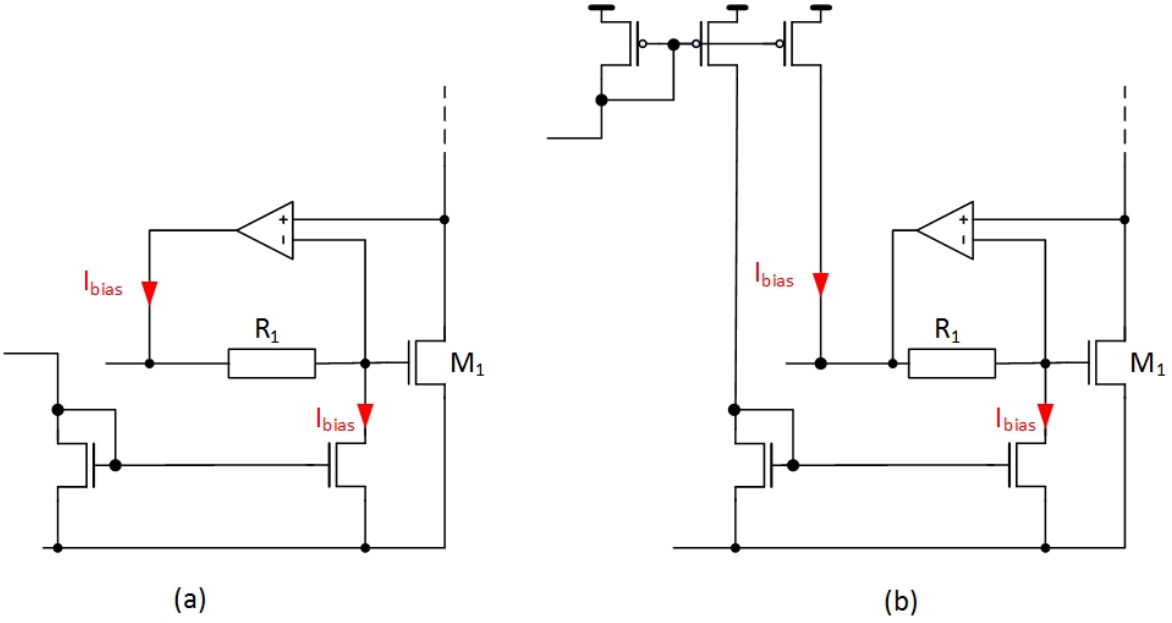


Figure 2.12: Bias concept of [SBM+09] (a) and proposed biasing of this work (b)

Of course, the resistance changes over temperature and thus, also the bias current. Despite, neglecting local mismatch and as long as the resistor R_2 changes with the same variation over temperature and process like R_1 does, the voltage drop over R_1 stays the same.

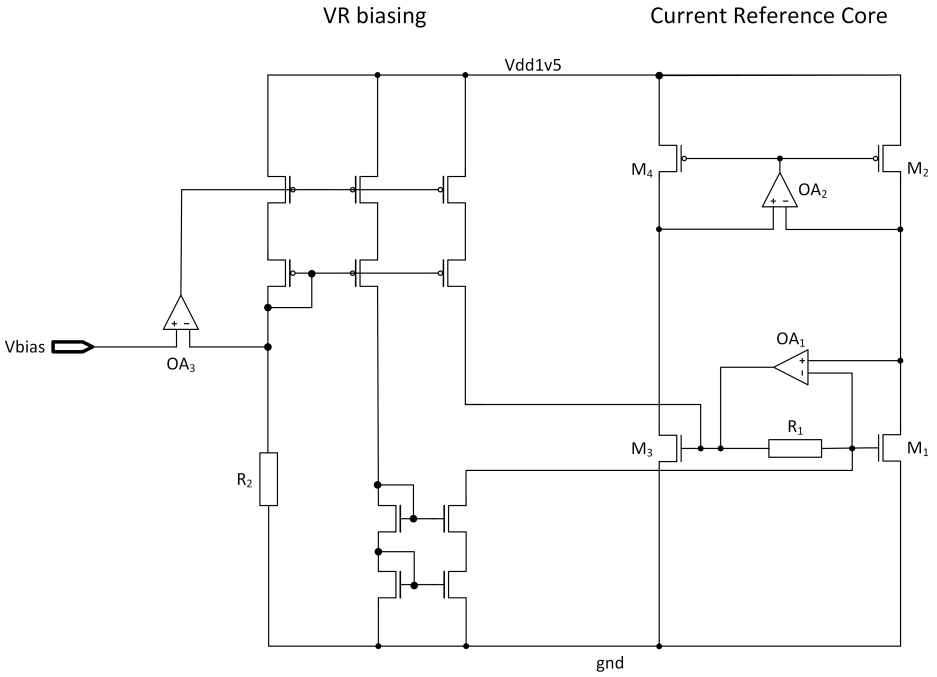


Figure 2.13: Implementation of the resistor biasing current

2.7.5 Power-down

A common feature in integrated circuits is the so-called power-down state, in which unnecessary circuit parts of the chip are turned off when they are not needed in order to minimize power consumption. Hence, the aim is to consume as litte current as possible. With the help of a digital power-down signal and transistors used as switches, certain nodes in the circuit are pulled up to the highest or down to the lowest possible voltage and cut off every branch in which current is flowing under normal conditions. The digital signal (enable) is logic high, when the circuit must be active and logic low to guarantee the circuit to be in power-down mode.

In the reference current circuit, first, the p-channel current mirror must be turned off by pulling the mirroring transistor's node to the supply voltage potential, which leads to $V_{gs,mirror} = 0V$ and $I_{DS} = 0A$ if the leakage current is neglected. The switch between V_{DD} and $V_{g,mirror}$ is implemented by a p-channel transistor and driven by the normal pwd_n signal.
Furthermore both nodes on the drain of M_1 and M_3 are pulled down to ground. In this case both the input voltages of the amplifier, which biases the p-channel current

mirror, is $V_{gs} = 0$ and also the positive input of the amplifier diode-connecting M_1 is set to zero. In order to have the negative input of this amplifier pulled down to ground as well, the gate of M_3 is pulled down and the current to bias the resistor R_1 is cut off, what leads to zero volt also at gate of M_1 and the second pin of the amplifier.

Regarding the amplifiers, they are kept in power-down state by pulling the OTA's p-channel current mirror to the supply rail. This step is only possible due to the previous mentioned strategy, that the input voltage for the amplifier is $V_{gs} = 0V$ and thus, the input transistors are turned off. If this would not be the case, current could flow from the supply voltage over the branch of the diode-connected transistor and the positive input transistor to the common source node of the input pair and further, since the low-voltage transistors are symmetrical and work the same when vertically flipped, over the negative transistor up to the output node. Especially for OA_1, who's output node is connected to the gate of M_3 and in power-down pulled down to ground, the current could rise to a significant level if the input voltages of the amplifier are not well defined.
In addition, the circuit, which provides the biasing for the amplifiers, suspends the current.

Regarding leakage currents, sensitive nodes, which has to be considered, are the drain of M_1 and M_3. As long as both transistors, which tie down these nodes in power-down mode, have the same size as well as the same drain-source voltage, the leakage during normal operation of the reference current is the same through both switches. The mirroring factor of the two branches is $I_{REF}/I_{OUT} = 2$. Thus the introduced error is half the leakage current $I_{LEAK}/2$ in I_{OUT}.

$$\frac{\Delta I}{I} = \frac{I_{leak}/2}{I_{OUT}}$$

For better understanding, an output current of $I_{OUT} = 2\mu A$ and the leakage through both transistors is $I_{LEAK} = 1nA$, the resulting error is $\Delta I/I = 0.00025 = 0.35‰$.

2.7.6 Startup Circuit

A critical condition, on which investigation is necessary, is the startup behavior of the circuit. Assuming the circuit is in power-down mode and then activated, it has to be guaranteed that the circuit has a well reliable and well defined startup independently of temperature and process variations. As soon as the current reference is triggered by the signal to become active, the current should rise and come to its stable operating point. At this time, both the drain voltages as well as the gate voltages of M_3 and M_1 are on ground potential due to the power-down switches. In addition the gate-source voltage of the current mirror on top is $V_{gs,mirror} = 0V$, what in total is a stable operating point of the circuit. Hence, without any additional circuitry the only possibility of a self-sufficient startup is by means of leakage currents, which lead nodes to drift to a voltage level and maybe turn on either one of the amplifiers or the current mirror.

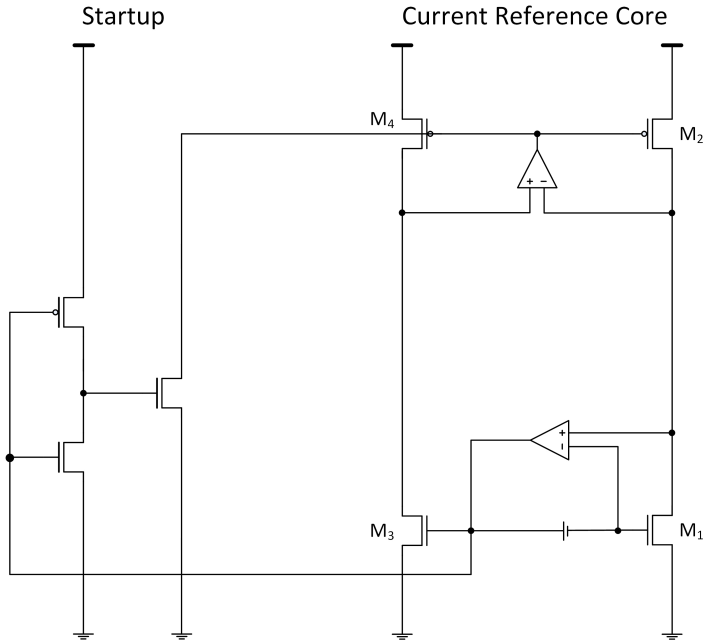


Figure 2.14: Startup circuit of the current reference

This non wanted stable operating point is avoided by the introduction of a startup circuit, which consists of an inverter and a switch. It senses the voltage at the gate of M_3 with the input of the inverter being connected to it and the output of this inverter drives a switch, which is placed between ground and the gate node of the p-channel current mirror transistors M_2 and M_4. As long as the voltage is too low at this node,

the switch is turned on and pulls down the gate of the current mirror and increases $V_{gs,mirror}$ and the current starts to flow. Thus, the voltage levels on the drain of M_3 and M_1 rise as well as the gate of them by means of the amplifiers.

The input of the inverter does not reach the supply's voltage level, what can lead to unwanted cross current. Therefore, the inverter is designed asymmetrical and the switching point is shifted to a lower level with a large n-channel transistor's W/L ratio and three stacked p-channel transistors with a very small ratio of the width and the length. The threshold voltage of the stacked p-channel transistors always has to be higher than the gate-source voltage of M_3 at the stable and active operating point in order not to have them always on.

$$|V_{thp,inv}| > V_{gs,M3}$$

The switch can be stacked to reduce leakage current when it is turned off.

2.7.7 Stability

The current reference consists of multiple control loops, which makes it rather difficult to determine the overall stability of the circuit. Therefore every loop is considered and checked separately. It is split up into the following three loops, which then are investigated:

- The negative feedback loop constituted by the amplifier OA_1 and M_1
- The negative feedback loop constituted by the amplifier OA_2 and M_4
- The positive feedback loop with the M_1, M_2 and the low-voltage current mirror

The first negative loop consists of the buffer configuration with OA_1 as a first stage, the resistor R_0 and M_1, which is the second stage as a common-source amplifier. Considering OTA's gain bandwidth calculated with the small signal model it can be written

$$A_{V,OTA} = \frac{G_M R_{OUT}}{1 + sR_{out}C_L}$$

with the output resistance $R_{OUT} = r_{ds2}||r_{ds4}$ and the transconductance $G_M = g_{m1,m2}$. Out of this equation the dominant pole is

$$\omega_{pd} = \frac{1}{R_{out}C_L}$$

On the output node of the OTA with the two transistors acting as a current source, the main contributors to the load capacitance are

$$C_L = C_{db2} + C_{db4} + C_{db\_bias,p}$$

and the output resistance is

$$R_{OUT} = r_{ds2}||r_{ds4}||r_{ds\_bias,p}||(r_{ds\_bias,n} + R_0)$$

Together with M_2 as a current source, the transistor M_1 is operated like a common-source stage with a gain of

$$\begin{aligned} |A_{V,CS}| &= \frac{g_{m1}}{g_{ds}} \\ &= g_{m1}(r_{ds1}||r_{ds2}) \\ &= 2\frac{I_D}{V_{GS1} - V_{T1}}(r_{ds1}||r_{ds2}) \end{aligned}$$

with the output resistance $r_{ds} = L/\lambda' I_D$ being proportional to the transistor's length and its current the gain is proportional as

$$|A_{V,CS}| \propto \frac{2I_D}{V_{GS1} - V_{T1}}\frac{L}{\lambda' I_D}$$

As the OTA and the second stage have no explicitly defined output capacitance, only the capacitances from the transistors are taken into account.
At nominal condition and ambient temperature, the OA_1 alone has a simulated DC gain of $A = 45dB$ and with a load capacitance coming only from the transistor's parasitics, which in that case is close to $C = 10fF$, the unity gain frequency is at $f_{UGF} = 310.42kHz$. The weak unity gain bandwidth derives from low biasing, through which the OTA is determined to set the dominant pole in this loop and to separate it

from the non-dominant pole of M_1.

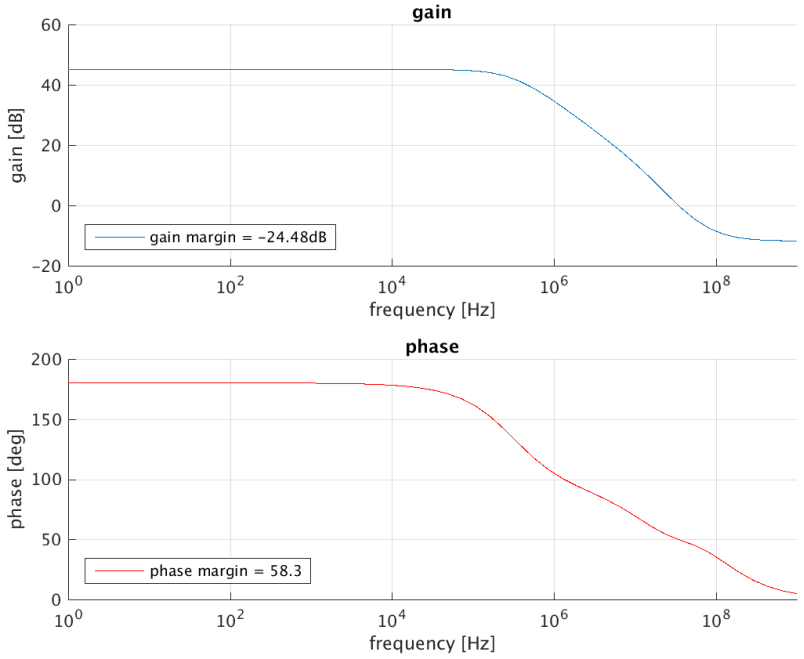
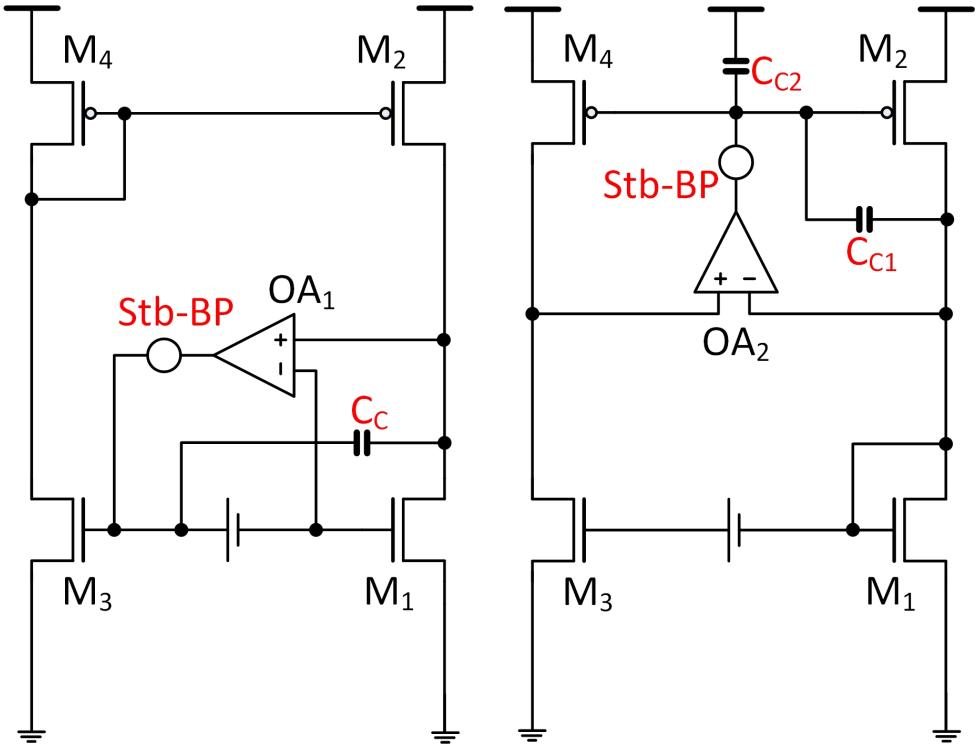


Figure 2.15: Bode Plot of the OTA

To circumvent the influence of that amplifier, which biases the current mirror, the amplifier is removed and the current mirror biased by diode-connecting the transistor M_4. For the analysis, the output of the OTA is chosen to break the loop and perform the stability analysis, depicted in figure 2.16a and described as stability breakpoint (Stb-BP).

Without any correction the phase margin of the loop is about $\phi_{PM} = 25°$. It is compensated to ensure a phase reserve of $\phi_{PM} > 60°$ by either adding a capacitance between the output of OA_1 and ground or by placing a Miller capacitance between the output of the OTA and the drain of M_1 being considered as the second stage output. The first possibility would result in a big capacitance occupying a lot of chip area, wherefore implementing the capacitor the other described way, the capacitance's value increases by the gain of the stage on each connected node due to the Miller effect. With a small capacitance of $C_C = 400fF$ a stable operation with a phase margin of $\phi_{PM} > 69.3°$ and a gain margin of $_{GM} > 69.°$for nominal process, supply and ambient temperature is guaranteed. For all process variations, a supply voltage of $V_{DD1v5} = 1.5 \pm 10\%$ and over the whole temperature range, the phase margin does not drop below $\phi_{PM} > 63.5°C$.



(a) stability analysis OA_1 (b) stability analysis OA_2

Figure 2.16: Modified current reference for the stability analysis

The second control loop relates to the upper current mirror in combination with the amplifier OA_2. In order to eliminate the other feedback loops, the current reference circuit is simplified with having M_1 diode-connected directly and the OA_1 removed. The circuit is depicted in figure 2.16b.

The starting point here is similar to that for the lower circuit part. As the OTA is biased weakly it is supposed to set the dominant pole and be separated from the other pole determined by the transistors. The approach is again to either place a capacitance between the supply node and the output of the OTA or inserting an additional C_{GD} at M_2 being amplified and separating the two poles.

Again, a capacitor to supply must be rather large for compensation, nevertheless, in this case it is beneficial in terms of suppressing disturbances from the supply, which is the well-known power supply rejection ratio (PSRR), and to keep the gate-source voltage of the mirroring transistors constant. Hence, a combination of two capacitances, a large one from V_{DD} to the output of the OTA and another rather small as C_{GD} of M_2. It has to be paid attention to choose their size properly and not to counter compensate

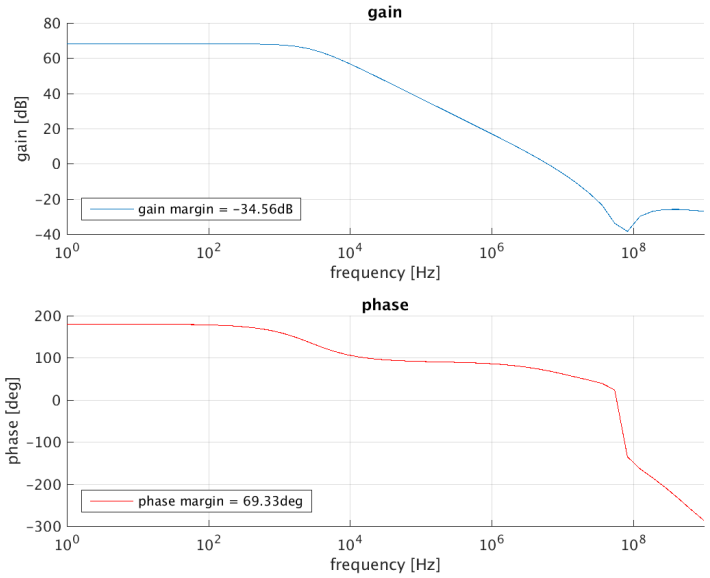


Figure 2.17: Bode Plot for stability analysis with OA_1

one with the other.
With the capacitors $C_{C1} = 100fF$ and $C_{C2} = 6pF$ a stable operation is achieved, what can be seen in the bode plot shown in figure 2.18.

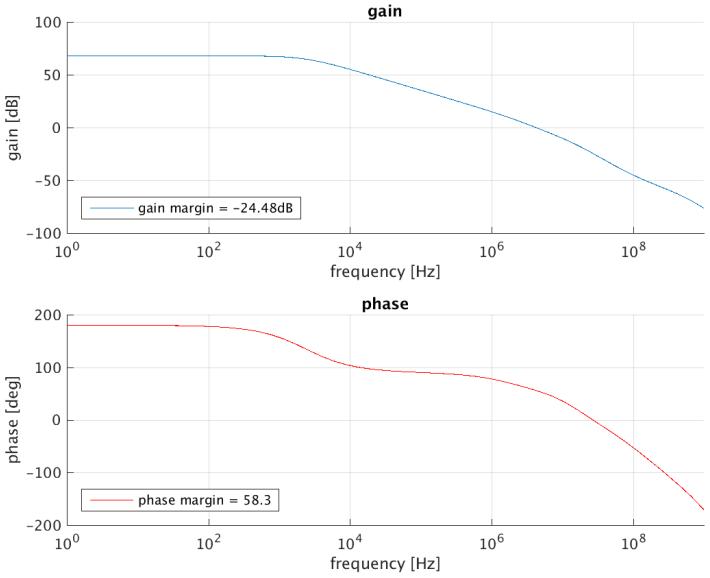


Figure 2.18: Bode Plot of stability analysis with OA_2

2.7.8 Schematics

The complete circuit for the current reference is depicted in figure 2.19.

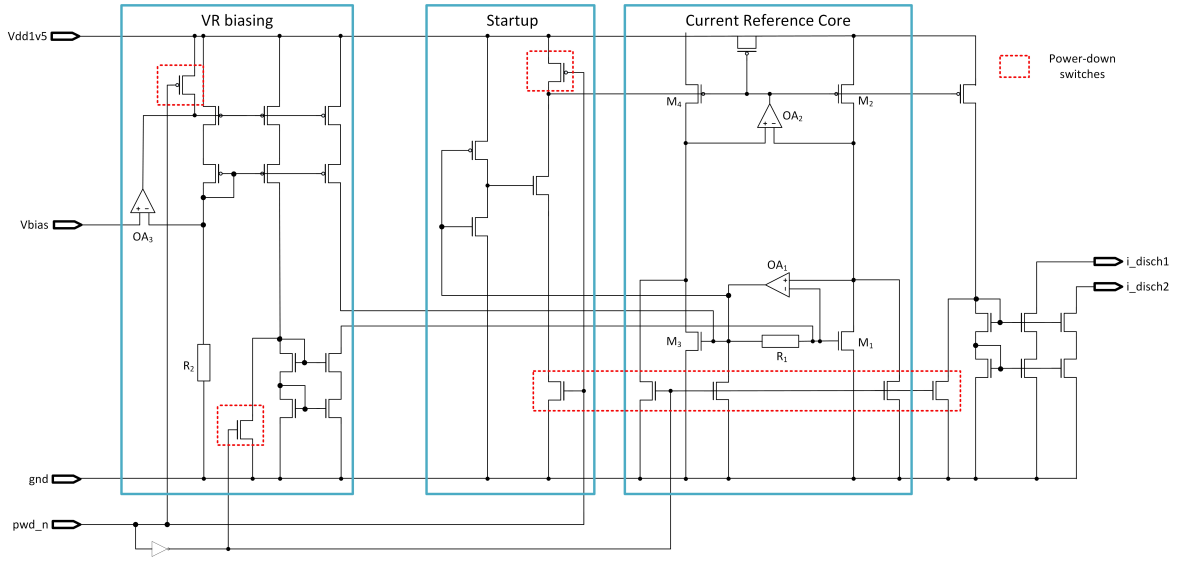


Figure 2.19: Schematic of the implemented current reference

The current to discharge the capacitors is provided by an additional current mirror, with which the current is scaled down. In addition, the current mirror avoids coupling of digital noise during switching on the core circuit of the current reference. The nominal value of the discharge current at ambient temperature is $I_{disch} = 200nA$.

2.7.9 Simulation

The generated current's behavior over temperature is illustrated in the following figure 2.20 with

- $V_{DD1v5} = 1.5V$
- process corner = nominal
- $V_{bias} = 650mV \rightarrow V_{R_1=162.5mV}$

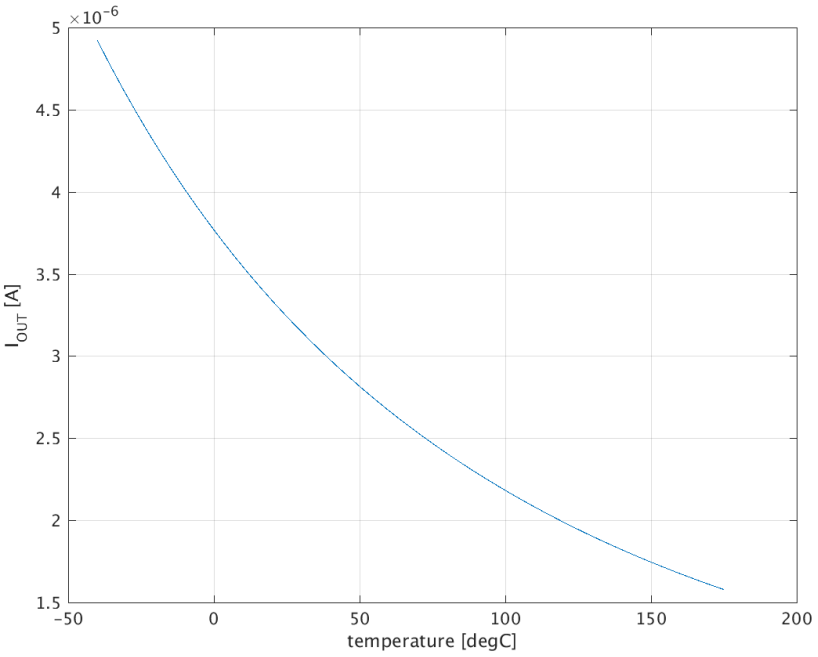


Figure 2.20: nominal current I_{OUT} over temperature

The influence of the process variation to the generated current is depicted in figure 2.21. The relative spread of the current is

$$I_{OUT} = I_{OUT}(T) \pm 20\%$$

With a single point trimming of the voltage drop V_{R_1} - in this case at a temperature of T=-40°C - the relative spread of the current is decreased to less than

$$I_{OUT} = I_{OUT}(T) \pm 0.7\%$$

at hot temperature, where the highest deviation is expected, if it is trimmed at cold temperature. This confirms, that altough the current has a strong temperature dependence, but is very well defined and can be compensated for.

As can be seen from the waveform, the current decreases from $I_{-40°C} = 4.92\mu A$ to $I_{175°C} = 1.578\mu A$, which is a factor of around 3. The ratio W/L of M_1 and M_3 is chosen small in order to keep the current in the two branches low and the multiplication factor between the two transistors is 1:1.

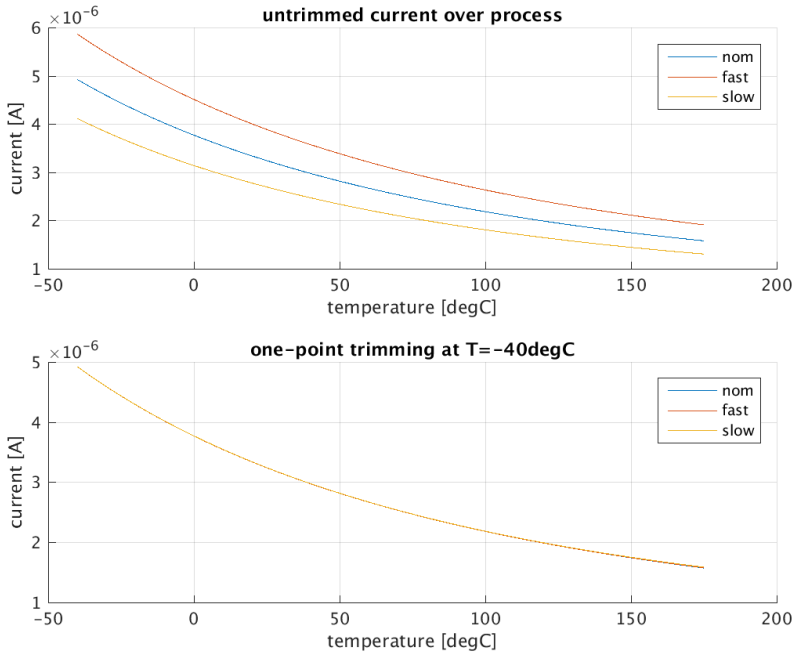


Figure 2.21: Current for different process corners and without and with one-point trimming

The current shall only be influenced by the mobility temperature dependence which is like previously mentioned.

$$\mu(T) = \mu(T_{rt}) \left(\frac{T}{T_{rt}} \right)^{-k_2}$$

A first coarse curve fitting based on this equation and a simulated current value of $I_{d1} = 4.7\mu A$ at room temperature leads to the following curve

$$I_{d1} = 3.16\mu A \left(\frac{T(^\circ K)}{300K} \right)^{-1.74}$$

Attached to the previous waveform it gives a rather good approximation of the actual behaviour of the current.

In addition, saving the effective mobility μ_{eff} from every operating point of the transistor M_1 and displayed together with the current, the temperature dependence principally only from the mobility is confirmed.

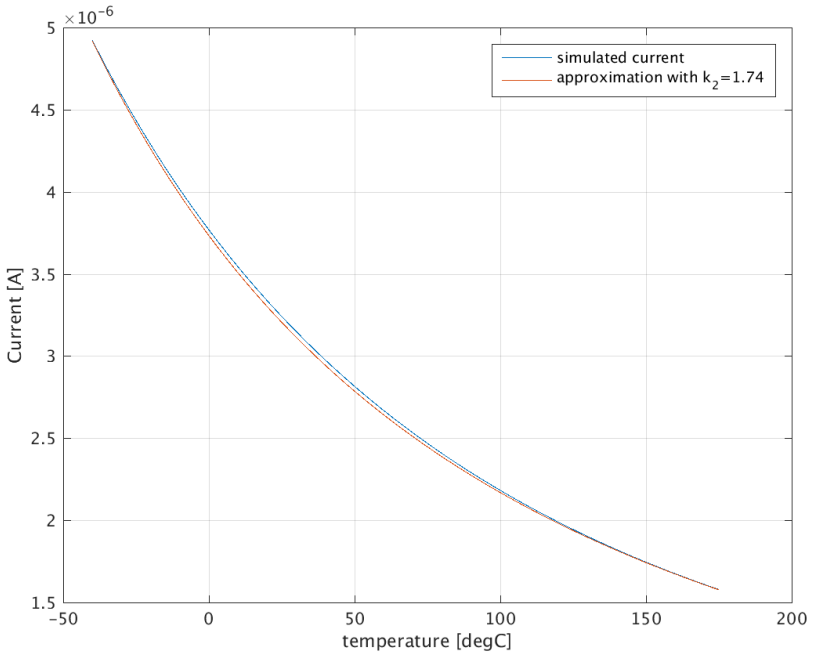


Figure 2.22: I_{OUT} with process variation and one-point trimming

One aspect with this circuit configuration, which is already mentioned before, is the stable overdrive voltage for the diode-connected M_1 in order only to be dependent on the carrier mobility and the oxide capacitance.

$$\begin{aligned} \frac{\mu_n C_{OX}}{2}\frac{W_1}{L_1}(V_{GS1}-V_{TH1})^2 &= \frac{\frac{\mu_n C_{OX}}{2}\frac{W_1}{L_1}V_{IDEAL}^2}{\left(\sqrt{\frac{n}{m}}-1\right)^2} \\ (V_{GS1}-V_{TH1})^2 &= \frac{V_{IDEAL}^2}{\left(\sqrt{\frac{n}{m}}-1\right)^2} \end{aligned}$$

Plotting the voltage drop over this transistor and comparing it with the threshold voltage, it becomes clear, that $V_{ov} = V_{gs} - V_{th}$ is stable as shown in figure 2.23.

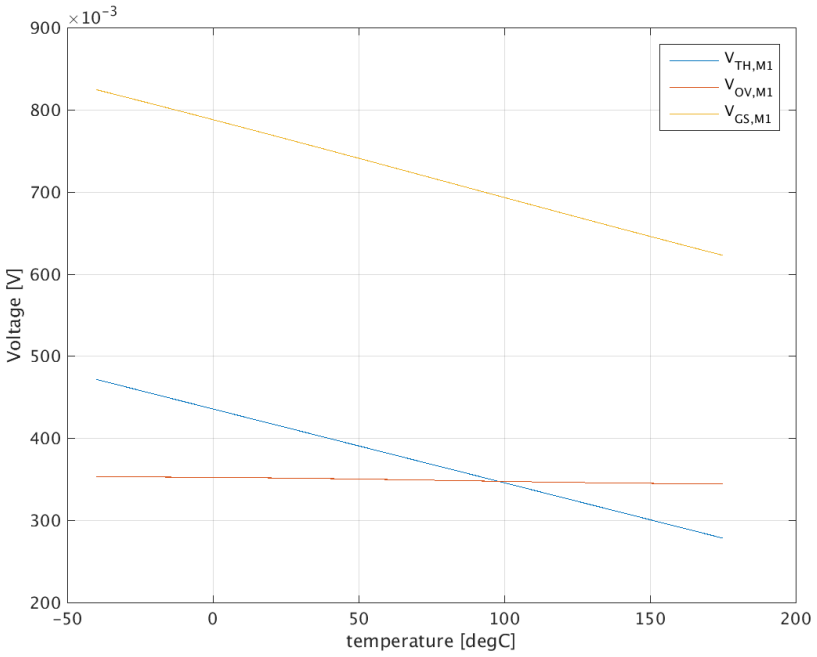


Figure 2.23: Stable overdrive voltage at the transistor M_1

To prove the benefit of the implemented amplifier's biased current mirror, a simple comparison is made between this concept and a standard current mirror regarding the ratio error between the reference and the mirrored current. The lower part of the reference is implemented the same for both of the circuits. For the DC simulation, only a nominal corner is used.

With a mirroring factor of $I_{REF}/I_{OUT} = 2$, the curves depicted in figure 2.24 show the error between the two currents in the left and the right branch of the current reference, which is defined as following

$$f = \frac{I_{REF}}{2 \cdot I_{OUT}}$$

The blue line represents the error of the approach proposed in this concept, whereas the red line shows the deviation in the standard current mirror. For the same transistor sizes, the difference is reduced from 18‰ to less than 0.1‰.

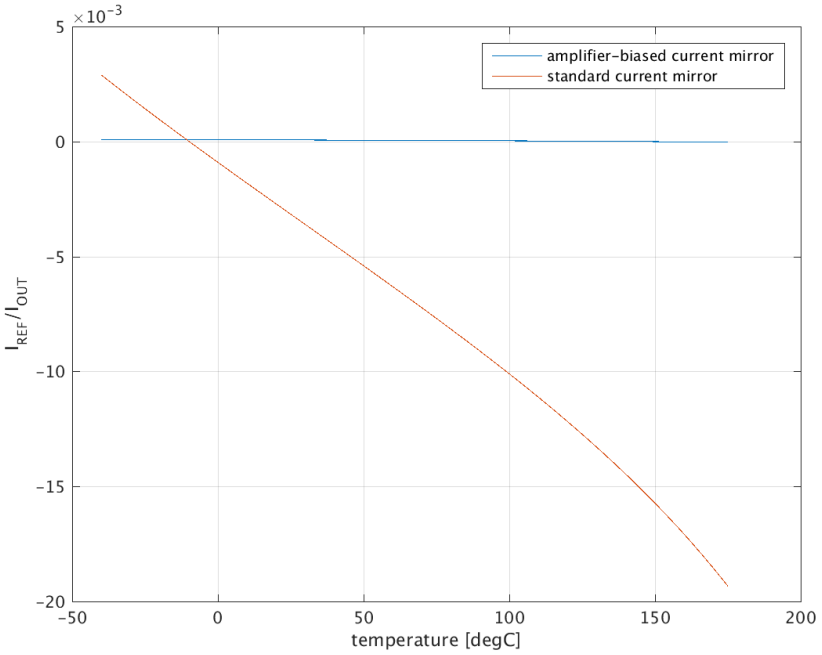


Figure 2.24: Mirroring error in the amplifier-biased and standard current mirror

3 MOS Capacitor

In general, integrated circuit technologies offer many different integrated capacitances being based on the principle of parallel plates with the simple relationship

$$C = \epsilon \frac{A}{d}$$

Depending on the field of application they reveal certain benefits and drawbacks. Besides the common metal-metal and poly-poly capacitors, there is also the possibility of using a MOS transistor as a capacitor for the implementation. Frequently they are deployed as stabilization for circuits like current mirrors with the same size like the mirroring transistors and placed in the common centroid layout instead of having additional dummy transistors.
The reason why this type of capacitor is used for this work, is the dependence of the capacitance's value on the oxide capacitance C_{ox}. The combination of the mobility-based reference current with the MOS transistor as capacitor cancels out process variation of C_{ox} and does not influence the generated frequency. A closer investigation on this type of capacitor is presented in this chapter.

As depicted in figure 3.1 (a) the gate of the MOS capacitor is the metal electrode. A MOS transistor is just a MOS capacitor with two PN junctions flanking the capacitor (figure 3.1 (b)).

The difference to other integrated capacitors is the capacity value as a function of the applied voltage between the body and the gate whereby a distinction between three major sectors can be drawn:

- surface accumulation

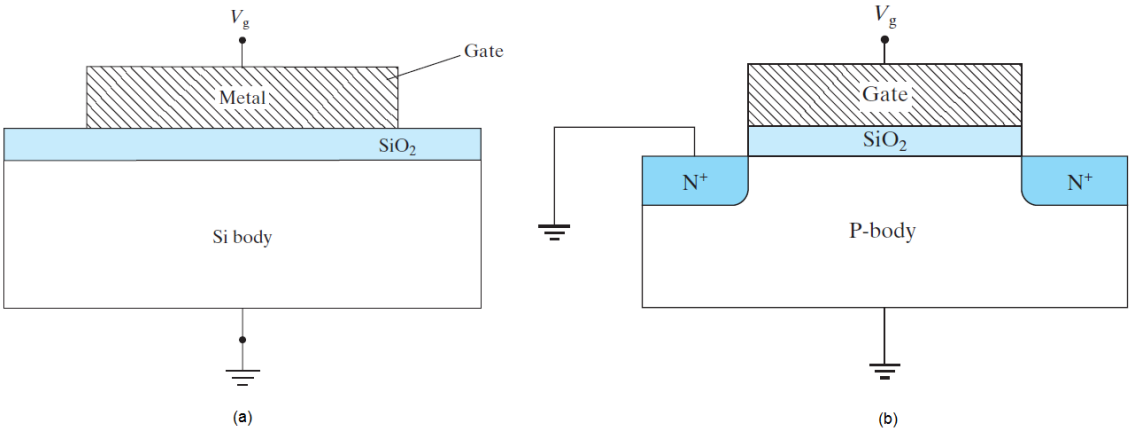


Figure 3.1: pure MOS capacitor and MOS transistor as capacitor [Hu09]

- surface depletion
- strong inversion

3.0.10 Surface Accumulation

For the following considerations an n-channel transistor will be used. Applying a more negative voltage than the flat-band voltage V_{fb}, the surface hole concentration, p_s, is larger than the bulk hole concentration, p_0, so the large number of holes attracted at or near the surface form an accumulation layer with the accumulation charge Q_{acc}.

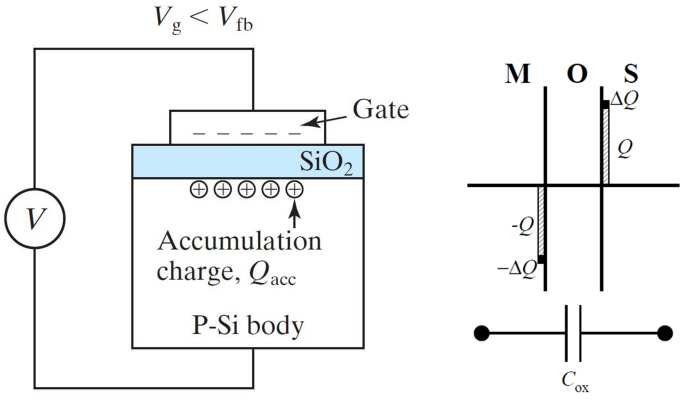


Figure 3.2: surface accumulation in a mos capacitor [Hu09] [unk03]

Neglecting the influence of the surface potential in first order, defining the voltage

V_{ox} = V_g - V_{fb} over the oxide thickness T_{ox} and using Gauss' Law it can be written

$$V_{ox} = \vec{\boldsymbol{E}} \cdot T_{ox} = -\frac{Q_{acc}}{\epsilon_{ox}} T_{ox} = -\frac{Q_{acc}}{C_{ox}}$$

This is the usual relationship between the voltage and the capacitance except the negative sign, which is explainable by the choice of the gate as the electrode, but the charges are in the substrate. With the previous definition of V_{ox} the charge per unit is a function of the oxide capacitance per unit and the gate voltage with an offset of the flat band voltage:

$$Q_{acc} = -C_{ox}(V_g - V_{fb})$$

what also explains, that if $V_g = V_{fb}$, the accumulation charge is zero.

The total capacitance for the accumulation region is

$$C = \frac{dQ_{acc}}{dV_g} = C_{ox}$$

3.0.11 Surface Depletion

If the applied between gate and body is more positive than the flat-band voltage, both the electron and hole densities are small and this leads to a depletion region at the surface. As the gate voltage is varied, the incremental charge is added or subtracted at a certain depth W_d in the substrate instead of near or at the surface. This forms the depletion width and an additional factor has to be included in calculation. This factor can be seen as an additional capacitance C_{dep} in series, which diminishes the overall capacitance value.

The capacitance consists of

$$\frac{1}{C} = \frac{dQ_{dep}}{dV_g}^{-1} = \frac{1}{C_{ox}} + \frac{1}{C_{dep}} = \frac{1}{C_{ox}} + \frac{W_d}{\epsilon_{Si}}$$

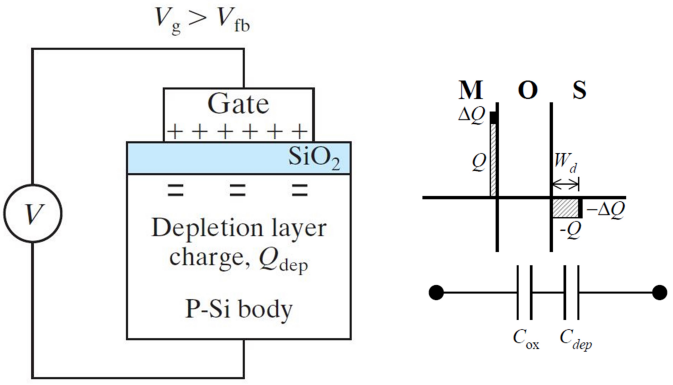


Figure 3.3: surface depletion in a mos capacitor [Hu09] [unk03]

3.0.12 Strong Inversion

Raising the voltage V_g beyond the threshold voltage, the width W_{dep} of the depletion region does not increase further but the surface has now a higher electron concentration than the bulk doping concentration resulting in an N-type surface. The inversion charge per unit Q_{inv} is

$$Q_{inv} = -C_{ox}(V_g - V_t)$$

with the known voltages and the oxide capacitance per unit.

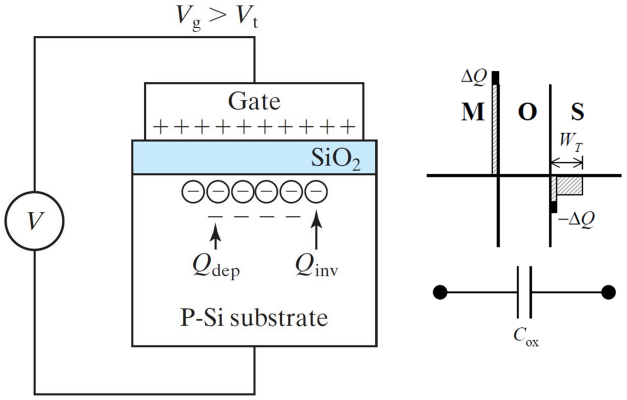


Figure 3.4: strong inversion in a mos capacitor [Hu09] [unk03]

The capacitance again is

$$C = \frac{dQ_{inv}}{dV_g} = C_{ox}$$

3.0.13 MOS capacitors vs MOS transistor

For fast charging and discharging the capacitance in inversion at higher frequencies, electrons always have to be present to react on the applied voltage to build the inversion layer which can be seen as the bottom electrode. In a pure MOS capacitor with a p-substrate there is a lack of free electrons and thus, inversion layer charges cannot be supplied or removed quickly enough to respond on a rapidly changing voltage. In contrast, the additional N regions in a MOS transistor provide the required charges. Therefore, the transistor's C-V characteristic behave the same both for DC and higher frequencies, whereas in a pure MOS capacitor it has to be distinguished between static and HF.

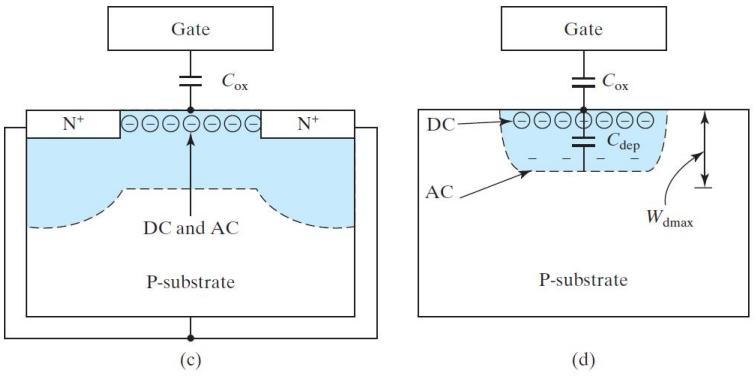


Figure 3.5: Influence of n-regions to DC and AC behavior on a MOS capacitor [Hu09]

For better visibility the capacitance as a function of the gate voltage is shown in figure 3.6.

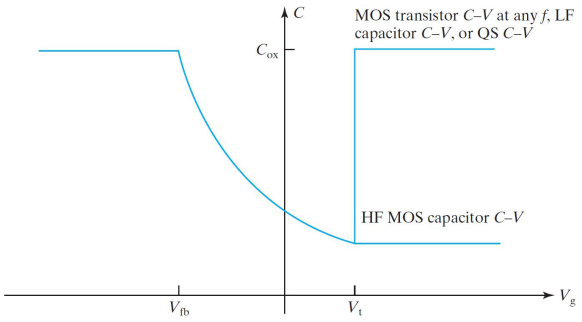


Figure 3.6: The mos capacitance versus the applied voltage V_g [Hu09]

3.0.14 Parasitic Capacitances

So far, the thickness of the inversion layer is assumed to be infinitely thin. However, the width of this layer is a function of V_g and the electric field and it becomes thinner with higher voltage resulting in an additional capacitance in series with the oxide capacitance. Another factor, which influences the overall capacity is the electric field over the poly-silicon at the gate forming a depletion width and therefore a further capacitor in series, whose impact is dominant at higher biasing voltages. Both the inversion width and the poly-silicon width together with the thickness of the actual gate oxide is difficult to differentiate. Hence, an effective thickness T_{oxe} is introduced which is

$$T_{oxe} = T_{ox} + \frac{W_{dpoly}}{3} + \frac{T_{inv}}{3}$$

where the divison by 3 is the ratio between ϵ_s/ϵ_{ox}. Especially for thin gate oxides the poly-silicon thickness and the inversion layer thickness T_{inv} in the silicon are not negligible anymore.

Moreover the transitions between the states are not that clear as depicted in figure 3.6. Considering the region around the threshold voltage, the capacitance will rise smoothly towards C_{ox} because the inversion layer evolves not exactly at the interface between the silicon and the oxide, but is located at some depth that varies with V_g. The different regions overlap what impedes to set the limits. A simulation with a DC sweep over the applied voltage is shown in figure 3.7.

3.1 MOS capacitor with mobility-based current reference

The actual advantage of this type of capacitor for this type of oscillator, which utilizes a defined current to charge a capacitor as a time constant, is the influence of the oxide capacitance C_{OX} on the capacitance value. As can be seen in equation 2.7, the generated current $I_{OUT} \propto C_{ox}$ is directly proportional to the oxide capacitance. The frequency of the oscillator is a function of the capacitance, the voltage levels and the

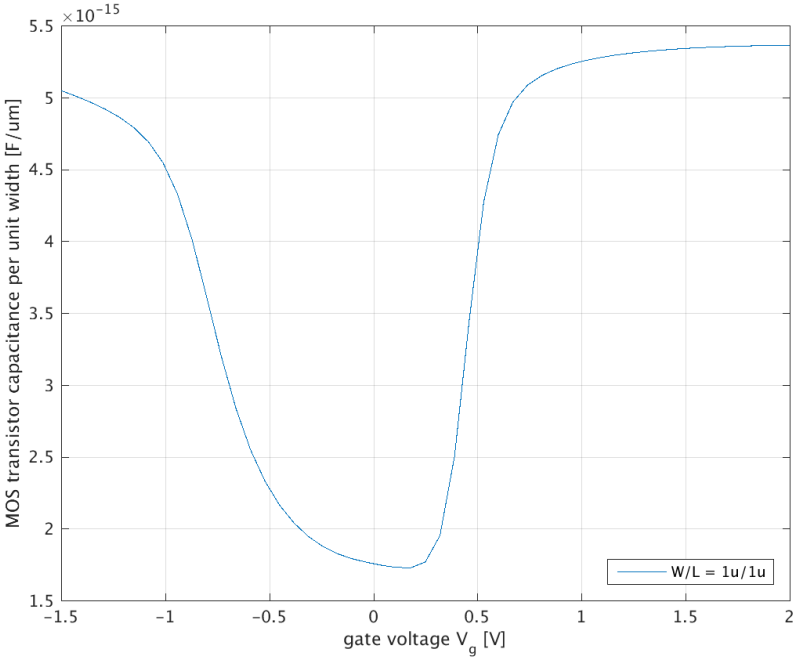


Figure 3.7: A simulation of the mos capacitance versus the applied voltage V_g

discharge current with

$$\begin{aligned} T_{osc} = \frac{1}{f_{osc}} &\propto \frac{C}{I} \\ &\propto \frac{2C_{ox}}{\mu_n C_{ox}} \\ &\propto \frac{2}{\mu_n} \end{aligned}$$

Changes in the generated current with $I_{OUT} \propto C_{ox}$ due to process variations in the oxide thickness or the permittivity also effects the MOS transistor's capacitance, which is affected in the same way ideally and the influence of process and temperature deviations from the oxide capacitance is canceled out.

3.1.1 Operating region

In principle, the capacitor should always be biased in a region in which it is stable. Due to the complexity of providing a negative voltage for the accumulation regime, the targeted region is inversion ($V_{DD1v5} > V_g > V_{min}$). The operating range, in which the

value of the capacitance is stable, is experimentally found with simulations by doing a DC sweep of the voltage applied on the gate as the positive plate and having drain, source and bulk connected to ground. The value is determined by the summation of all gate-related capacitances

$$C_G = C_{GS} + C_{GD} + C_{GB}$$

Compared with a calculation of the input capacitance of a MOS transistor in triode region ($V_{th} < V_{GS}$, $0 < V_{DS} < V_{GS} - V_{th}$), the individual terms can be expressed as [KK02].

$$\begin{aligned} C_{GD} &= \frac{1}{2}WL \cdot C_{ox} + W \cdot (C_{GD,ov} + C_{GD,fr}) \\ C_{GS} &= \frac{1}{2}WL \cdot C_{ox} + W \cdot (C_{GS,ov} + C_{GS,fr}) \end{aligned}$$

with the gate oxide capacitance C_{ox} and the overlap and fringe capacitances $C_{XX,ov}$ and $C_{XX,fr}$ respectively. These can be seen in figure 3.8 with the cross section of a transistor and all the involved capacitors drawn.

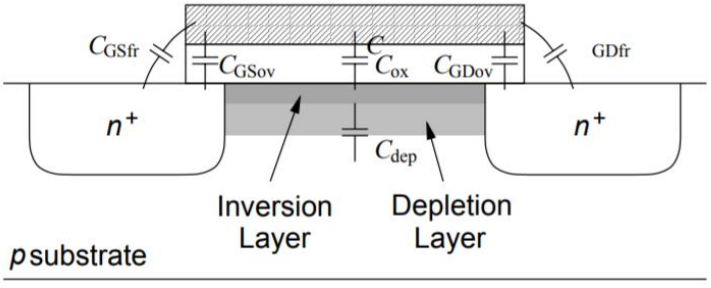


Figure 3.8: Cross section of a mos transistor with the most important input capacitances [KK02]

Resulting from the simulation in figure 3.7, the maximum voltage limit to which the MOS capacitor is discharged is set to $V_{g,min} = 0.9V$.

3.2 Dimensioning

As depicted in figure 3.7, the unit capacitance is $C_{ox} = 5.25fF$. The voltage limitations of $V_{g,min} = 0.9V$ and $V_{g,max} = V_{DD1v5} = 1.5V$ set the range in which the capacitor is discharged.
With the previous defined discharge current of $I_{disch}(T = 27°C) = 200nA$, the calculation of the demanded capacitance for a frequency of $f_{out} = 100kHz$ is

$$f_{out} = \frac{I_{REF}}{2C_{REF}} \frac{1}{V_{REF,HI} - V_{REF,LO}}$$

Changing the equation to calculate the capacitance results in

$$\begin{aligned} C_{ref} &= \frac{200nA}{2 \cdot 100kHz} \frac{1}{1.5V - 0.9V} \\ &= 1.66pF \end{aligned}$$

With a unit capacitance of $C_{ox} = 5.25fF$, the overall transistor area is

$$\begin{aligned} W_{C_{ref}} \cdot L_{C_{ref}} &= \frac{C_{ref}}{C_{ox}} \\ &= \frac{1.66pF}{5.25fF} \\ &= 317\mu m^2 \end{aligned}$$

4 Toplevel

The focus of this work is the current reference in combination with the n-channel transistor as a capacitor. Nevertheless, it will be placed on a testchip to compare the results of the simulations with real measurements on silicon, wherefore a fully functioning oscillator with certain restrictions is designed. These restrictions concern the instances in terms of not being efficient regarding the power consumption, but they are designed in order not to have an impact or influence on the testchip results. The mentioned instances include the comparator and the reference voltages.

The toplevel concept is depicted in figure 4.1 [SBM+09] [SBM+11]. The digital circuitry, which drives the signals for the switches, is not included.

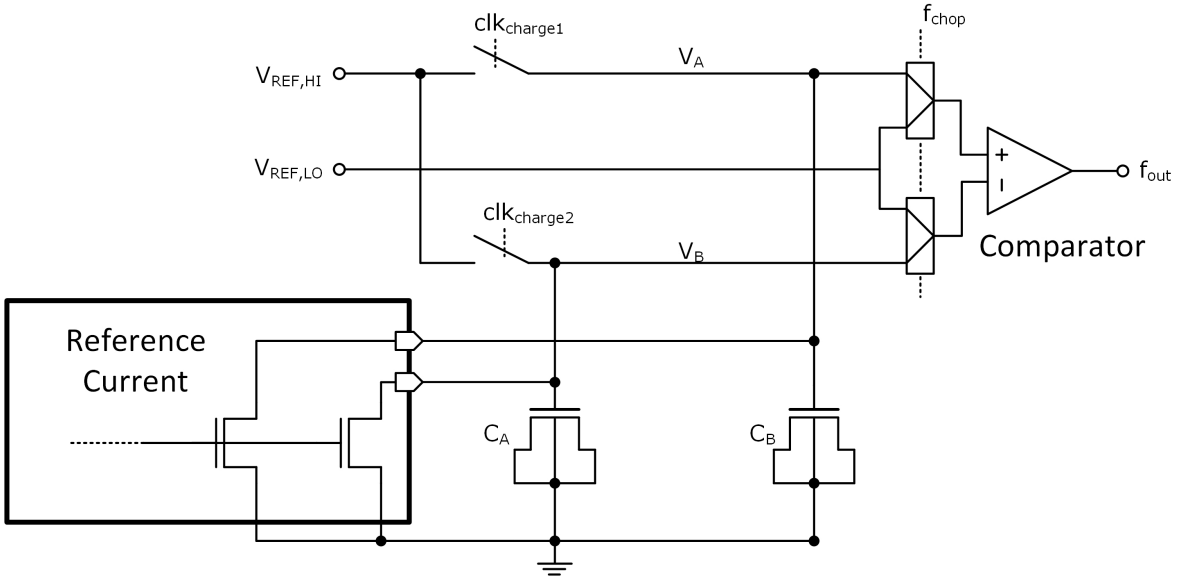


Figure 4.1: Concept circuit of the proposed implementation [SBM+09]

For better understanding, a timing diagram of the signals, which drive the switches and

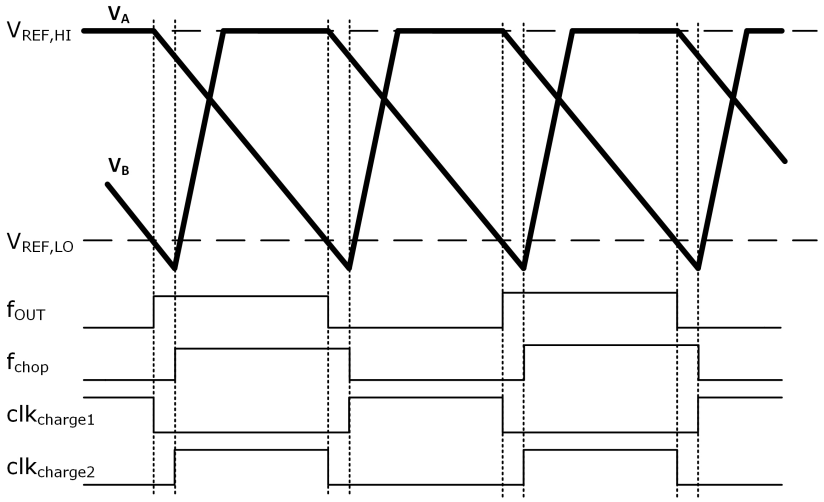


Figure 4.2: Time diagram for the implementation [SBM+09]

the multiplexer, is depicted in figure 4.2. The two reference capacitances are precharged alternatively to the high reference voltage $V_{REF,HI}$ and then linearly discharged by the defined reference current. At the time when the discharging capacitor's voltage drops below the low reference voltage $V_{REF,LO}$, the comparator switches its output and immediately the other capacitor starts to linearly be discharged. The recharge of the first capacitor is slightly delayed to ensure that non-idealities of the comparator do not have any effect on the slope of the discharge at the crossing. In addition it does not influence the time period T_{osc}.
It consists of

- the current reference
- the n-channel capacitors
- the comparator
- two analog multiplexers for the input of the comparator
- the switches between the reference voltages and the capacitors
- the digital circuitry to generate the driving signals for the switches

4.1 Voltage Reference

The two voltage references $V_{REF,HI}$ and $V_{REF,LO}$ define the range, in which the reference capacitor is charged and discharged, respectively. A general consideration for these voltages is the influence of disturbances, which come from non-idealities like clock

feedthrough and channel charge injection of the switches. The higher the difference between the low and the high reference is, the lower is the error $\Delta V/V$.
For the testchip measurement it is decided to supply the voltages from external. The first reason for the decision is, that for test purposes $V_{REF,HI} - V_{REF,LO} = max$ is achieved by providing the highest possible voltage. The possibility of connecting the high reference voltage pin directly to the internal supply shall be avoided, because any noise on this supply voltage or any deviation on the voltage level influences the frequency, what is not wanted and can not be corrected for the measurements. Hence, for a nominal supply of $V_{DD1v5} = 1.5V$, it is decided to be isolated and to vary for this voltage from external.
The second consideration is to see the influence of the low voltage reference, which is set to $V_{REF,LO} = 900mV$ due to the characteristics of the n-channel transistor being used as the reference capacitor and its region, in which it is in strong inversion and has a stable capacitance. For this purpose it can be varied the easiest way, if it is defined from external.

4.2 Comparator

On the date of the tape-out of the testchip, the comparator, which is intended to be implemented, was still in the design phase and not yet layouted as well. Hence, only for the testchip a simple miller operational amplifier is implemented as comparator.
Regarding its technical properties, the design focuses on two major characteristics, which are relevant for the testchip only:

- common mode input voltage
- speed and propagation delay

The power consumption of the comparator is not taken into consideration at this time being. Though, if the decision is taken after the results in the laboratory, that it is worth to improve and continue with the mobility-based oscillator, the comparator will be replaced by a more efficient circuit.

The comparator is a two-stage amplifier with two inverters at its output. The latter amplify the output voltage of the amplifier and ensure to have a "rail-to-rail" output at

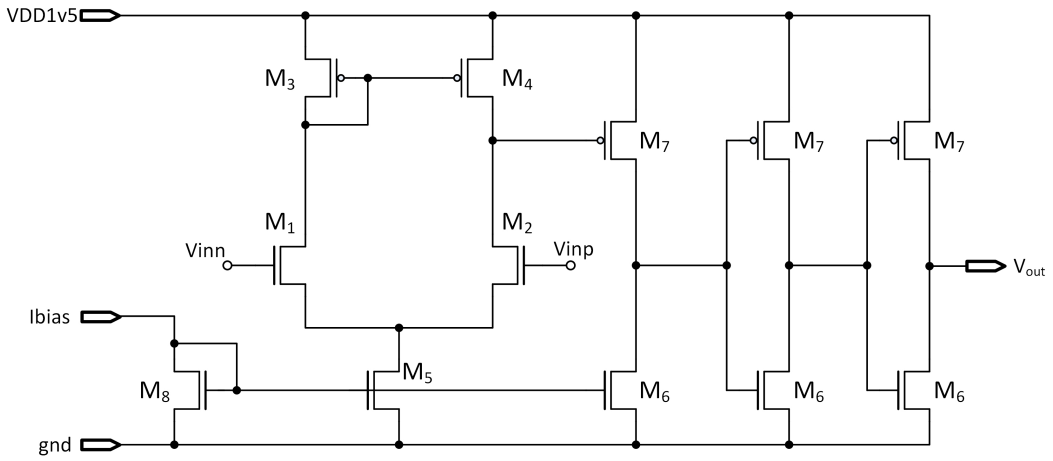


Figure 4.3: Schematic of the comparator

V_{out}. This is necessary since the amplifier's output is limited and cannot go higher or lower than the saturation voltage of the second stage's transistors.

$$\begin{aligned} V_{out,max} &= V_{DD1v5} - V_{dsat,M6} \\ V_{out,min} &= V_{dsat,M7} \end{aligned}$$

Since the testchip's objective is the measurement of the mobility-based current generation and the behaviour in combination with the n-channel transistor as a capacitor, the influence of the comparator's non ideal performance such as its delay is wanted to be neglectable. The determined frequency is $f_{osc} = 100kHz$ and with a duty cycle of 50% the time constant is $T_{osc}/2 = 5\mu s$ at room temperature.
The calculated error of $\dfrac{\Delta T}{T_{osc}/2} < 0.5\%$ introduced by the comparator due to the delay leads to

$$\begin{aligned} t_{del,comp} &< 0.005 \cdot 5\mu s \\ t_{del,comp} &< 25ns \end{aligned}$$

The slew rate of the second stage is defined by its bias current and the capacitances on this node

$$\frac{\Delta V}{\Delta t} = \frac{I_{bias,M6}}{C_L}$$

with the parasitic capacitances of the transistors as the load capacitance

$$\mathrm{C}_L = C_{gd,M6} + C_{gd,M7} + C_{in,INV}$$

With a nominal supply of $V_{DD1v5} = 1.5V$ and the derived time delay, the minimum bias current can be calculated.
If the delay is constant over temperature and process, it can be seen as a static offset and as such taken into account.

The limits for the common-mode input voltage are set by the first stage and defined as following

$$\begin{aligned} V_{cm,max} &= V_{DD1v5} - V_{th,M3} - V_{dsat,M3} + V_{th,M1} \\ V_{cm,min} &= V_{dsat,M5} + V_{th,M1} + V_{dsat,M1} \end{aligned}$$

Since the input of the comparator is chopped, the switching voltage level does not change but it is always the low reference voltage. If the output of the oscillator is high, on the positive input of the comparator V_A is in the discharge phase till it drops below the low reference voltage and the comparator switches and the inputs are chopped. Hence, the comparator is optimized for this common-mode voltage level.

The input common-mode voltage is determined by the low reference voltage level. The low reference voltage level is chosen, that the n-channel transistor as the reference capacitor is in the saturation region with a certain safety distance to the roll-off of the depletion region. The low voltage reference is set to $V_{REF,LO} = 900mV$ and for that voltage the comparator is designed.

A further consideration is the influence of the parasitic capacitances of the input pair transistors M_1 and M_2. If the size of these transistors is chosen too large, their fringe and overlap capacitances influences the overall capacitance on this sensitive node and hence the time constant. For this reason, the transistors' area is kept small, but the bias current is increased to have a certain gain. In addition, the second stage's

as well as the first inverter's transistors are designed small to reduce their parasitics capacitances, which have an effect on the delay of the comparator, to a minimum.

4.3 Analog Multiplexer

As can be seen in figure 4.1, two multiplexers are placed at the input of the comparator to mitigate the effect of its offset. The offset voltage can be seen as a constant voltage source V_{os} on one of the inputs of the comparator. Thus, the output is not switched anymore exactly on the lower reference voltage, but on $V_{REF,LO} - V_{os}$. The advantage with the multiplexers is, that the offset is once added and once subtracted from the reference voltage with the resulting switching levels $V_A = V_{REF,LO} - V_{os}$ and $V_B = V_{REF,LO} + V_{os}$. This shifts of course the duty cycle of the oscillator, but the period remains the same.

$$\begin{aligned} T_{osc,high} &= \frac{C}{I_{ref}} \left(V_{REF,HI} - V_{REF,LO} - V_{os}\right) \\ T_{osc,low} &= \frac{C}{I_{ref}} \left(V_{REF,HI} - V_{REF,LO} + V_{os}\right) \end{aligned}$$

With matching discharge currents and capacitances, the oscillator's time constant is

$$T_{osc} = \frac{2C}{I_{ref}} \left(V_{REF,HI} - V_{REF,LO}\right)$$

and the offset voltage is mitigated.

One multiplexer has two inputs and one output and is designed with two switches in parallel, which are driven with complementary logic signals. The two multiplexers are depicted in figure 4.7.
The approach here is, to influence neither the input of the comparator nor the capacitor, which is discharged, with charge injection from the switch. Regarding the charge injection on the capacitor's side,

4.3.1 Channel charge injection

The charge injected on the node V_A and V_B, respectively, is deposited on the capacitor, introducing an error in the voltage level. Especially in the phase, when the capacitor is discharged, it changes the voltage on that node, what leads to a deviation of the oscillator's frequency. The resulting error equals

$$\Delta V = \frac{WLC_{ox}(V_{DD} - V_A - V_{th})}{2C_{ref}} \text{[Raz01]} \tag{4.1}$$

For this equation it is assumed, that exactly half of the channel charge is injected on the one side and the other half at the other side whereas in reality, the fraction of charges is a complex function of different parameters such as the impedance seen on drain and source of the transistor and the clock signal, which drives it.

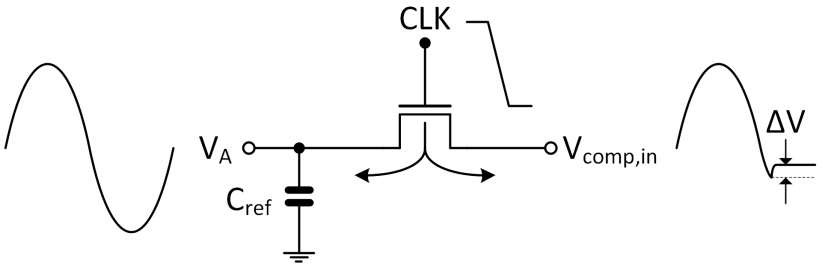


Figure 4.4: Channel charge injection when a switch turns off [Raz01]

Out of this definition, the amount of charges is directly proportional to the size of the transistor as well as its oxide capacitance. As the latter parameter cannot be influenced, the transistor's width and length is kept as small as possible to minimize the effect. Altough this design influences the resistance r_{on} of the switch, it can be neglected as long as it connects the V_A and the gate of the comparator's input transistors and hence no current flows through the switch. However, the aspired frequency is $f_{out} = 100kHz$. For higher frequencies a high r_{on} is a limiting factor and must be taken into account.

4.3.2 Clock feedthrough

Another effect, which accompanies switches is the clock feedthrough. During the clock transition, the clock signal is coupled on the connected nodes through its gate-drain and gate-source overlap capacitances. With the assumption, that the overlap capacitance is constant, it can be expressed as

$$\Delta V = V_{CLK} \frac{WC_{ov}}{WC_{ov} + C_{ref}}$$

with the overlap capacitance per unit width.

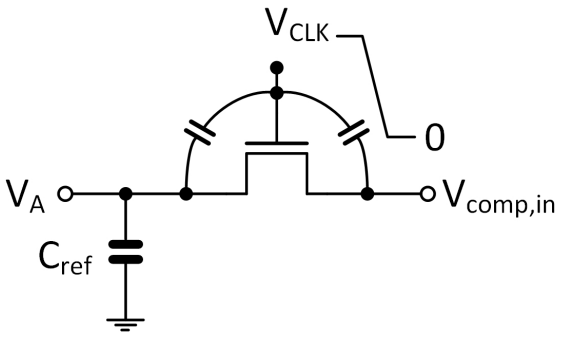


Figure 4.5: Clock feedthrough when a switch turns off [Raz01]

Again, the approach to keep this effect at a low level, the transistor's size, especially its width, is chosen small. Furthermore, since the reference capacitance is large compared to C_{ov} the introduced error can be set to a minimum.

A possibility to lower the effect of channel charge injection is the switch being built up as a transmission gate. This transmission gate uses a n-channel and a p-channel transistor in parallel such that the opposite charge packets injected by the two cancel each other as depicted in figure 4.6. Considering equation 4.1, it must be $W_1 L_1 C_{ox}(V_{DD} - V_A - V_{th,n})$ $= W_1 L_1 C_{ox}(V_A - |V_{th,p}|)$, thus the cancellation is only valid for one input voltage level.

In addition the transmission gate has better characteristics in terms of feed through all voltage levels. A n-channel transistor does not fully turn on, if the voltage signal is close to the digital high voltage level and the same problem occurs with a p-channel transistor with a very low voltage level on the capacitor. Thus, depending on the node

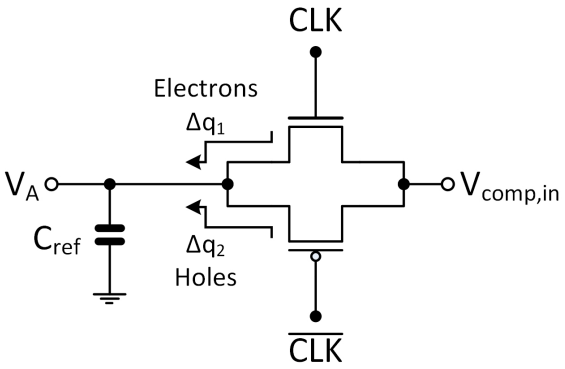


Figure 4.6: Use of complementary switches to reduce charge injection [Raz01]

potential, at least one of the transmission gate's transistors is fully turned on.

The analog multiplexer is a combination of two transmission gates for each input of the comparator and they are driven with inverted signals like shown in figure 4.7.

The driving signals for the transmission gates are generated by a digital circuitry, which is explained later in detail.

4.4 Digital Circuitry

The signals for the switches and the multiplexer are provided by a digital circuit. According to the time diagram in figure 4.2, the following considerations are relevant:

- The switch between the high reference voltage $V_{REF,HI}$ and the capacitor, which has to be discharged, has to opened as fast as possible right after the comparator has switched its output.

- To recharge the other capacitor the second switch is closed with a certain delay in order not to have disturbances at the comparator's input.

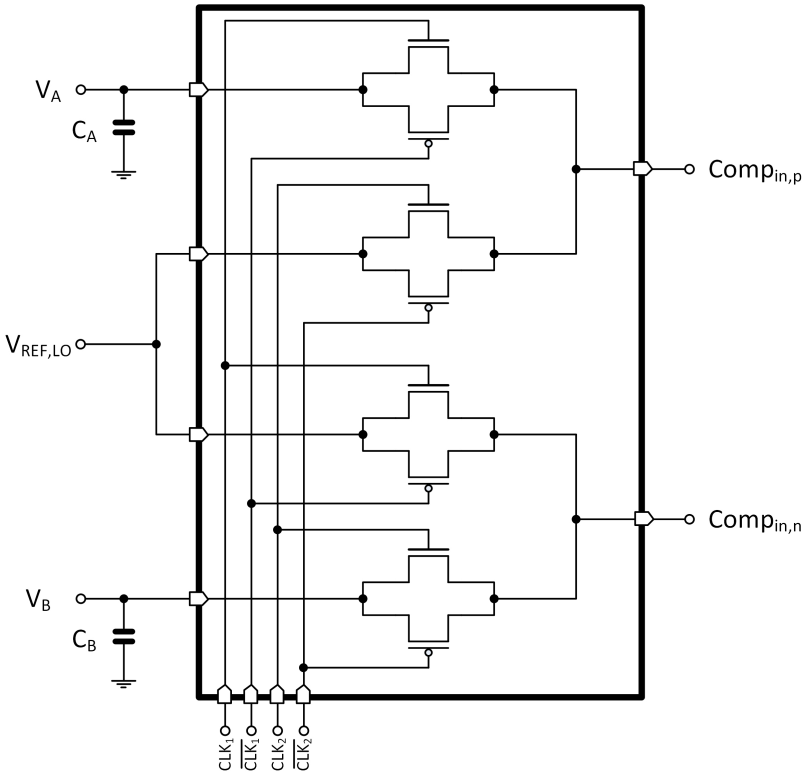


Figure 4.7: Two analog multiplexers

- Regarding the time, when the comparator's input must be chopped, ideally it happens before the switch for the discharged capacitor closes. If the capacitor starts to recharge while it is still connected to the comparator and depending on the time in which the capacitor is recharged, it might cause disturbances on the comparator.

The two signals f_{chop1} and f_{chop2} are generated by a non-overlapping clock circuit. It is depicted in figure 4.8.

With this non-overlapping clock it is guaranteed, that the transmission gates, which are closed, first open and then the others are closed. If both were closed in the same time, the capacitor's node would have been directly connected to $V_{REF,LO}$ and discharged uncontrolled by the low reference voltage.

As mentioned before, the multiplexer is built of transmission gates and for the p-channel transistors in it, the complementary signals $\overline{f_{chop1}}$ and $\overline{f_{chop2}}$ must be provided, what is done by just inverting the original signals.

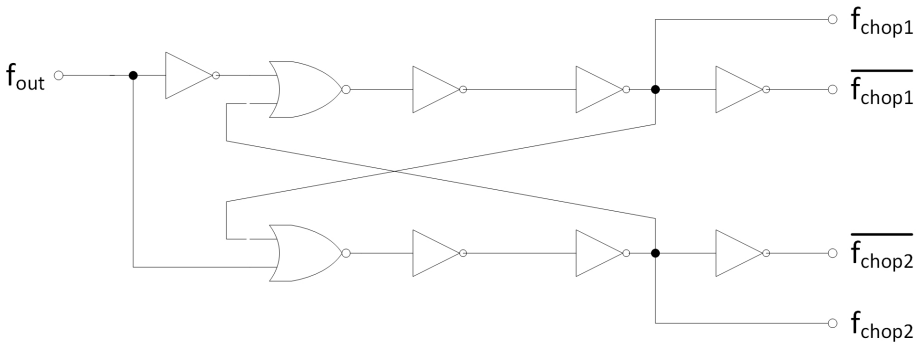


Figure 4.8: signal generation for the multiplexer

The signals $clk_{charge1}$ and $clk_{charge2}$ are delayed in an alternate sequence, depending on whether the oscillator's output switches from high to low or vice versa. According to this requirement, this delay must be designed. A common way is the utilization of three inverters with an resistor-capacitor (RC) element in between.
Like it is depicted in figure 4.9, the delay time depends on the speed of the first inverter and how fast it can charge and discharge the capacitor and also on the time constant $\tau = RC$ of the low pass. The second and third inverter after the low pass are expected

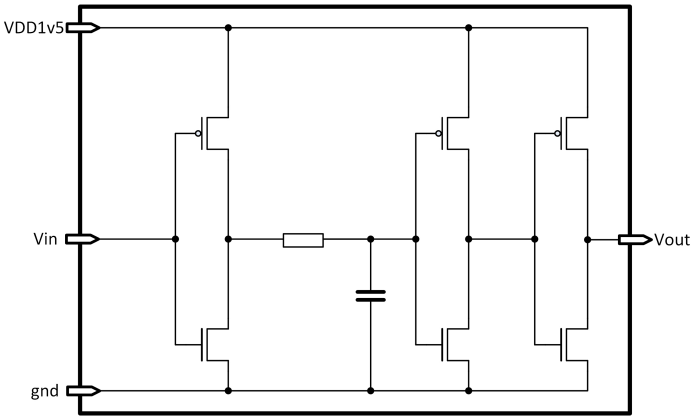


Figure 4.9: Time delay generating circuit

to provide a steep edge on the output of the delay cell. It has to be taken into account, that the total chain is inverting.

Since the propagation delay for a logic gate is in the range of 10ps-100ps, the cell is designed to delay for 2ns. This ensures to delay this signal longer than any other delay arisen from routing or switching. In this case, no accurate value is needed, since it only

has to ensure, that all the other switches have already changed their state before. The logic circuit is shown in figure 4.10.

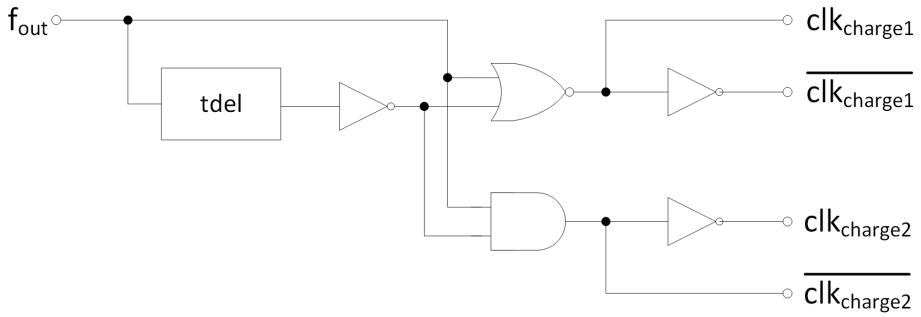


Figure 4.10: Signal generation for the switches to recharge the capacitors

The fast switching transitions at the input of the comparator can potentially lead to spikes at the output due to capacitive coupling at the input transistors. As a result and to avoid to have disturbances on the logic, which is driven by the comparator's output, a spike filter is added between the output of the comparator and the logic. The filter is expected to immediately let pass the first switch of the comparator, but filter possible spikes after the transition is made. The implementation is done with the logic shown in figure 4.11.

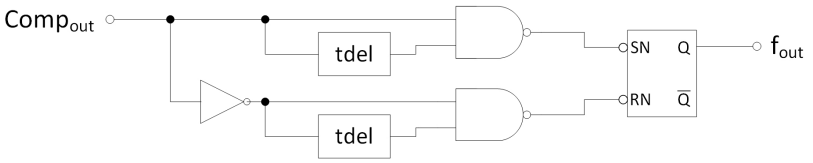


Figure 4.11: Circuit to filter out spikes on the output of the comparator

The same delay like shown before is utilized in this filter circuit. Both for setting or resetting the RS-flipflop spikes are canceled out if they do not last longer than the delay.
One configuration, which is the same as in the logic of the voltage reference's switches, is the delay cell together with the inverter. The delayed and inverted signal from the spike filter is also used in the other logic and there the two cells can be removed.

Considering the startup behavior, the logic must provide the correct states for the switches and the multiplexer, otherwise it could lead to a deadlock of the circuitry. Hence, additional logic gates with the enable signal are added to always start from a defined state

- V_A=$V_{REF,HI}$ $\rightarrow$ $clk_{charge1} = 1$ $clk_{charge2} = 0$
- $V_{comp,inp}$= V_A $V_{comp,inn}$= $V_{REF,LO}$ $\rightarrow$ $f_{chop1} = 1$ $f_{chop2} = 0$

The switch between V_A and $V_{REF,HI}$ can be closed also when the oscillator is disabled as the current reference does not provide a discharge current in the power-down mode. Otherwise it would have consumed unwanted current consumption. The overall digital circuitry is shown in figure 4.12.

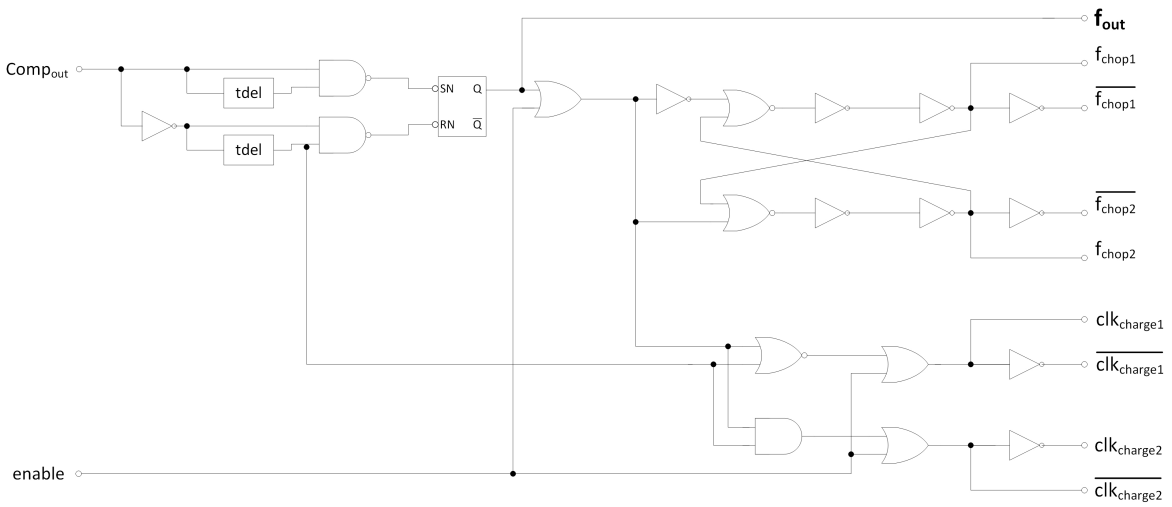


Figure 4.12: The complete digital circuitry

The signals of the digital circuit are shown in the plot 4.13. The input voltage on the pin $comp_{out}$ is connected to an ideal voltage source with a rise fall time of $t_{ft} = 500ps$. First, the multiplexer's and the reference voltage switch's transmission gates open, when f_{chop1} and $clk_{charge1}$ go to 0. After that, the other transmission gates of the multiplexer closes with f_{chop2} and after a certain delay, the discharged capacitor is recharged by closing the switch with $clk_{charge2}$.

4.5 Simulation

At the temperature $T = 27°C$ and nominal process corner the waveforms of the oscillator are shown in figure 4.14. It can be seen, that at the beginning of a discharge phase of a capacitor, the voltage breaks down for roughly 20mV. This effect can be explained by channel charge injection and clock feedthrough and further by parasitic capacitances which are recharged. One of these is the input capacitance of the comparator, which

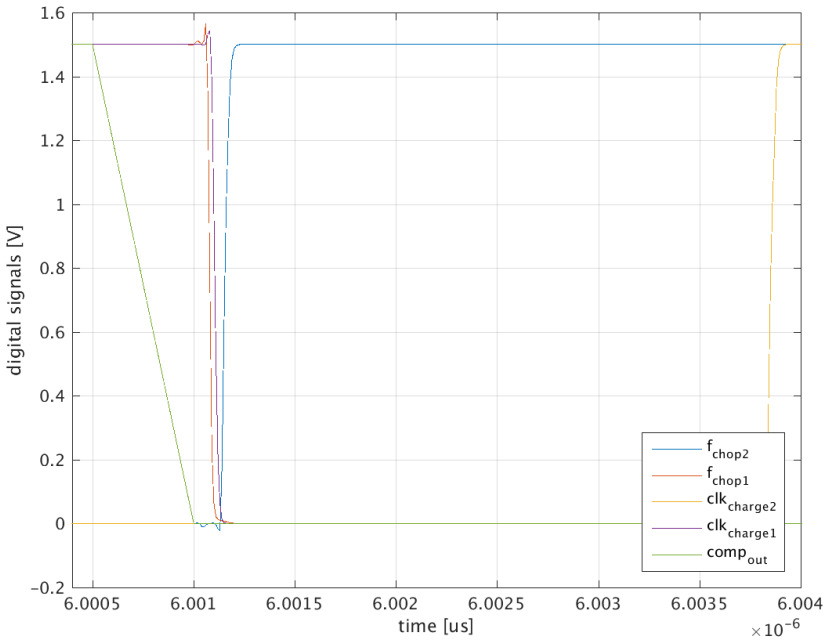


Figure 4.13: Waveforms of the generated clock signals

first is charged to $V_{REF,LO}$. As soon as the multiplexer switches, it must be charged from the low reference voltage to the high reference voltages and takes the necessary current from the reference capacitance. This effect can be limited, if the switch between the high voltage reference $V_{REF,HI}$ and V_A is opened after the multiplexer switches the comparator's inputs.

Regarding the reference current and the capacitance as the main part of the time constant, simulations of the current in combination with the capacitance value as a function of temperature is accomplished and depicted in figure 4.15. In a previous chapter, in which the reference current is explained and simulations are made, variations on the process introduce an error of

$$I_{OUT} = I_{OUT}(T) \pm 20\%$$

Together with the capacitance value and a constant factor for the discharge voltage range of $V_{REF,HI}V_{REF,LO} = 0.6V$, the frequency is calculated. Due to the elimination of the oxide capacitance as part of the frequency's definition, the relative deviation is

$$f_{OUT} = f_{OUT}(T) \pm 10\%$$

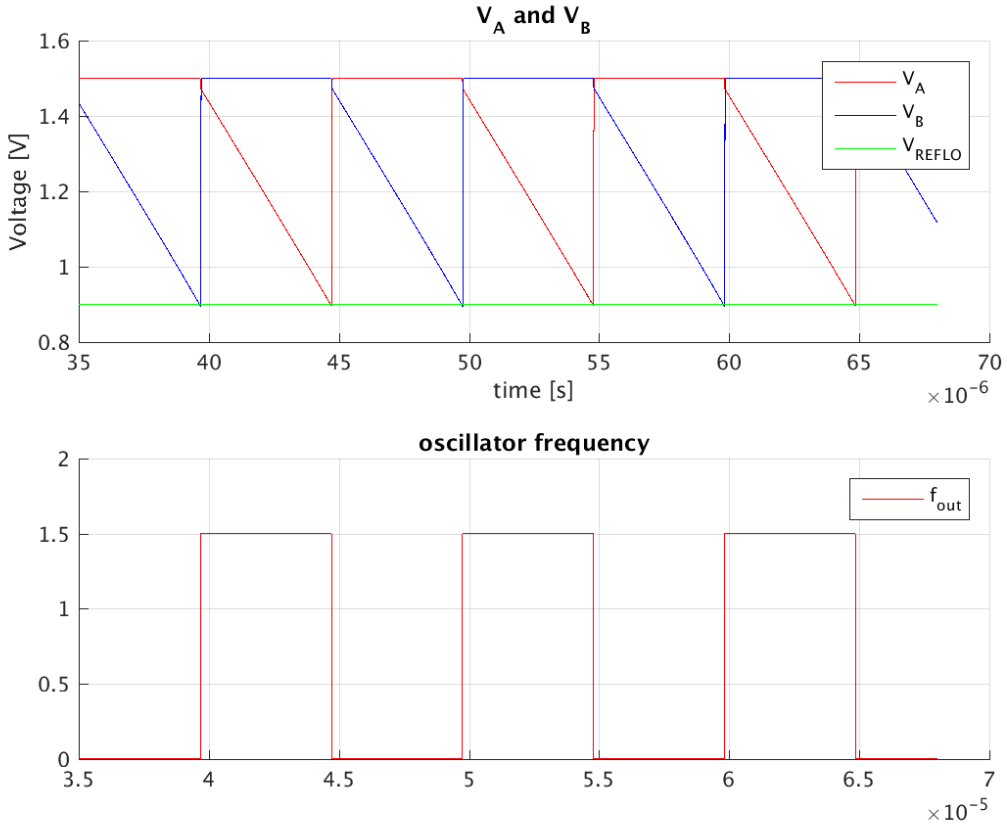


Figure 4.14: Voltage at the two reference capacitors and the output frequency signal

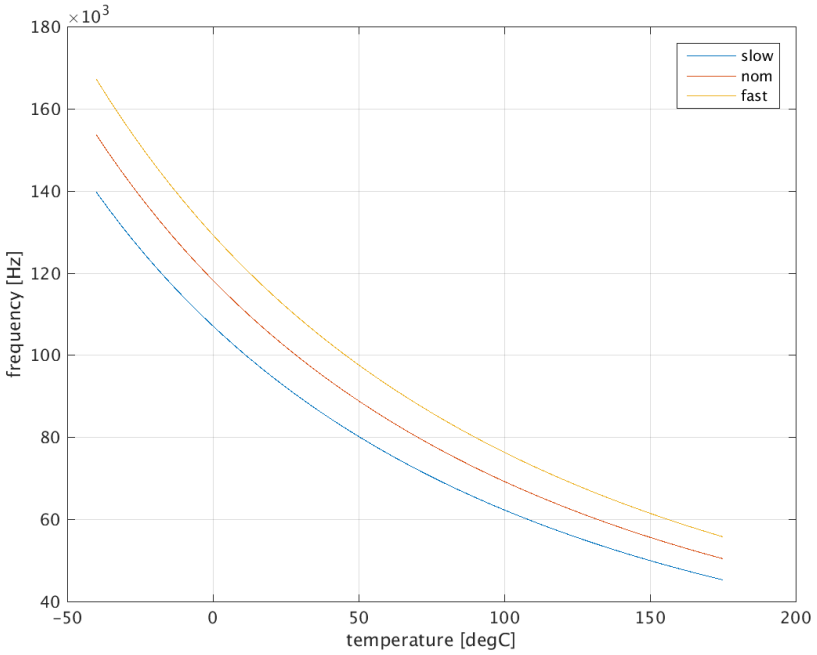


Figure 4.15: Untrimmed frequency with different process corners and $V_{R_1} = 162.5mV$

With a simple optimization of the bias current, which generates the voltage drop over the resistor R_1 between the gates of M_1 and M_3, the following curves are achieved. The resulting relative accuracy of the frequency as function of the reference current and the capacitance only is

$$f_{OUT} = f_{OUT}(T) \pm 0.65\%$$

This confirms, that despite the strong carrier mobility's temperature dependence, it is well defined and can be compensated for.

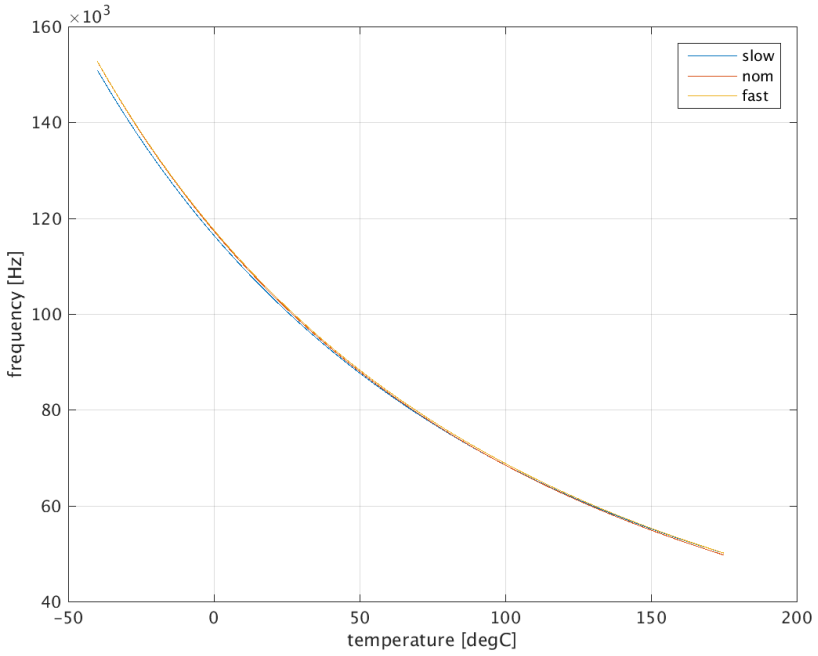


Figure 4.16: Frequency with different process corners and single-point trimming

4.6 Layout

The layout of the testchip circuit is depicted in figure 4.17.

- green: the MOS transistors as reference capacitor.
- red: biasing transistors for the amplifiers.
- orange: n-channel current mirror to provide the discharge current for the MOS capcacitors.
- blue: n-channel transistors M_1 and M_3.

- purple: current mirror p-channel transistors M_4 and M_2 with common centroid layout. The other transistors, which surround them, are utilized as capacitor C_{C_2} and for better matching of the mirroring transistors.
- yellow: digital circuitry with a proper distance to the sensitive analog area.

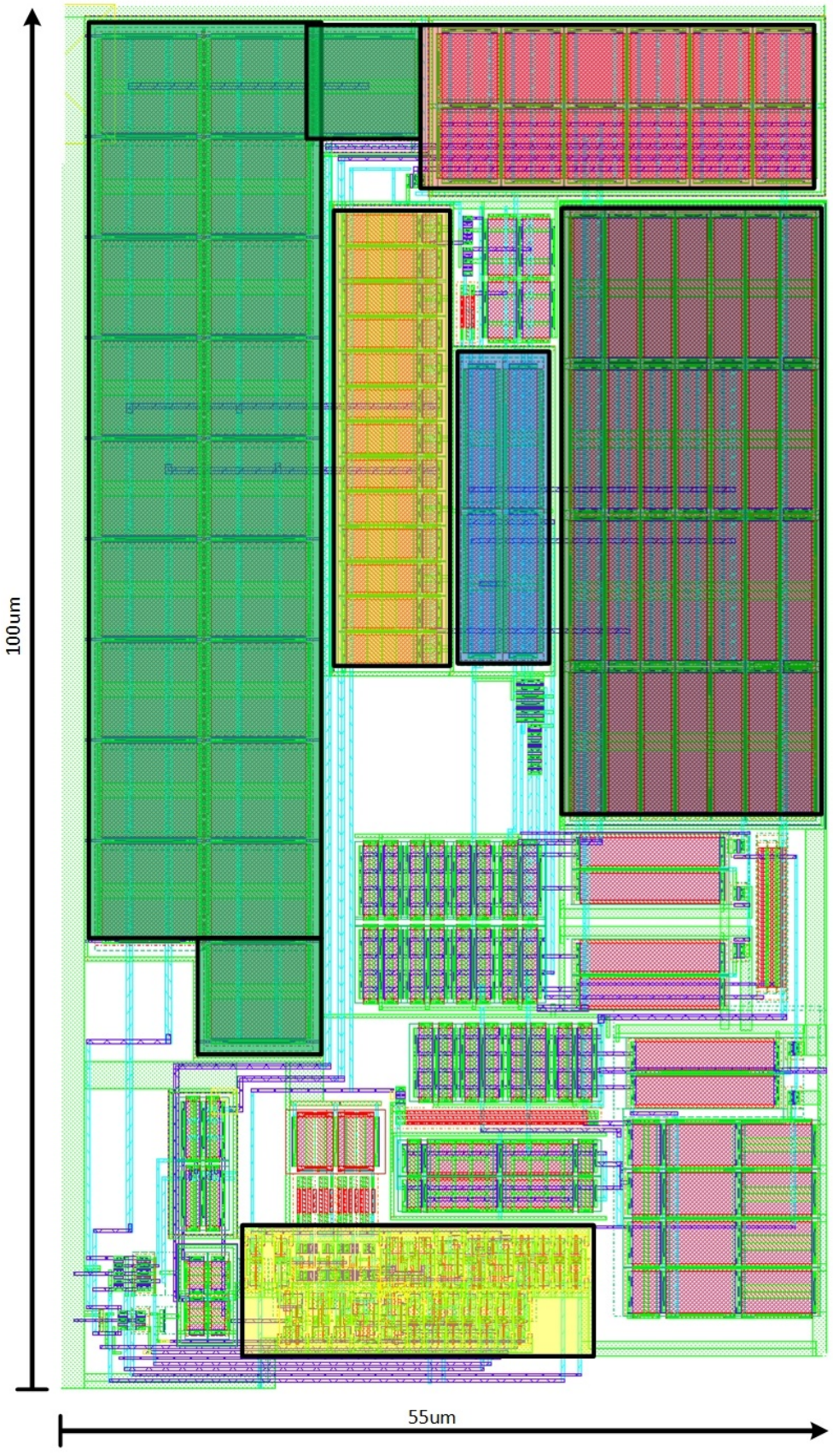


Figure 4.17: Testchip Layout

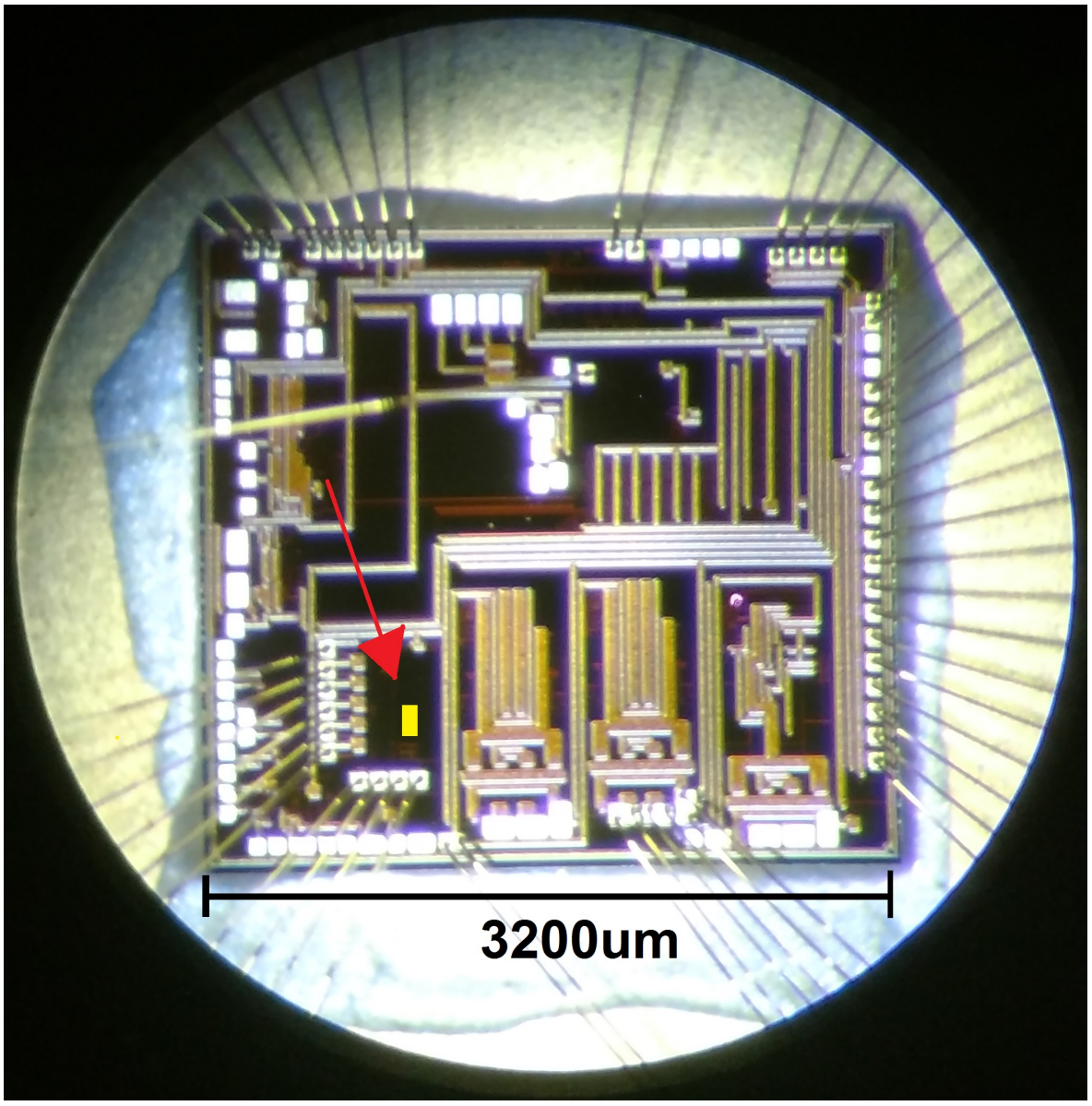


Figure 4.18: Testchip

5 Temperature Compensation

In the following chapter, a possible compensation of the oscillator's temperature behavior will be presented. Only the current reference and the n-channel transistor as reference capacitance are considered for the temperature compensation, as it is the major impact on the time constant of the oscillator. The voltage reference is assumed to be ideal and taken into account with a constant factor for all the presented simulation results.

In the previous section it has been shown that the frequency of the mobility-based oscillator is strongly temperature dependent. Although the temperature behavior is well defined, a temperature-stable oscillator is required to be usable as a clock reference in most applications. To repeat once again the frequency as a function is defined as the following:

$$f_{osc} = \frac{I_{REF}}{2C_{REF}} \frac{1}{V_{REF,HI} - V_{REF,LO}}$$

With the reference current from the mobility-based current reference

$$I_{REF} = \frac{\mu_n C_{OX}}{2} \frac{W_1}{L_1} \frac{V_{R_1}^2}{\left(\sqrt{\frac{n}{m}} - 1\right)^2}$$

and the n-channel transistor as the reference capacitance with a certain width W_C and length L_C and operated in strong inversion and the oxide capacitance C_{ox} per unit area

$$C_{REF} = C_{OX} W_C L_C$$

the frequency can be rewritten

$$\begin{aligned} f_{osc} &= \frac{\mu_n C_{OX}}{4} \frac{1}{C_{OX} W_C W_C} \frac{W_1}{L_1} \frac{V_{R_1}^2}{\left(\sqrt{\frac{n}{m}} - 1\right)^2} \frac{1}{V_{REF,HI} - V_{REF,LO}} \\ &= \frac{\mu_n}{4} \frac{1}{W_C W_C} \frac{W_1}{L_1} \frac{V_{R_1}^2}{\left(\sqrt{\frac{n}{m}} - 1\right)^2} \frac{1}{V_{REF,HI} - V_{REF,LO}} \end{aligned}$$

Assuming the reference voltages $V_{REF,HI}$ and $V_{REF,LO}$ are generated from a bandgap voltage and $V_{REF,HI} - V_{REF,LO} = const$ is considered to be stable as well as the voltage drop V_{R_1} over the resistance R_1, the frequency depends only on the temperature behaviour of the carrier mobility μ_n.

$$f_{osc}(T) \propto \mu_n(T)$$

The common approximation for the carrier mobility μ_n as function of temperature is

$$\mu(T) = \mu(T_{rt}) \left(\frac{T}{T_{rt}}\right)^{-k_2}$$

The temperature dependence of the frequency must follow the carrier mobility behavior and it can be written

$$f_{osc}(T) = f_{osc}(T_{rt}) \left(\frac{T}{T_{rt}}\right)^{-k_2}$$

5.1 Curve Fitting

To derive the factor k_2, which mainly depends on the doping concentration of the silicon, the frequency is simulated over temperature and the data is post-processed in *Matlab* and compared to a curve, which is based on the approximation and adapted to fit as good as possible to the simulation results. The measured frequency and the curve are depicted in figure 5.1.

As a result, a very good fitting is achieved with a temperature coefficient of $k_2 = -1.7$.

As mentioned earlier, this factor varies in the range of $-2.5 < k_2 < -1.5$ for integrated technologies.

$$f_{osc}(T) = 100kHz \cdot \left(\frac{T}{T_{rt}}\right)^{-1.7}$$

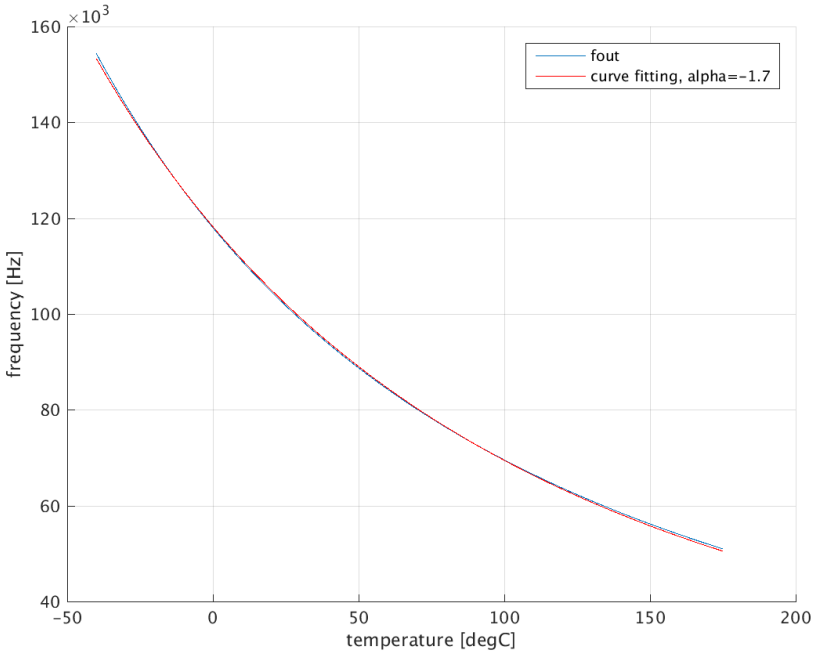


Figure 5.1: Curve fitting for the mobliity-based current generation

This knowledge about the mobility's temperature behavior serves as a basis for further procedures regarding the compensation scheme.

For this oscillator, various concepts how to implement a temperature compensation, are available:

- introduction of a temperature gradient in the reference voltages $V_{REF,HI} - V_{REF,LO} = V(T)$
- digital clock division circuit with discrete division steps in combination with a temperature sensor
- temperature-dependent voltage V_{R_1}

The approach of introducing a temperature dependent voltage range, in which the reference capacitance is discharged, is very difficult to design. As the frequency is inversely proportional to the voltage, it must change the same way like the carrier mobility does. Above that, between cold and hot temperature the frequency decreases by a factor of 3,

accordingly the voltage must decrease by a factor of 3. As shown in a previous section, the voltage level for the n-channel transistor being utilized as capacitor is restricted. The low voltage reference $V_{REF,LO}$ is set to $900mV$ and the highest maximum possible high reference voltage $V_{REF,HI}$ is limited due to the supply voltage of $V_{DD1v5} = 1.5V$ and its lower specification limit $V_{DD1v5,-10\%} = 1.35V$, $\Delta V_{max,cold} = 450mV$ for low temperature compensation and the resulting $\Delta V_{max,hot} = 450mV/3 = 150mV$. At such low voltage range, in which the capacitance is discharged, disturbances of channel charge injection, clock feedthrough, inaccuracy of the comparator and other influences increase. As a conclusion this approach is not dealt with any further.

Secondly, a digital clock division circuit is a possible compensation method. With an integrated temperature sensor, which in a CMOS process usually takes advantage of a lateral bipolar transistor, and an analog-to-digital converter (ADC), a digital circuitry divides the clock based on the information of the ADC's output. This principle is expected to compensate very well because a single point trimming of the oscillation frequency results in the same frequency at every temperature for different process variations. On the contrary, it is a big effort to design the sensor, the converter and define the clock division mapping and implement its digital circuitry. In the publication from [65nm CMOS Temperature-Compensated], a possible clock division circuitry is described, in which first a 7^{th} order polynomial and then a 4^{th} order polynomial is utilized while post-processing the measured data. As a consequence, for this is work this compensation scheme is not taken into account.

5.2 Calculation

A further approach of compensation is the introduction of a non-constant voltage drop V_{R_1} between the gates of M_1 and M_3 with a certain temperature gradient. To derive the demanded behavior, the starting point is the curve fitting, which is depicted in figure 5.1. If the voltage is temperature dependent, the previous defined oscillator as a function of temperature needs to be expanded like the following

$$f_{osc}(T) \propto \mu_n(T_{rt}) \left(\frac{T}{T_{rt}}\right)^{-1.7} \cdot (V_{R_1}(T))^2$$

To receive a stable frequency it must be

$$\begin{aligned} f_{osc}(T) &= const. \\ \frac{\delta f_{osc}(T)}{\delta T} &= 0 \\ \frac{\delta f_{osc}(T)}{\delta T} &\propto \frac{\delta\left(\mu_n(T_{rt}) \left(\frac{T}{T_{rt}}\right)^{-1.7} \cdot (V_{R_1}(T))^2\right)}{\delta T} \end{aligned}$$

With a simplification for the mobility and a definition of $V_{R_1}(T)$ with a certain temperature gradient x_1 as

$$\begin{aligned} \frac{\mu_n(T_{rt})}{T_{rt}^{-1.7}} &= K_1 \\ V_{R_1}(T) &= K_2 \cdot T^{x_1} \end{aligned}$$

the equation can be expressed with

$$\begin{aligned} 0 &= \frac{\delta\left(\frac{K_1}{K_2^2} \cdot T^{1.7} \cdot T^{2x_1}\right)}{\delta T} \\ 0 &= \frac{\delta\left(\frac{T^{2x_1}}{T^{1.7}}\right)}{\delta T} \\ 0 &= \frac{2x_1 T^{1.7} T^{2x_1-1} - 1.7 T^{0.7} T^{2x_1}}{T^{3.4}} \\ 0 &= 2(x_1 - 0.85) T^{2x_1-2.7} \end{aligned}$$

The only solution for this equation is

$$x_1 = 0.85$$

The resulting temperature dependent voltage V_{R_1} is

$$V_{R_1}(T) = V(T_{rt}) \cdot \left(\frac{T}{T_{rt}}\right)^{0.85}$$

For a better vizualization, exemplary curves of this function with three different voltages $V(T_{rt})$ are shown in figure 5.2.

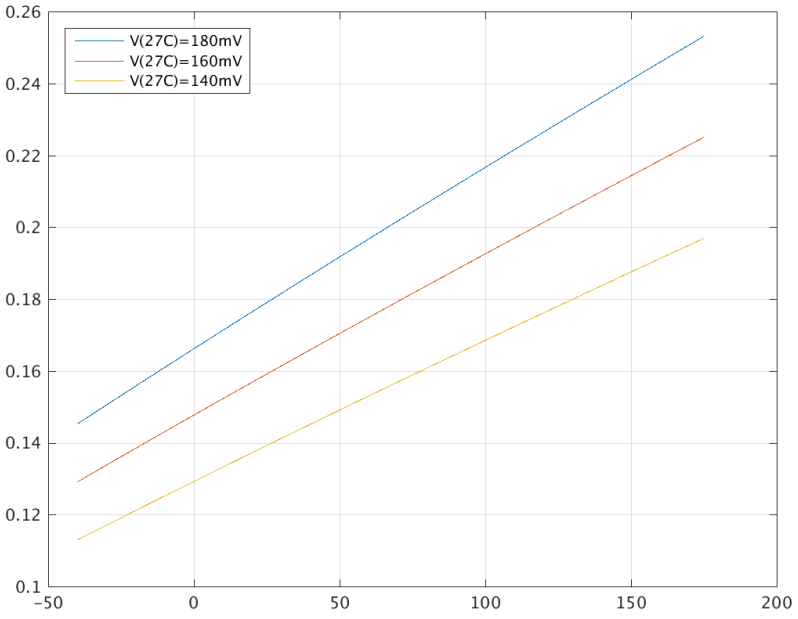


Figure 5.2: Ideal compensation voltage for three different trimming voltages $V(T_{rt})$

The request not to have only one single curve is due to the fact, that the frequency changes with process variation, but is aligned very good with one single point trimming at a certain temperature, what is in most cases room temperature T=27°C. This trimming is shown in a previous section.

However, creating a voltage or a current with a temperature gradient of T^{x_1} with $x_1 = 0.85$ is complicated. The proposal here is now to have a simple proportional-to-absolute-temperature (PTAT) signal to fit to the calculated curve as good as possible. A linear function is the basis for the approach.

$$V_{approx}(T) = kx + d = V_{new}(T_{rt}) \cdot \left(\frac{T}{T_{rt}}\right) + V_{OS}$$

To obtain the same function, it is set equal to the former equation

$$V(T_{rt}) \cdot \left(\frac{T}{T_{rt}}\right)^{0.85} = V_{new}(T_{rt}) \cdot \left(\frac{T}{T_{rt}}\right) + V_{OS}$$

At room temperature, the same voltage must be set, hence, the offset voltage V_{OS} is the difference between the two constant voltages.

$$V_{OS} = V(T_{rt}) - V_{new}(T_{rt})$$

The offset voltage V_{OS} is necessary to come close to the proposed ideal compensation voltage as can be seen in figure 5.3. In this plot the x in the legend is T/T_{rt}.

By solving the equation, $V_{new}(T_{rt})$ can be calculated by

$$V_{new}(T_{rt}) = V(T_{rt}) \frac{\left(\frac{T}{T_{rt}}\right)^{0.85} - 1}{\frac{T}{T_{rt}} - 1}$$

Depending for which temperature the equation is solved, it results in different behavior. At low temperature, the ideal compensation voltage curve has a higher gradient than at hot temperature. Thus, the middle of the temperature range, which is T=67°C = 340°K, is chosen.

The resulting function, in this case for $V(T_{rt}) = 180mV$ is

$$V_{approx}(T) = 151.5mV \left(\frac{T}{T_{rt}}\right) + 28.5mV$$

The resulting error as a ratio between the ideal compensation voltage in the temperature range -40°C <T<175°C and the fitted linear PTAT curve is shown in figure 5.4.

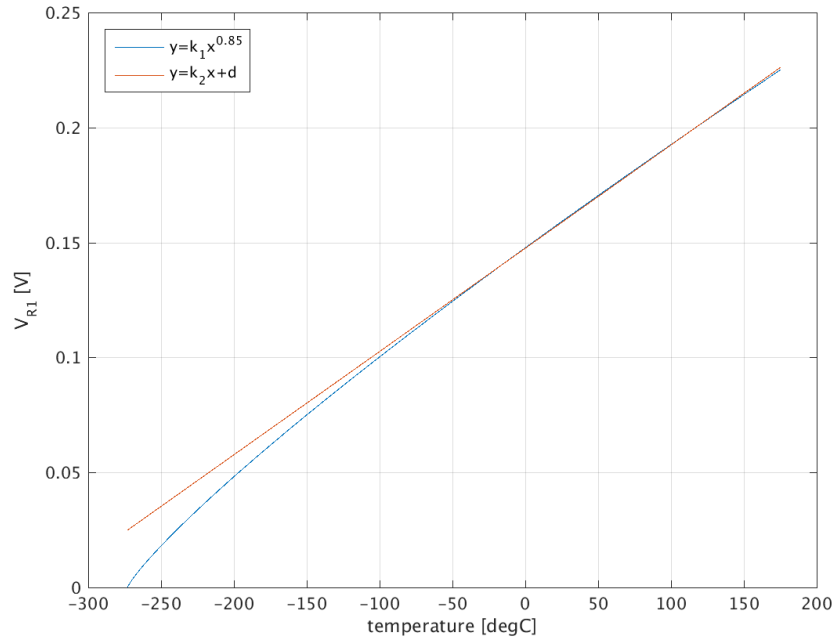


Figure 5.3: Approximation of the ideal compensation voltage with a linear function

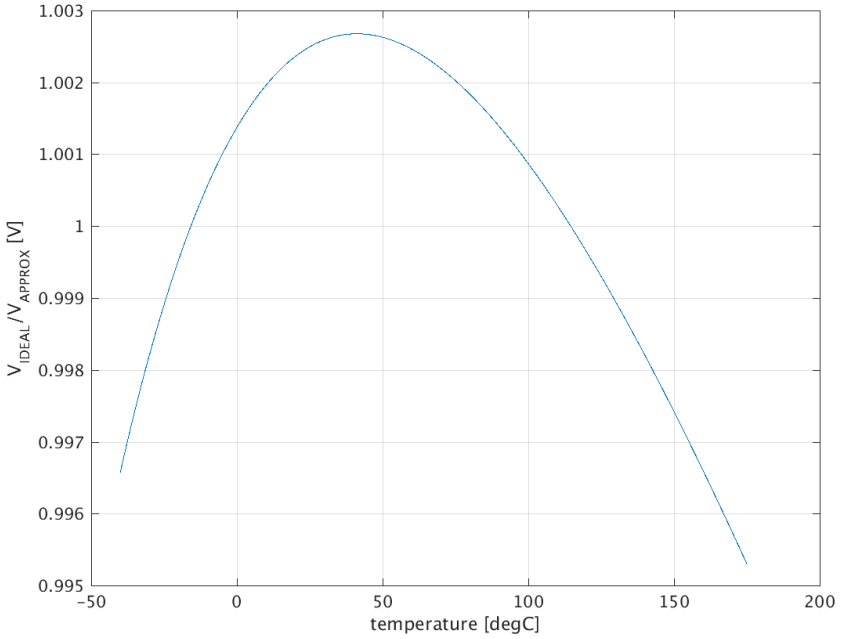


Figure 5.4: Error between the ideal curve and the approximated linear function

5.3 Simulation

To prove the curve fitting and the calculation, a simple simulation is made, in which the resistor between M_1 and M_3 is replaced by an ideal voltage source with the defined value, which is declared in equation 5.3.

$$V_{R_1}(T) = V(T_{rt}) \cdot \left(\frac{T}{T_{rt}}\right)^{x_1}$$

The result of the simulated frequency is depicted in figure 5.5. The reason, why the calculated voltage gradient (blue line) does not compensate the frequency, is the carrier mobility's μ_n dependence of electric fields. Considering the uncompensated circuit, the drain-source voltge follows the threshold voltage of the transistor M_1, what leads to a difference of the electric field between cold and hot temperature. In contrast to that, the drain-source voltage even increases slightly during the compensation, hence, the electric field is different to the uncompensated circuit and the necessary compensation voltage changes.It becomes clear, that the frequency still decreases with increasing temperature and the compensation voltage gradient is too low.

After an optimization after simulation of the gradient $x_1 = 0.905$, an accuracy is achieved of

$$f_{osc} = 100.7kHz \pm 0.16\%$$

The fact, that the frequency curve over temperature is not a straight line also for the adjusted temperature coefficient (red line), derives from the approximation of the compensation scheme. As shown in figure 5.1, already the approximation with the proposed function does not fit perfectly. Nevertheless, the result is promising and like expected.

The approximation with a linear function must be calculated again for the adapted curve resulting in a linear function of

$$V_{approx}(T) = 145.7mV \left(\frac{T}{T_{rt}}\right) + 16.28mV$$

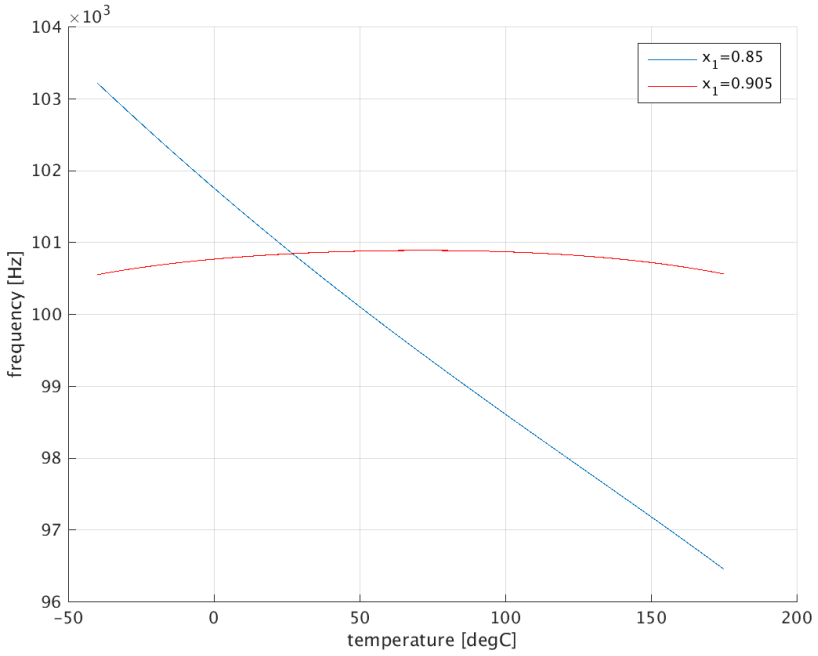


Figure 5.5: Output frequency with the calculated and optimized temperature coefficient x_1

This voltage is for the nominal process corner condition.

With the adapted compensation voltage, different corner simulations are done with a coarse one point trimming at room temperature. As shown in figure 5.6, with different corners, the frequency is more inaccurate. However, the resulting accuracy is

$$f_{osc} = 100.6kHz \pm 0.69\%$$

The reason, why the frequency shows different behavior over process variation, is explainable with leakage currents. Slow corner process shows less leakage at hot temperature than a fast corner. Hence, the frequency, which is $f_{out} \propto I_{disch}$ increases at the slow corner simulation and decreases at a fast process at hot temperature.

As at the beginning of the chapter mentioned, this accuracy only derives from the current reference circuit and the n-channel transistor as reference capacitor and their process variations. In addition, the inaccuracy of a real voltage reference generation must be included in the overall performance of the oscillator. However, a frequency

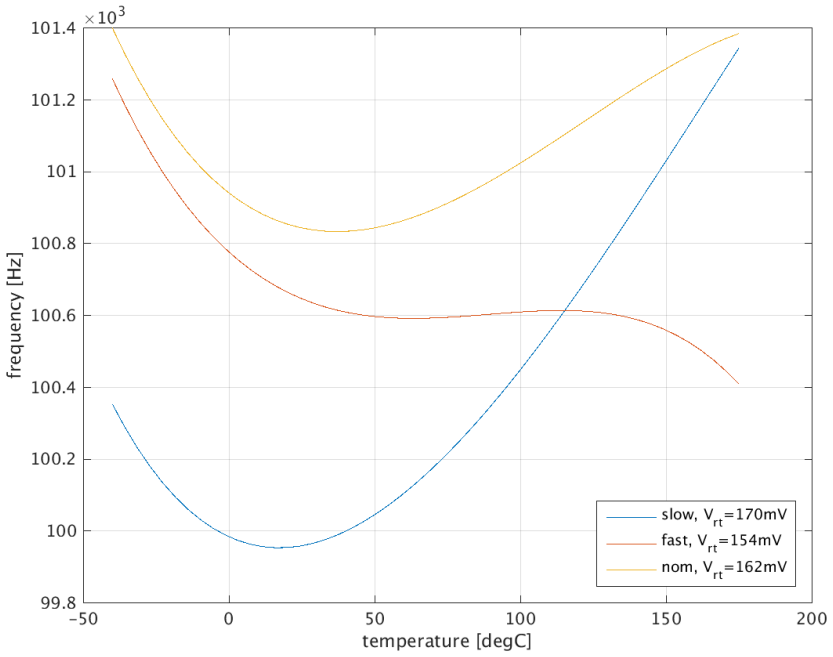


Figure 5.6: Output frequency with linear compensation voltage over different process conditions

error of $100.6kHz \pm 0.69\%$ with a simple PTAT voltage between the gates of M_1 and M_3 is achieved.
The different behavior of the simulations can be explained by the carrier mobility's changing temperature coefficient as a function of process. Different corners require different compensation curves. In figure 5.6, the temperature coefficient of the compensation voltage stays the same, only constant voltage factor is changed.

5.4 Concept

As previous explained, a compensation voltage with PTAT behavior must be designed. This linear voltage increase with temperature can be implemented by a PTAT current over the resistor R_1.

For the proposed compensation scheme, two considerations are relevant for it

- a well defined temperature gradient over process variation

- the possibility of a one-point trimming according to figure 5.2

Apart from a large number of different concepts for a PTAT voltage or current, one possible concept is the utilization of a well-known bandgap structure, which is shown in figure 5.7.

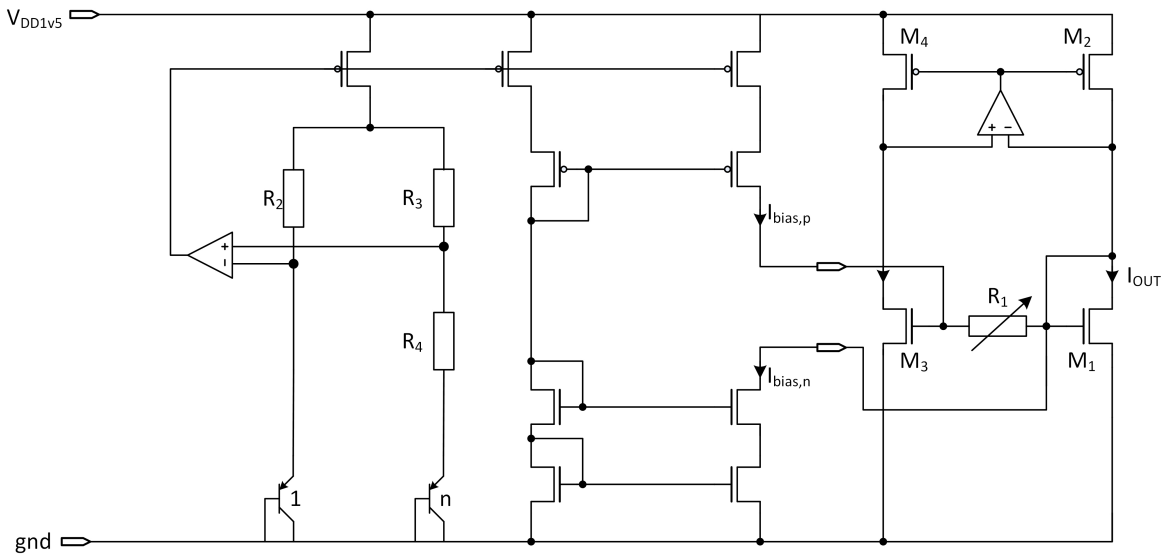


Figure 5.7: Schematic of temperature compensating biasing for current reference

With different current densities in the pnp bipolar transistors and a defined ratio between the resistors R_2, R_3 and R_4 a defined temperature coefficient of the current in each the left and the right branch is achieved. The sum of these currents flows through the p-channel transistor on top, which can be seen as the second stage of the amplifier, and is mirrored with a certain factor to provide the bias current for the resistor R_1 between the two gates of M_1 and M_3. It has to be mentioned, that the bandgap circuit does not provide a bandgap voltage anymore, it is utilized for a combination of two different PTAT currents to generate the demanded temperature coefficient. The one-point trimming can be implemented by a variable resistor R_1.

This circuit replaces the bias current generation, which is explained in the previous chapter. In addition, if the bias currents match ($I_{bias,p}$=$I_{bias,n}$), the amplifier, which diode-connects M_1, is removed, because no current flows from the biasing through the diode-connected transistor. With this modification, the current consumption of the circuit is reduced and the stability analysis is facilitated. The results from the simulation with the adapted bandgap circuit and the reference current without the

amplifier are presented in plot 5.8.

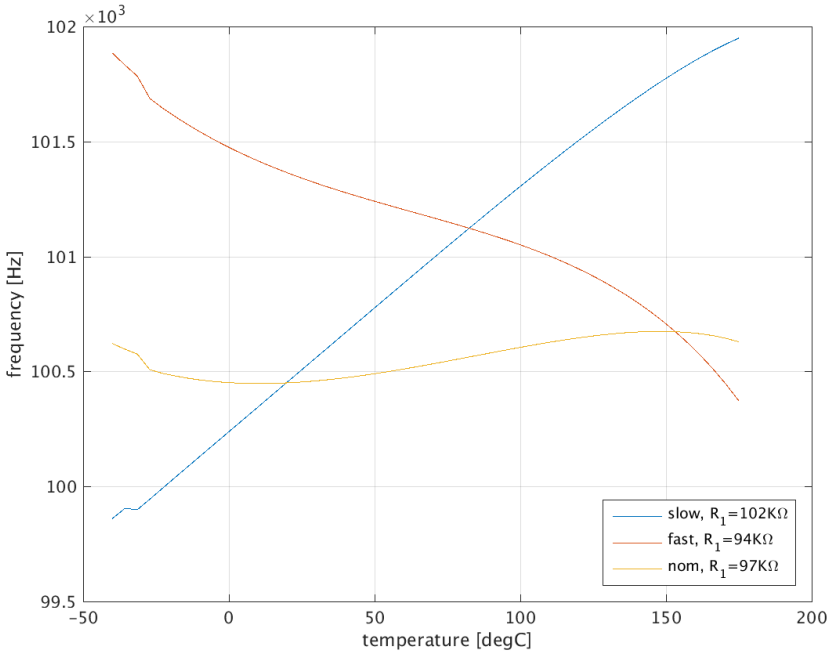


Figure 5.8: Compensated frequency with different process corner and one-point-trimming

The corner simulation with the bandgap circuit reveal the actual accuracy of a possible implemented oscillator. The frequency curve over temperature is similar to that one with an ideal voltage source (fig. 5.6), but has a higher spread. Nevertheless, with a one-point trimming of the resistor R_1, the frequency is

$$f_{osc} = 100.6kHz \pm 1.04\%$$

If no trimming is applied, the spread with only process variation is

$$f_{osc} = 100.6kHz \pm 9.1\%$$

5.4.1 Non-idealities

The shown results only include process variation. Local mismatch introduces an error in the bandgap structure and in the temperature coefficient of the bias current, which

is not considered in this work.

Of course the offset voltage V_{OS} is not constant, but changes with different trimming voltages. Depending on the concept of the compensation circuit, V_{OS} changes with the trimming or remains stable.

6 Conclusion

Finally, the results and key issues of this work are recapitulated.

Nowadays, time references are an essential part of basic circuits like bandgaps or bias current generation and present in almost every integrated circuitry. Among many different concepts, how a time reference is built, this work presents a mobility-based current-controlled oscillator.
The essential core of this oscillator is the generation of the discharge current in combination with the utilization of a MOS transistor as the reference capacitor. Although the current, which depends on the carrier mobility of the silicon, shows a strong temperature dependence and therefore also the output frequency, it is well defined and can be compensated for. In addition, the MOS capacitance eliminates deviations due to process variations of the oxide capacitance in the reference current generation.

First, the current reference is analyzed and initial steps are taken in the used technology. A circuit is built with respect to possible design weaknesses due to the temperature dependence of the carrier mobility and changing voltages and currents. Variations of the supply voltage within its specified range does not lead to any deviation in the generated current. With a single-point trimming at room temperature, the relative accuracy of the current is below $< \pm 1\%$ over process variation.

A closer examination of a MOS transistor utilized as capacitor is presented. Operated in strong inversion, the capacitance shows a stable value. The main reason why this type of integrated capacitance is chosen, is its value as a function of the oxide capacitance. Since both the generated current and the capacitance are influenced by the oxide capacitance and its process variation, it is canceled out and has no impact on the generated frequency. The combination of the current, which has a relative deviation of $\pm 20\%$ when not compensated, is reduced to $\pm 10\%$ with the oxide capacitance as a

function of the frequency.

To compare simulation results with measurements, a functional oscillator is designed. For better testability, the reference voltages are applied from external and the comparator being used is for testchip purpose only and replaced by a more efficient in case it is decided to continue with this work after the measurements are evaluated. In addition, a circuit to trim the discharge current from external is implemented.

Finally, a concept for the temperature compensation is presented. Besides the investigation of the carrier mobility's temperature behavior in the used technology, a bandgap circuit is utilized to generate a PTAT current with a defined temperature coefficient. Considering process variations only, an untrimmed accuracy of $f_{out}(T) = f_{out} \pm 10\%$ is achieved. With a one-point trimming the deviation is reduced to $f_{out}(T) = f_{out} \pm 1.04\%$.

The current consumption for the reference and the compensation circuit at ambient temperature is $I_{supply} = 17.34\mu A$. Compared to an in this technology already deployed RC-oscillator with a nominal frequency of $f_{out} = 100kHz$, the reference current generation part only is higher by a factor of 6. However, an important advantage of this circuit is the untrimmed accuracy of $\pm 10\%$ in contrast to $\pm 20\%$. Since trimming is a cost factor in production and can be avoided for circuits, which are do not depend on an accurate time reference, it is an alternative to the other oscillator. In addition, after single point trimming the proposed work is twice as accurate as the RC-oscillator.

6.1 Further investigations

Measurements on the testchip are not performed yet. They will show, whether it is worth to continue with this concept.
In the mean time, it is investigated in how the current consumption can be further reduced. Also the possibility of building the voltage reference with the generated current is considered.
A further point is the implementation of an efficient comparator, if the decision is made to continue and improve this concept.

Bibliography

[Hu09] Hu, Chenming C.: *Modern Semiconductor Devices for Integrated Circuits*. Pearson, 2009. – ISBN 0136085253

[KK02] Kogure, H. ; Kobayashi, H.: Analysis of CMOS ADC Nonlinear Input Capacitance. In: *EICE TRANS. ELECTRON.* E85-C (2002), may, Nr. 5, S. 1182 – 1190

[Raz01] Razavi, Behzad: *Design of Analog CMOS Integrated Circuits*. McGraw-Hill International Edition, 2001. – ISBN 0–07–118815–0

[SBM+09] Sebastiano, F. ; Breems, L. ; Makinwa, K. ; Salvatore, D. ; Domine, L. ; Bram, N.: A Low-Voltage Mobility-Based Frequency Reference for Crystal-Less ULP Radios. In: *IEEE JOURNAL OF SOLID-STATE CIRCUITS* 44 (2009), july, Nr. 7, S. 2002 – 2009. http://dx.doi.org/10.1109/JSSC.2009.2020247. – DOI 10.1109/JSSC.2009.2020247

[SBM+11] Sebastiano, F. ; Breems, L. ; Makinwa, K. ; Salvatore, D. ; Domine, L. ; Bram, N.: A 65-nm CMOS Temperature-Compensated Mobility-Based Frequency Reference for Wireless Sensor Networks. In: *IEEE JOURNAL OF SOLID-STATE CIRCUITS* 46 (2011), july, Nr. 7, S. 1544 – 1552. http://dx.doi.org/10.1109/JSSC.2011.2143630. – DOI 10.1109/JSSC.2011.2143630

[SOES88] Sansen, W. ; Op'T Eynde, F. ; Steyaert, M.: A CMOS Temperature-Compensated Current Reference. In: *IEEE JOURNAL OF SOLID-STATE CIRCUITS* 23 (1988), june, Nr. 3, S. 821 – 824

[Tsi] Tsividis, Yannis: *Operation and Modeling of The MOS Transistor*. OXFORD UNIVERSITY PRESS

[unk03] UNKNWON: *Course Lecture*. online, 2003

[URAC09] UY, K.J.S. ; REYES-ABU, P.A. ; CHUNG, W.Y.: A High Precision Temperature Insensitive Current and Voltage Reference Generator. In: *International Journal of Electronics and Communication Engineering* 3 (2009), Nr. 2, S. 324 327

[WA12] WOLPERT, D. ; AMPADU, P.: *Managing Temperature Effects in Nanoscale Adaptive*. Springer, 2012. – ISBN 978–1–4614–0747–8

[Wap07] WAPPIS, Herwig: *Auswirkungen und Einflüsse von automobilen Zuverlässigkeitsanforderungen auf analoge Schaltungskonzepte in hochvoltfähigen Deep-Sub-Micron Technologien*, Technische Universität Graz Institut für Elektronik, Diplomarbeit, September 2007