Dipl.-Ing. (FH) Harald Enzinger

Ultra Low Power Voltage Regulator for a Wireless Sensor Node

Master Thesis

Graz University of Technology

Institute of Electronics
Head: Univ.-Prof. Dipl.-Ing. Dr.techn. Wolfgang Bösch

Supervisor: Dipl.-Ing. Dr.techn. Mario Auer
External Supervisor: Dipl.-Ing. Günter Hofer

This work was created in cooperation with
Infineon Technologies Austria AG.

Graz, August 2012

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

STATUTORY DECLARATION

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 marked all material which has been quoted either literally or by content from the used sources.

date

signature

Abstract

This thesis describes the analysis and design of an ultra low power linear voltage regulator that is targeted for the application in a wireless sensor node.

In the introductory chapter, the concepts of wireless sensor nodes, power management and ultra low power design are summarized.

The second chapter contains fundamental considerations concerning the design of linear voltage regulators. A small signal model of the control loop is presented and methodologies for stability analysis and performance measures are introduced. The AC characteristics of both NMOS and PMOS pass device transistors are investigated by analytical calculations. A symbolic circuit analysis methodology is introduced that allows the analytical treatment of larger circuits. This methodology is applied to derive analytical models of the power supply rejection of simple CMOS OTAs.

In the third chapter, the design of the voltage regulator based on an NMOS pass device is described. A new circuit topology for the control loop is presented that allows the biasing of the pass device for low dropout operation. The proposed topology is analytically investigated concerning stability, power supply rejection and output impedance. Additional circuit blocks like an oscillator, a charge pump and a startup circuit are described.

The fourth chapter contains simulation results of the complete system. The functionality is shown by transient simulations and parameter variations are considered.

Finally in the last chapter a conclusion of the presented results is given.

Kurzfassung

Diese Arbeit beschreibt die Analyse und den Entwurf eines Linearspannungsreglers mit besonders niedriger Leistungsaufnahme der für den Einsatz in einem drahtlosen Sensorknoten ausgelegt ist.

Im einleitenden Kapitel werden die Konzepte von drahtlosen Sensorknoten, elektronischer Energieversorgung und Schaltungsentwurf mit besonders niedriger Leistungsaufnahme zusammengefasst.

Das zweite Kapitel beinhaltet grundlegende Überlegungen bezüglich des Entwurfs von Linearspannungsreglern. Ein Kleinsignalmodell der Regelschleife wird präsentiert und Methoden für die Stabilitätsanalyse und Maße für die Regelgüte werden eingeführt. Das AC Verhalten von sowohl NMOS als auch PMOS Regeltransistoren wird mittels analytischer Rechnungen untersucht. Eine Methodik zur symbolischen Schaltungsanalyse wird eingeführt, welche die analytische Behandlung von größeren Schaltungen erlaubt. Diese Methodik wird angewandt um analytische Modelle der Versorgungsspannungsunterdrückung von einfachen CMOS OTAs herzuleiten.

Im dritten Kapitel wird der Entwurf des Spannungsreglers, basierend auf einem NMOS Regeltransistor beschrieben. Eine neue Schaltungstopologie für die Regelschleife wird präsentiert, welche die Ansteuerung des Regeltransistors für den Betrieb mit niedrigem Spannungsabfall ermöglicht. Die vorgeschlagene Topologie wird analytisch in Hinblick auf Stabilität, Versorgungsspannungsunterdrückung und Ausgangsimpedanz untersucht. Zusätzliche Schaltungsblöcke wie ein Oszillator, eine Ladungspumpe und eine Hochlaufschaltung werden beschrieben.

Das vierte Kapitel beinhaltet Simulationsergebnisse des kompletten Systems. Die Funktionalität wird anhand von Simulationen im Zeitbereich gezeigt und Parameterschwankungen werden berücksichtigt.

Schlussendlich wird im letzten Kapitel eine abschließende Feststellung der erzielten Resultate gegeben.

Acknowledgment

This work was created in the course of the master studies of Telematics at Graz University of Technology. It was conducted at Infineon Technologies in Graz at the department Contactless and RF Exploration (CRE).

I would like to thank Dipl.-Ing. Günter Hofer and Dipl.-Ing. Martin Wiessflecker for the kind support throughout my time at Infineon. I am very grateful for the opportunity to gain experience in the field of analog chip design. I want to thank the head of CRE, Dipl.-Ing. Gerald Holweg for giving me this opportunity.

At the university this thesis was supervised by Dipl.-Ing. Dr.tech. Mario Auer, whom I would like to thank for his support. The knowledge I have gained from his lectures helped me a lot throughout this work.

Finally and above all, I want to thank my mother, who always supported me, always believed in me and whom I owe so much.

Contents

1 Introduction

1.1 Wireless Sensor Nodes

Wireless sensor nodes are electronic devices with the ability to sense a physical quantity and to transmit the measured value to other electronic devices, without the need for electrical wiring between the devices. This significantly reduces installation costs and enables the deployment of large scale sensor networks.

Application Areas

Typical application areas are automotive and industrial, were the number of sensors used for monitoring and control is high. In some cases the wireless transmission of the measured data does not only save costs, but is also required by the application. For example in the case of tire pressure monitoring systems, the sensor node is integrated in the tire of a car. With a wired sensor such an installation would not be possible.

A further application area is structural health monitoring of airplanes, bridges, heavy machinery or buildings. A large number of sensors distributed over the object allows continuous monitoring of the stress that is applied to its functional parts. This knowledge can be used to prevent accidents or to indicate the need for replacement.

In the health care sector, wireless sensor nodes can be used to monitor the vital signs of patients. Even an implantation of the sensor is possible. For example patients suffering from diabetes would significantly benefit from an implanted blood-sugar monitor.

A long term vision for the application of wireless sensor nodes is formulated by the term *ambient intelligence*. The idea behind this is that the environment adapts itself autonomously to the needs of human beings. Wireless sensor nodes can act as the "eyes and ears" of a computational intelligence that supports people in their everyday life.

Technological Considerations

From a technological point of view, a wireless sensor node requires many functional blocks. Beside the sensor and the read-out electronics that together form a transducer, a data processing unit, a memory, a transceiver and the implementation of a network protocol are required. Furthermore the sensor node needs a source of electrical energy for its operation. This energy source can be implemented in three different ways. First the energy can come from a battery that is integrated on the wireless sensor node. Second it can be drained from the electromagnetic field that is emitted by a reader device. Third it can be converted from other forms of energy (e.g. vibrational, thermal, optical) that are available in the ambient of the sensor node. The latter is called *energy harvesting* and is the preferable choice for sensor nodes that require a very long life time. In any case, the available energy is a limiting factor for the size and the complexity of the sensor node.

1.2 Power Management

The term power management summarizes all measures that are taken to ensure an efficient use of electrical power in electronic circuits. Such measure can be located at the circuit level (e.g. the design of highly efficient voltage conversion systems), but they can also reach up to the system level. The holistic design approach suggests to look across block boundaries and to optimize systems at all levels of abstraction. A good example for this is power aware computing were the power consumption of computational tasks is considered already at the hardware/software co-design level. As a result, communication protocols can be optimized to require less CPU cycles, or several power-down-modes can be implemented to ensure minimal power consumption in idle mode.

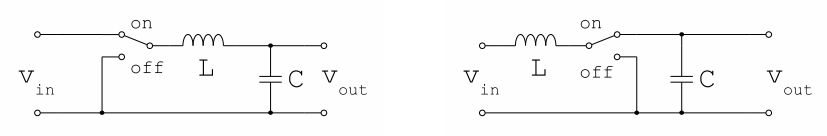
In the present work however only the circuit level will be considered. At this level the main task is to efficiently convert and/or regulate supply voltages. For this task there are basically two approaches: switched mode converters and linear regulators.

Switched Mode Converters

At switched mode converters, transistors are used as switches to periodically change the connectivity of energy storing devices. Based on the types of devices used, it can be differentiated between inductive/capacitive converters and purely capacitive converters.

As an example for the first kind, the Buck and the Boost converter architecture are given in Figures 1.1a and 1.1b. These circuits use an inductive/capacitive low-pass structure and a switch to control the current flow.

In one phase energy is stored in the magnetic field of the inductor and in the other phase, energy is transfered to the capacitor by the current driven by the inductor. The ratio of output to input voltage is set by the duty cycle of the switching process. At the Buck converter $v_{\text{out}}/v_{\text{in}}$ is equal to $T_{\text{on}}/(T_{\text{on}}+T_{\text{off}})$. This allows an output voltage that is lower than the input voltage. At the Boost converter $v_{\text{out}}/v_{\text{in}}$ is equal to $(T_{\text{on}}+T_{\text{off}})/T_{\text{on}}$, which allows a higher output voltage.



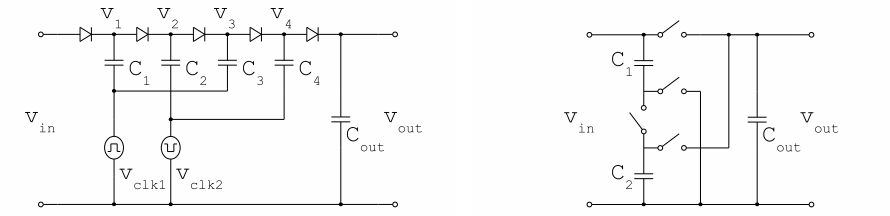
(a) Buck converter. (b) Boost converter.

Figure 1.1: Principles of inductive/capacitive power converters.

With inductive/capacitive switched mode converters, high efficiencies can be achieved and large load currents can be supplied. For low load currents however, the efficiency decreases. Another disadvantage is the fact that inductors of sufficient size cannot be integrated. Especially for wireless sensor nodes, a fully integrated solution is desirable. Because of this, inductive/capacitive converters are not considered in this work.

Purely capacitive converters, also called charge pumps, can be fully integrated. The common architecture of a Dickson charge pump is depicted in Figure 1.2a. It consists of a number of stages, each composed of a diode and a capacitor. When the clock at the bottom plate of the capacitor is low, the capacitor is charged by the preceding stage. When the clock changes to a level of v_{in}, also the potential at the top plate is increased by v_{in}. This way, every stage increases the voltage by v_{in}, leading to $4\,v_{\text{in}}$ at the output.

On the other hand also a lower output voltage can be achieved by an architecture like the one depicted in Figure 1.2b. The switches alternately connect the capacitors in a serial configuration to the input and in a parallel configuration to the output. This way, the input voltage is divided by a factor of 2.



(a) Dickson charge pump. (b) Voltage divider charge pump.

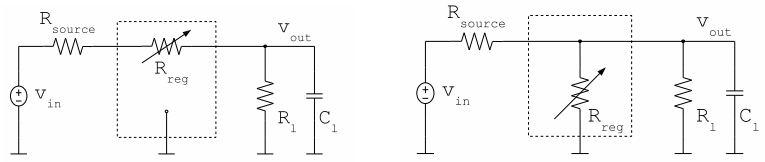
Figure 1.2: Principles of capacitive power converters.

The efficiency of charge pumps depends on the dissipative losses in the diodes or switches. Switches dissipate less energy compared to diodes, because they have a lower on-resistance. The remaining losses depend on the voltage drop across the switch and the current flow through the switch. The amount of current is mainly defined by the external load. The amount of voltage drop is defined by the current, the capacitor sizes and the switching frequency. High efficiencies can be achieved for low to medium load currents. Larger load currents require larger capacitor sizes and higher switching frequencies. Common switching frequencies are in the kilohertz range, up to several megahertz.

Linear Regulators

In contrast to switched mode converters, linear regulators do not convert but dissipate energy. This way only an output voltage which is lower than the input voltage can be achieved. The resulting voltage drop is called the dropout voltage [1].

For the implementation of a linear regulator there are basically two ways: The excess of input voltage can be dissipated by a pass device in series with the load (Figure 1.3a), or there can be an additional device in parallel with the load (Figure 1.3b). The series regulator increases the overall resistance which gives a better efficiency. With the parallel (also called shunt-) regulator the overall resistance can only be decreased. This way a lot of energy is wasted. Shunt regulators are therefore only suitable for applications like over-voltage protection.



(a) Series regulator. (b) Shunt regulator.

Figure 1.3: Principles of linear voltage regulators.

If series voltage regulators are compared to charge pumps, the main advantages are the high load current ability and the fast regulation. However because of the dissipative nature of linear regulators, they suffer from bad efficiency at high dropout voltages.

A possible solution for this problem is to cascade a charge pump with a linear regulator. This way the input voltage is converted efficiently to a level slightly above the output voltage and is then regulated by the linear regulator. The additional charge pump however adds complexity. If large variations of the input voltage are expected, also a regulation mechanism for the charge pump is needed.

1.3 Ultra Low Power Design

The need for electronic circuits which require very little power for their operation first came up with the advent of electronic wrist watches. In this context a long duration with a single battery is desirable. Similar applications are electronic pacemakers, hearing aids and of course wireless sensor nodes. All these devices require a power dissipation in the range of several hundred microwatt or below. To achieve this goal, low operation voltages and very low bias currents are used.

The use of low operation voltages helps to decrease the dynamic power consumption in digital circuits. Since the energy stored in parasitic capacitors is equal to $v_{\mathrm{dd}}^2\, C/2$, a lower operation voltage v_{dd} reduces the power dissipation quadratically.

For analog blocks, the power dissipation can mainly be reduced by a reduction of bias currents. With ultra low bias currents in the range of several nanoampere, the MOS transistors are operated in the weak inversion region. Equations 1.1 and 1.2 give a comparison between strong and weak inversion [2].

Strong Inversion	Weak Inversion	
$I_{\mathrm{D}} = \dfrac{\mu\, C_{\mathrm{ox}}}{2}\, \dfrac{W}{L}\, (V_{\mathrm{GS}} - V_{\mathrm{th}})^2$	$I_{\mathrm{D}} = I_{\mathrm{D0}}\, \dfrac{W}{L}\, \mathrm{e}^{\frac{V_{\mathrm{GS}} - V_{\mathrm{th}}}{n\, \Phi_{\mathrm{T}}}}$	(1.1)
$\dfrac{g_{\mathrm{m}}}{I_{\mathrm{D}}} = \dfrac{2}{V_{\mathrm{GS}} - V_{\mathrm{th}}}$	$\dfrac{g_{\mathrm{m}}}{I_{\mathrm{D}}} = \dfrac{1}{n\, \Phi_{\mathrm{T}}}$	(1.2)

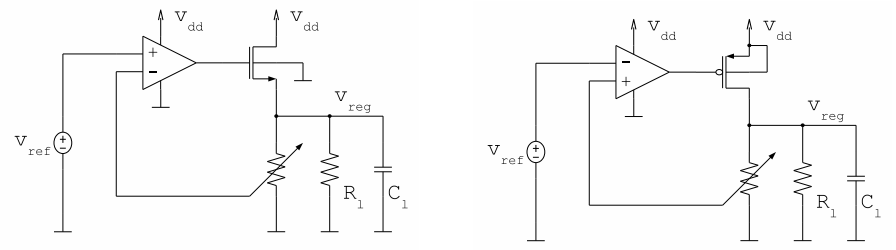
In the case of strong inversions, the drain current I_{D} shows a quadratic dependence on the overdrive voltage $V_{\mathrm{GS}} - V_{\mathrm{th}}$, whereas in the case of weak inversion the dependence becomes exponential. As a consequence the ratio of transconductance over drain current $g_{\mathrm{m}}/I_{\mathrm{D}}$ also shows a different behavior. For strong inversion it depends inversely on the overdrive voltage, but for weak inversion it does not depend on the operating point of the transistor any more. For weak inversion it is defined by the thermal potential Φ_T and the subthreshold slope factor n. A typical value for $g_{\mathrm{m}}/I_{\mathrm{D}}$ at strong inversion is 10 V^{-1}, whereas at weak inversion it is much closer to the value of a bipolar transistor of $1/\Phi_{\mathrm{T}}$ which is approximately 38 V^{-1} at room temperature.

From the values of $g_{\mathrm{m}}/I_{\mathrm{D}}$ it can be concluded that the bias current is most effectively used if the transistor is operated in weak inversion. On the other hand, in weak inversion the ratio of $g_{\mathrm{m}}/I_{\mathrm{D}}$ is defined by a constant that is independent of transistor dimensions. Because of this, only relatively low values of g_{m} can be achieved with low bias currents. This implies a low transition frequency and only permits the operation at relatively low frequencies. Another factor that prevents the operation at high frequencies is the output conductance g_{ds} which linearly depends on the bias current. Low bias currents lead to low output conductances, which means high ohmic nodes, or low frequency poles.

2 Fundamentals

2.1 Linear Voltage Regulators

A linear voltage regulator is composed of a pass device, a feedback network, an error amplifier and a voltage reference. Figure 2.1 shows basic circuits, featuring an NMOS pass device (Figure 2.1a) and a PMOS pass device (Figure 2.1b).



(a) NMOS pass device. (b) PMOS pass device.

Figure 2.1: Basic types of linear voltage regulators.

For the design of the control loop, the pass device acts as the plant and the error amplifier acts as the regulator. Therefore the choice whether to use an NMOS or a PMOS transistor as pass device has important consequences for the design of the control loop. The NMOS transistor acts as a source-follower (i.e. a buffer) which has unity gain and a low output impedance. The PMOS transistor acts as a common-source amplifier with inverting gain and high output impedance. From these considerations it can be concluded that the option with the NMOS pass device is easier to compensate. In this case the dominating pole is located at the error amplifier and the non-dominating pole is located at the pass device. Because of the low output impedance of the pass device, a sufficient separation of the two poles can be achieved quite easily. With the PMOS pass device, the dominating pole can be located either at the error amplifier or at the pass device. Since the latter is also dependent on the load, the compensation becomes more difficult.

Beside the AC behavior of the two pass device options, also the DC operating point needs to be considered. With the NMOS pass device the gate voltage needs to be biased above the output voltage. This is only possible if the supply of the error amplifier is also above the output voltage. At low dropout operation the input voltage approaches the output

voltage and there is not enough voltage headroom to bias the gate of the pass device. Therefore low dropout operation with an NMOS pass device is only possible, if the input voltage is pumped up internally by a charge pump. With a PMOS pass device, there is no problem concerning low dropout operation.

The efficiency of a linear voltage regulator is given by Equation 2.1. From this equation it can be seen that the efficiency depends on the two ratios of output to input voltage and load current to overall current. To maximize the first ratio, low dropout operation is needed. To maximize the second ratio, the quiescent current needs to be minimized.

$$\eta = \frac{P_{\text{out}}}{P_{\text{in}}} = \frac{v_{\text{out}}}{v_{\text{in}}} \frac{i_{\text{l}}}{i_{\text{l}} + i_{\text{q}}} \tag{2.1}$$

2.1.1 System Model

An integrated circuit composed of MOS transistors is an inherently nonlinear system. For the mathematical description however it is often useful to consider a linearized version of the system, at a specific operating point. This way, an abstract model of the voltage regulator can be set up and considerations from basic control system theory can be applied. In Figure 2.2 the voltage regulator is modeled by individual blocks representing transfer functions of the linearized system.

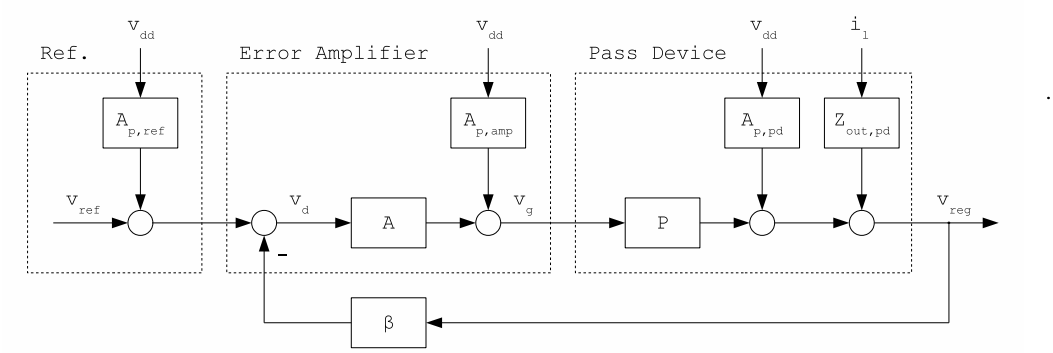


Figure 2.2: Model of the control loop.

The model from Figure 2.2 includes the reference voltage v_{ref}, the error amplifier A, the pass device P and the feedback network β. Additionally to the signal path, also the influence of the power supply is modeled. The transfer functions $A_{\text{p,ref}}$, $A_{\text{p,amp}}$ and $A_{\text{p,pd}}$ model the path from the supply voltage v_{dd} to the output of the respective circuits. The influence of varying load currents is modeled by the output impedance Z_{out} which also can be interpreted as a transfer function.

For the derivation of closed loop transfer functions, the well known formula from control system theory in Equation 2.2 is used. The denominator $1 + A_{\text{open loop}}$ is called the effectiveness of the control loop. For frequencies below the unity gain frequency of the open loop, this term can be approximated by the open loop gain $A_{\text{open loop}}$. Above the unity gain frequency, the control loop looses its effectiveness and the closed loop transfer function $A_{\text{closed loop}}$ can be approximated by the direct path $A_{\text{direct path}}$.

$$A_{\text{closed loop}} = \frac{A_{\text{direct path}}}{1 + A_{\text{open loop}}} \tag{2.2}$$

The command transfer function T from v_{ref} to v_{reg} is given by Equation 2.3. With the assumption of an effective control loop it is equal to the inverse of the voltage divider ratio β of the feedback network.

$$T = \frac{v_{\text{reg}}}{v_{\text{ref}}} = \frac{A\ P}{1 + A\ P\ \beta} \approx \frac{1}{\beta} \tag{2.3}$$

The overall power supply transfer function A_{p} is derived by the superposition of all signal paths from v_{dd} to v_{reg}. In detail this involves $A_{\text{p,ref}}$, $A_{\text{p,amp}}$ and $A_{\text{p,pd}}$ as shown in Equation 2.4. With the assumption of an effective control loop, the formula can be simplified to the expression given in Equation 2.5. It can be seen that the power supply gain from the voltage reference is not affected by the control loop. The power supply gain of the error amplifier is decreased by the gain of the error amplifier and the power supply gain of the pass device is decreased by the gain of the error amplifier and the gain of the pass device. The term $A_{\text{p,amp}}/A$ represents the power supply rejection ratio (PSRR) of the error amplifier and $A_{\text{p,pd}}/P$ represents the PSRR of the pass device.

$$A_p = \frac{v_{\text{reg}}}{v_{\text{dd}}} = \frac{A_{\text{p,ref}}\ A\ P}{1 + A\ P\ \beta} + \frac{A_{\text{p,amp}}\ P}{1 + A\ P\ \beta} + \frac{A_{\text{p,pd}}}{1 + A\ P\ \beta} \tag{2.4}$$

$$\approx \frac{1}{\beta}\left(A_{\text{p,ref}} + \frac{A_{\text{p,amp}}}{A} + \frac{1}{A}\frac{A_{\text{p,pd}}}{P}\right) \tag{2.5}$$

The output impedance is also affected by the operation of the control loop, as is shown in Equation 2.6. For an effective closed loop, the output impedance of the pass device is decreased by the loop gain.

$$Z_{\text{out}} = \frac{v_{\text{reg}}}{i_{\text{l}}} = \frac{Z_{\text{out,pd}}}{1 + A\ P\ \beta} \approx \frac{1}{\beta}\frac{Z_{\text{out,pd}}}{A\ P} \tag{2.6}$$

2.1.2 Stability

An important property of a voltage regulator is its stability. As has been mentioned before, circuits composed of MOS transistors are inherently nonlinear. At nonlinear systems in general, stability is not a system property, but a property of operating points. In practice, the stability of the system is therefore checked by an investigation of the stability of linearized models at different operating points. Additional to this, also transient simulations of step responses should be made.

A common definition of stability is the BIBO stability. The term BIBO stands for "bounded input - bounded output". If all excitations of the system with a bounded input lead to a bounded output, the system is called BIBO-stable. Based on the transfer function of a system, BIBO stability can be guaranteed, if all poles have a negative real part, i.e. all poles are located in the left half of the s-plane [3].

In practice however the transfer function of a closed loop control system is not known. And even if it would be known, it is difficult to predict the influence of changes that are applied to the open loop. A common approach to control system design is therefore to shape the frequency response of the open loop and to use a mathematical formalism to see how this influences the closed loop. The connection of open loop transfer function $L(s)$ and closed loop transfer function $T(s)$ is given in Equation 2.7.

$$L(s) = \frac{\mu(s)}{\nu(s)} \qquad T(s) = \frac{L(s)}{1 + L(s)} = \frac{\mu(s)}{\nu(s) + \mu(s)} \tag{2.7}$$

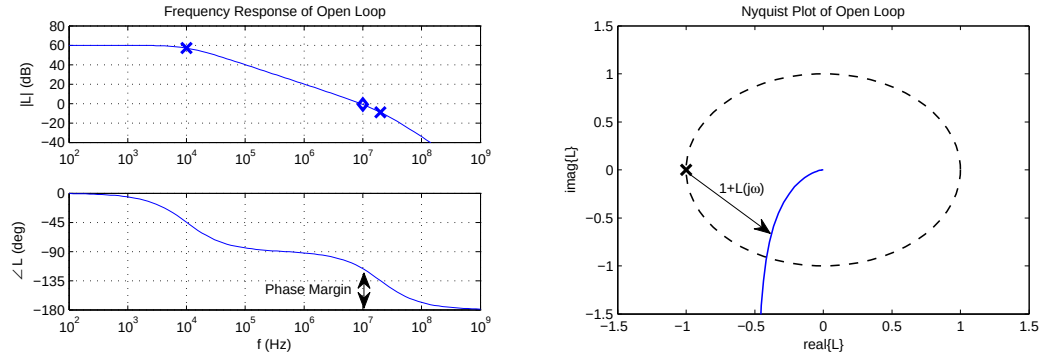
The Nyquist criterion can be used to determine the stability of the closed loop, based on the frequency response of the open loop $L(j\omega)$ and basic knowledge about the distribution of the poles of $L(s)$. In Equation 2.8 this criterion is given.

$$\Delta\text{arc}\{1 + L(j\omega)\} = (2n_{\text{right}} + n_{\text{axis}})\frac{\pi}{2} \quad \Longleftrightarrow \quad \text{Closed loop is BIBO stable} \tag{2.8}$$

The operator $\Delta\text{arc}\{\}$ determines the continuous phase change that results from an evaluation of the term in curly braces from $\omega = 0$ to $\omega = +\infty$. The variables n_{right} and n_{axis} stand for the number of poles of $L(s)$ that are located in the open right halfplane, or directly on the imaginary axis. To clarify the application of this criterion, a simple example with a two-stage-amplifier will be presented. Such an amplifier can be described by the open loop transfer function given in Equation 2.9.

$$L(s) = A_0 \frac{1}{(1 + s/\omega_{\text{p1}})(1 + s/\omega_{\text{p2}})} \tag{2.9}$$

The amplifier of the present example has a gain of 60 dB, a dominating pole at 10 kHz and a non-dominating pole at 20 MHz. In Figure 2.3 two representations of the open loop transfer function are given. Figure 2.3a shows the Bode plot and Figure 2.3b shows a clipped part of the Nyquist plot. At low frequencies, the magnitude of $L(j\omega)$ is 60 dB (in a linear scale 1000), and the angle is 0°. At the Nyquist plot which uses Cartesian coordinates, this point $1000 + j\,0$ is located at the far right side, beyond the visualized region. At higher frequencies, the magnitude and phase of $L(j\omega)$ decrease. In the Nyquist plot this results in a curve, which originates at the point $1000 + j\,0$ and spans clockwise to the point $0 + j\,0$. The last part of this arc is shown in the figure.



(a) Bode plot.

(b) Nyquist plot.

Figure 2.3: Frequency domain representations of the open loop.

To check the stability of the closed loop by use of the Nyquist criterion, first the number and type of poles of the open loop must be determined. In the present example the amplifier has two poles which are located at the left side of the s-plane. There is no pole at the right side ($n_{\text{right}} = 0$) and no pole on the imaginary axis ($n_{\text{axis}} = 0$). Therefore the Nyquist criterion can be formulated as $\Delta\text{arc}\{1 + L(j\omega)\} = 0$. To check if this requirement is fulfilled, the Nyquist plot of Figure 2.3b is used. If the point $-1 + j\,0$ is considered as the origin and an arrow is drawn from this origin to the curve of $L(j\omega)$, the continuous phase change of $1 + L(j\omega)$ can be determined. At $\omega = 0$ the arrow points horizontally to the right. With increased frequency, the arrow rotates clockwise which is mathematically defined as a negative angle. With further increased frequency, the arrow moves counterclockwise again and finally at $\omega = \infty$ it reaches the horizontal position. This means the continuous phase change is equal to zero and the requirement for BIBO stability of the closed loop is fulfilled.

As a measure for stability, the distance between $L(j\omega)$ and the point $-1 + j\,0$ can be identified. If the radius of the curve of $L(j\omega)$ is increased, at some point it would encircle the point $-1 + j\,0$. Then the continuous phase change of $1 + L(j\omega)$ would not be zero but $-2\,\pi$ and the closed loop would not be stable any more.

For open loop transfer functions of the "simple type", a simplified version of the Nyquist criterion can be formulated. A transfer function is of the "simple type" if the following requirements are fulfilled [3].

- It has lowpass character, i.e. $\lim_{s\to\infty} L(s) = 0$.
- It has positive gain and maximally one integrator.
- It is BIBO stable. (with the exception of maximally one integrator)
- There is only one intersection with the unity-gain line.

With the simplified Nyquist criterion, the stability of the closed loop can be checked, simply by determination of the phase margin. The phase margin is defined as the angular distance from $L(j\omega)$ to –180°, evaluated at the unity gain frequency, were the magnitude of $L(j\omega)$ is equal to one. If the phase margin is positive, the closed loop is BIBO-stable.

The simplified Nyquist criterion is the stability criterion which is most often used at the design of electronic circuits. One reason for this is its simplicity, but on the other hand also a connection between the open loop frequency response and the closed loop step response can be made. The phase margin PM is connected to the percentage overshoot PO of the step response and the unity gain frequency ω_0 is connected to the rise time t_r of the step response. For an estimation of the step response parameters, the formulas from Equation 2.10 can be used.

$$\mathrm{PM} + \mathrm{PO} \approx 70 \qquad\qquad \omega_0\, t_\mathrm{r} \approx 1.5 \tag{2.10}$$

For practical applications a phase margin of approximately 60° is desirable. In the case of a two-stage amplifier this requires a sufficient separation of the unity gain frequency and the non-dominating pole. In the previous example from Figure 2.3a the unity gain frequency is marked by a diamond and the poles are marked by crosses. The non-dominating pole at 20 MHz is well separated from the unity gain frequency at 10 MHz. This results in a phase margin of approximately 65°.

All considerations from above rely on the determination of the open loop frequency response. In the case of large scale control systems, this open loop frequency response is often determined by simply opening the connection of the feedback loop. In the case of electronic circuits however, this approach is not ideal since the opening of the feedback loop changes the load conditions and the operating points on both sides of the cut. For a reliable determination of the open loop frequency response, the Spectre circuit simulator provides a stability analysis simulation. This simulation is based on the mathematical formalism by Middlebrook and is described in [4].

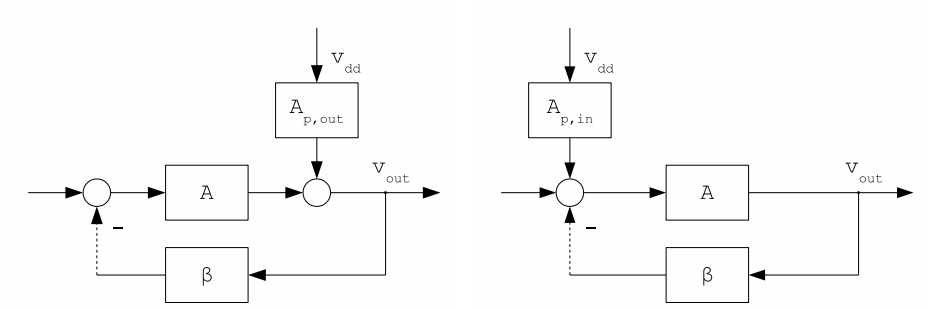
2.1.3 Performance Measures

The purpose of a linear regulator is to keep the output voltage constant, independent of variations of the input voltage and the load current. The insensitivity to input voltage variations is called line regulation. The insensitivity to load current variations is called load regulation. Both line- and load regulation can be specified as frequency responses.

Line Regulation

For the AC specification of line regulation, often the terms power supply rejection (PSR) and power supply rejection ratio (PSRR) are used. In the context of linear regulators, PSRR also stands for power supply ripple rejection. To avoid confusion, a closer look at the definitions of these terms will be given.

In Figure 2.4 two ways to model the coupling from the power supply of a closed loop amplifier are shown. In Figure 2.4a the coupling is modeled as a transfer function that adds directly to the output node (output-referred) and in Figure 2.4b it is modeled as a transfer function that adds to the input of the amplifier (input-referred).



(a) Output-referred. (b) Input-referred.

Figure 2.4: Definitions of power supply transfer functions.

The power supply rejection (PSR) is defined as the inverse of the output-referred power supply transfer function $A_{\mathrm{p,out}}$, given in dB as shown in Equation 2.11 [5]. The inverse is taken because $A_{\mathrm{p,out}}$ specifies a gain and PSR specifies a rejection. Higher values of PSR mean a better rejection of input voltage variations.

$$PSR = 20 \log \left(\frac{1}{A_{\mathrm{p,out}}} \right) \tag{2.11}$$

The power supply rejection ratio (PSRR) is defined as the ratio of signal gain A to output-referred power supply gain $A_{\mathrm{p,out}}$, given in dB as shown in Equation 2.12. The reason for this kind of definition is that the closed loop power supply transfer function is equal to $A_{\mathrm{p,out}}/(1 + A\ \beta)$. With an effective regulation this is approximately equal to $1/\beta \cdot A_{\mathrm{p,out}}/A$. Since most amplifiers are designed to be operated in closed loop the dominating performance measure is the PSRR, not the PSR.

$$PSRR = 20 \log \left(\frac{A}{A_{\mathrm{p,out}}} \right) \tag{2.12}$$

Another way to define the PSRR is to use the input referred power supply transfer function. If the block diagrams of Figure 2.4 are considered, it can be seen that an equivalent definition of the PSRR can be formulated as in Equation 2.13.

$$PSRR = 20 \log \left(\frac{1}{A_{\mathrm{p,in}}} \right) \tag{2.13}$$

If PSRR is translated to power supply ripple rejection, it refers to the closed loop PSR of a linear regulator [6]. As pointed out above, this is approximately equal to the power supply rejection ratio, multiplied by the feedback divider ratio β.

Load Regulation

As a measure for the load regulation, simply the closed loop output impedance can be used. Its definition is given in Equation 2.14.

$$Z_{\mathrm{out}} = \frac{v_{\mathrm{out}}}{i_{\mathrm{out}}} \tag{2.14}$$

Step Response

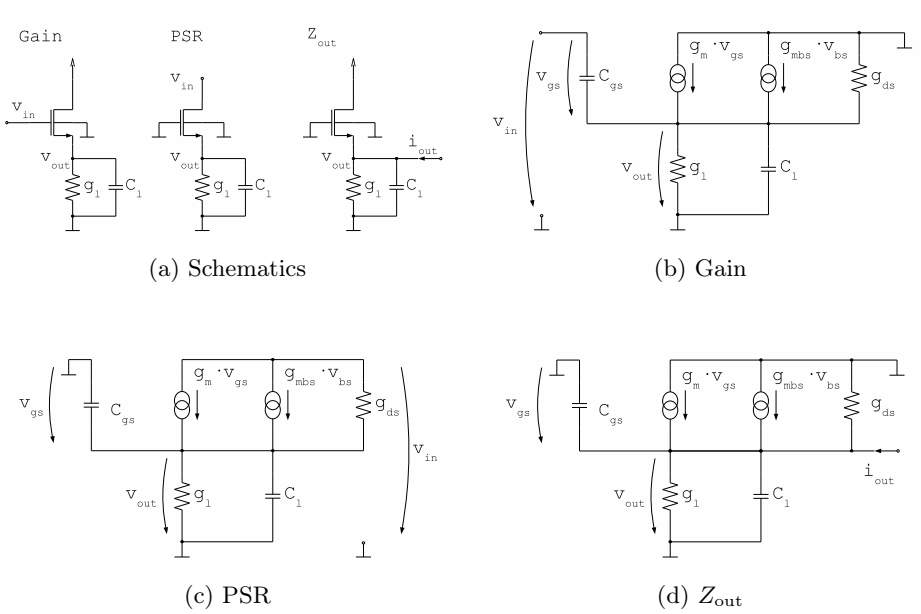
The AC characteristics presented above give a good insight in the functionality of the regulator, but on the other hand they are based on a linearized model. This means that nonlinearities are neglected and the large signal behavior can not be derived from it. A more application oriented approach is to look at the transient response of the regulated voltage, if an instant change of supply voltage or load current is applied.

2.2 Pass Device Characteristics

The AC behavior of the pass device is of great importance for the stability of the control loop, as well as for line- and load regulation performance. Because of this, the following section contains an investigation on the gain, the power supply rejection and the output impedance of both NMOS and PMOS pass device transistors. It should be noted however, that an isolated treatment of the pass device, can only be used for a basic comparison between NMOS and PMOS transistors. In a complete regulator circuit, the AC behavior of the pass device strongly depends on the surrounding circuitry.

2.2.1 NMOS Pass Device

Figure 2.5 contains the models that are used to investigate the AC behavior of an NMOS pass device. In Figure (a) the circuits for an evaluation of the gain, the power supply rejection and the output impedance are depicted. Figures (b) to (d) contain the small signal models of the respective circuits. In the small signal models, the transconductance g_{m} and the backgate transconductance g_{mbs} are modeled as current sources. The output conductance g_{ds} is modeled by a resistor. An ohmic/capacitive load is assumed at the output. Furthermore the parasitic capacitance between gate and source C_{gs} is included.



(a) Schematics

(b) Gain

(c) PSR

(d) Z_{out}

Figure 2.5: Small signal models of NMOS pass device.

Based on the nodal equation of the source node in the small signal models, the analytical expressions in Equations 2.15 to 2.17 have been derived. The signal gain A and the power supply gain A_p are defined as the ratio of output voltage v_out to input voltage v_in in the respective models. The output impedance Z_out was derived as the ratio of v_out to i_out in the small signal model of Figure 2.5d. The variable $g_\mathrm{out,N}$ is an abbreviation for the term in Equation 2.18. The value of C_gs has been neglected compared to C_l.

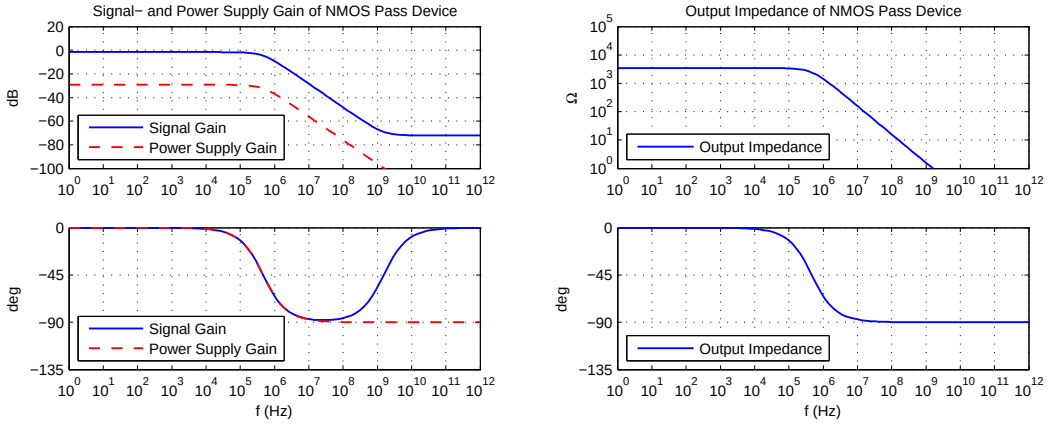
$$A = \frac{g_\mathrm{m}}{g_\mathrm{out,N}} \frac{1 + s\frac{C_\mathrm{gs}}{g_\mathrm{m}}}{1 + s\frac{C_\mathrm{l}}{g_\mathrm{out,N}}} \tag{2.15}$$

$$A_\mathrm{p} = \frac{g_\mathrm{ds}}{g_\mathrm{out,N}} \frac{1}{1 + s\frac{C_\mathrm{l}}{g_\mathrm{out,N}}} \tag{2.16}$$

$$Z_\mathrm{out} = \frac{1}{g_\mathrm{out,N}} \frac{1}{1 + s\frac{C_\mathrm{l}}{g_\mathrm{out,N}}} \tag{2.17}$$

$$g_\mathrm{out,N} = g_\mathrm{m} + g_\mathrm{mbs} + g_\mathrm{ds} + g_\mathrm{l} \tag{2.18}$$

In Figure 2.6, graphical evaluations of the analytical formulas from above are presented. From Figure (a) it can be seen that at low frequencies, the signal gain is approximately 0 dB and the power supply gain is approximately –30 dB. The output impedance in Figure (b) has a low frequency value of $1/g_\mathrm{out,N}$ and decreases at higher frequencies. The bandwidth of the system is defined by $g_\mathrm{out,N}/C_\mathrm{l}$.



(a) Signal- and power supply gain.

(b) Output impedance.

Figure 2.6: AC behavior of NMOS pass device.

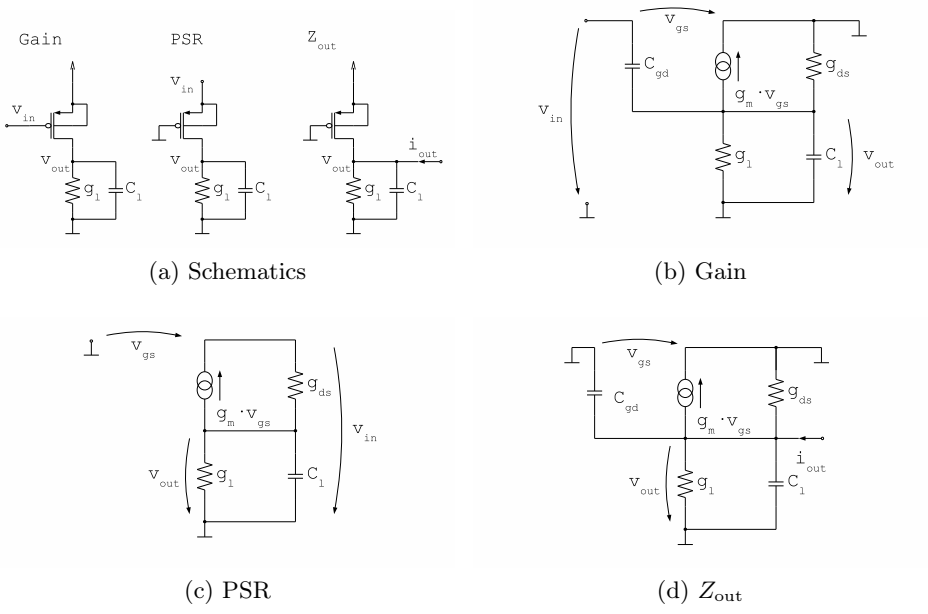
The small signal parameters used for the graphical evaluations of Figure 2.6 were obtained by a DC simulation in Spectre. The pass transistor dimensions were set to W/L = 40µm/400nm and the ohmic/capacitive load was choosen to be 10 µS and 100 pF. The input voltage was set to 1.2 V and the gate voltage was regulated for an output voltage of 1 V. The resulting small signal parameters are listed in Table 2.1.

Parameter	g_{m}	g_{mbs}	g_{ds}	C_{gs}
Value	242 µS	30 µS	10 µS	25 fF

Table 2.1: Small signal parameters of NMOS pass device.

2.2.2 PMOS Pass Device

Figure 2.7 shows small signal models of a PMOS pass device which are similar to the models presented for the NMOS pass device. The main difference is that no backgate transconductance has to be considered, because the bulk can be connected to the source of the transistor. As parasitic capacitance, C_{gd} is of relevance.



(a) Schematics

(b) Gain

(c) PSR

(d) Z_{out}

Figure 2.7: Small signal models of PMOS pass device.

Analytical expressions for signal gain, power supply gain and output impedance are given in Equations 2.19 to 2.21. The variable $g_{\mathrm{out,P}}$ is an abbreviation for the term in Equation 2.22. The value of C_{gd} has been neglected compared to C_{l} and the value of g_{ds} has been neglected compared to g_{m}.

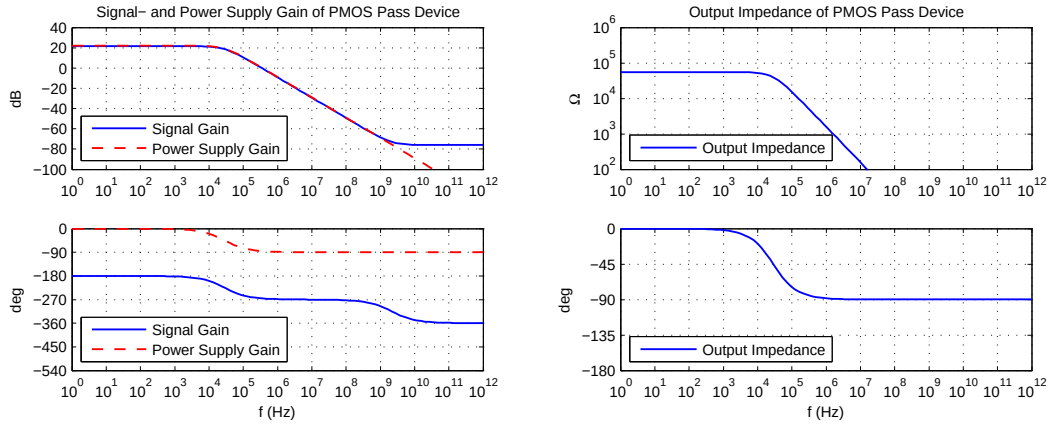
$$A = -\frac{g_{\mathrm{m}}}{g_{\mathrm{out,P}}} \frac{1 - s\frac{C_{\mathrm{gd}}}{g_{\mathrm{m}}}}{1 + s\frac{C_{\mathrm{l}}}{g_{\mathrm{out,P}}}} \tag{2.19}$$

$$A_{\mathrm{p}} = \frac{g_{\mathrm{m}}}{g_{\mathrm{out,P}}} \frac{1}{1 + s\frac{C_{\mathrm{l}}}{g_{\mathrm{out,P}}}} \tag{2.20}$$

$$Z_{\mathrm{out}} = \frac{1}{g_{\mathrm{out,P}}} \frac{1}{1 + s\frac{C_{\mathrm{l}}}{g_{\mathrm{out,P}}}} \tag{2.21}$$

$$g_{\mathrm{out,P}} = g_{\mathrm{ds}} + g_{\mathrm{l}} \tag{2.22}$$

Figure 2.8 contains graphical evaluations of the formulas from above. From Figure (a) it can be seen that both signal- and power supply gain have the same magnitude of approximately 20 dB. However the signal path shows an inverting behavior which can be seen at the phase response that starts at –180°. The output impedance in Figure (b) has a low frequency value of $1/g_{\mathrm{out,P}}$ and decreases at higher frequencies. The bandwidth is given by $g_{\mathrm{out,P}}/C_{\mathrm{l}}$.



(a) Signal- and power supply gain.

(b) Output impedance.

Figure 2.8: AC behavior of PMOS pass device.

The evaluations above are based on the same component values as with the NMOS transistor. The small signal parameters are given in Table 2.2.

Parameter	g_m	g_ds	C_gd
Value	218 µS	8 µS	16 fF

Table 2.2: Small signal parameters of PMOS pass device.

2.2.3 Comparison of Pass Devices

Table 2.3 contains a comparison of the AC characteristics of NMOS and PMOS pass devices. The most important difference between these two options is the output conductance, because all other characteristics depend on the output conductance. At the NMOS pass device it is dominated by the transconductance g_m. At the PMOS pass device it is composed of the drain-source conductance g_ds and the load conductance g_l. At low output currents g_ds is dominating, at high output currents g_l is dominating.

The NMOS pass device has approximately unity gain at the signal path and a power supply rejection that is defined by the intrinsic transistor gain $g_\mathrm{m}/g_\mathrm{ds}$. The PMOS pass device has equal gain at the signal- and at the power supply path. If the power supply rejection ratio is considered, this leads to $g_\mathrm{m}/g_\mathrm{ds}$ for the NMOS pass device and a value of 1 for the PMOS pass device. In other words, if the power supply rejection is considered, the NMOS pass device gives an advantage of one intrinsic transistor gain.

With respect to the bandwidth, the NMOS pass device has a higher value than the PMOS pass device. This makes the compensation of the control loop easier with an NMOS pass device. Both bandwidths are increased, if the load current is increased.

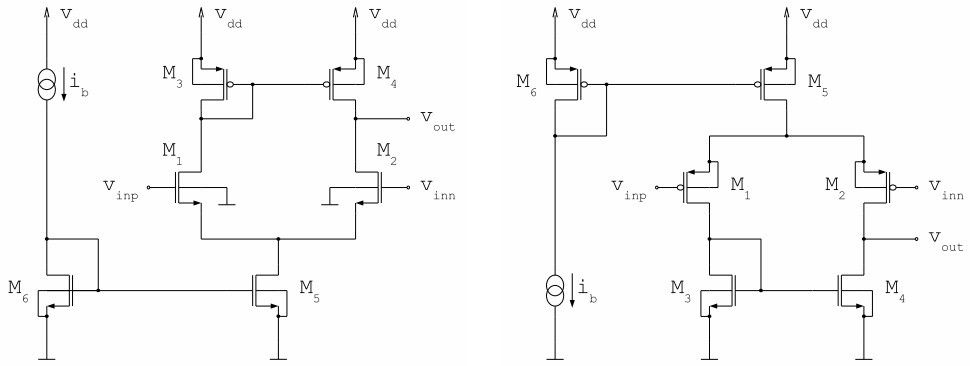
	NMOS	PMOS
Output Conductance	$g_\mathrm{m} + g_\mathrm{mbs} + g_\mathrm{ds} + g_\mathrm{l}$	$g_\mathrm{ds} + g_\mathrm{l}$
Signal Gain	$g_\mathrm{m}/g_\mathrm{out,N} \approx 1$	$-g_\mathrm{m}/g_\mathrm{out,P}$
Power Supply Gain	$g_\mathrm{ds}/g_\mathrm{out,N} \approx g_\mathrm{ds}/g_\mathrm{m}$	$g_\mathrm{m}/g_\mathrm{out,P}$
PSRR	$g_\mathrm{m}/g_\mathrm{ds}$	1
Bandwidth	$g_\mathrm{out,N}/C_\mathrm{l} \approx g_\mathrm{m}/C_\mathrm{l}$	$g_\mathrm{out,P}/C_\mathrm{l}$

Table 2.3: Comparison of pass device characteristics.

2.3 Error Amplifier

Beside the AC characteristics of the pass device, the AC behavior of the error amplifier is also of great importance for the performance of the linear regulator [7]. Its gain defines the effectiveness of the control loop at low frequencies and its bandwidth defines the frequency range where the control loop works effectively. The power supply rejection of the linear regulator at low frequencies is mainly defined by the power supply rejection ratio of the error amplifier. Beside of this, the power supply rejection ratio of amplifiers is also of importance for other applications.

The aim of the present section is to derive analytical models of the signal- and power supply gain of simple OTAs like depicted in Figure 2.9. Similar to the treatment of the pass device, a small signal model of the OTA is set up and the transfer functions of the signal- and the power supply path are derived analytically. In contrast to the small signal model of a pass device, the model of the OTA incorporates three nodes. This significantly increases the length and the complexity of the symbolic expressions. Since an evaluation of these long expressions by hand would be very time consuming and error prone, the Symbolic Math Toolbox of Matlab is used to solve the equations. But even if a solution is derived, the expression has no practical meaning, unless it is approximated to the most important terms. To illustrate the used methodology, first an abstract example of symbolic circuit analysis [8] is considered. After this, the models and analytical results for simple OTAs with NMOS and PMOS input devices are presented.



(a) NMOS input devices.

(b) PMOS input devices.

Figure 2.9: Schematics of simple OTAs.

2.3.1 Symbolic Circuit Analysis

A linear circuit with three unknown voltages can be described with an equation system like the one given in Equation 2.23 [9]. The variables v_1 to v_3 are the unknown voltages, v_{in} is a driving force applied to the circuit and y_{xx} are the conductances of the circuit. In the present case were the circuit is composed of resistors, capacitors and voltage-dependent current sources, the conductances have a form of $y = g + sC$.

$$\begin{bmatrix} +y_{11} & -y_{12} & -y_{13} \\ -y_{21} & +y_{22} & -y_{23} \\ -y_{31} & -y_{32} & +y_{33} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} y_1\, v_{\text{in}} \\ y_2\, v_{\text{in}} \\ y_3\, v_{\text{in}} \end{bmatrix} \tag{2.23}$$

The equation system can be set up, by writing down the nodal equations of nodes v_1 to v_3 individually and then combining them to a matrix equation. Alternatively it can be set up directly, by inspection of the circuit. The variables y_{11} to y_{33} in the main diagonal of the matrix contain all conductances that are connected to nodes v_1 to v_3, respectively. The off-diagonal entries of the matrix contain the conductances between individual nodes. The variables y_1 to y_3 on the right side of the equation contain the conductances between v_{in} and the respective node.

The solution for the unknown voltages can be obtained by multiplication with the inverted matrix from the left, as it is shown in Equation 2.24.

$$Y \cdot \underline{v} = \underline{i} \quad \Rightarrow \quad \underline{v} = Y^{-1} \cdot \underline{i} \tag{2.24}$$

The transfer function from v_{in} to the individual nodes is simply the obtained solution for the respective node, divided by v_{in}. As a simplification, v_{in} can be set equal to 1 from the beginning on, which eliminates the need for the last step.

The obtained transfer function has the form given in Equation 2.25. The coefficients b_3 to b_0 and a_3 to a_0 are abbreviations for relatively long terms, composed of the component values of the circuit.

$$H = \frac{b_3 s^3 + b_2 s^2 + b_1 s + b_0}{a_3 s^3 + a_2 s^2 + a_1 s + a_0} \tag{2.25}$$

For an interpretation of the obtained result, it is desirable to bring the transfer function to the form of equation 2.26. With this form, the gain A_0 and the pole/zero locations can be read out directly.

$$\tilde{H} = A_0 \frac{(1 + s/\omega_{z1})\,(1 + s/\omega_{z2})\,(1 + s/\omega_{z3})}{(1 + s/\omega_{p1})\,(1 + s/\omega_{p2})\,(1 + s/\omega_{p3})} \tag{2.26}$$

An exact mapping from H to $\tilde{H}$ would require to find the roots of the numerator and denominator in Equation 2.25. To avoid this, an extended version of dominant-pole approximation is used. By this approximation, the gain and pole/zero locations can be calculated with Equations 2.27 to 2.30.

$$A_0 = b_0/a_0 \tag{2.27}$$

$$\omega_{z1} = b_0/b_1 \qquad \omega_{p1} = a_0/a_1 \tag{2.28}$$

$$\omega_{z2} = b_1/b_2 \qquad \omega_{p2} = a_1/a_2 \tag{2.29}$$

$$\omega_{z3} = b_2/b_3 \qquad \omega_{p3} = a_2/a_3 \tag{2.30}$$

2.3.2 Simple OTA with NMOS Input

Figure 2.10 shows the small signal model that is used to analyze the AC behavior of a simple OTA with NMOS input devices. The indices of the variables follow the transistor numeration scheme that was introduced in Figure 2.9.

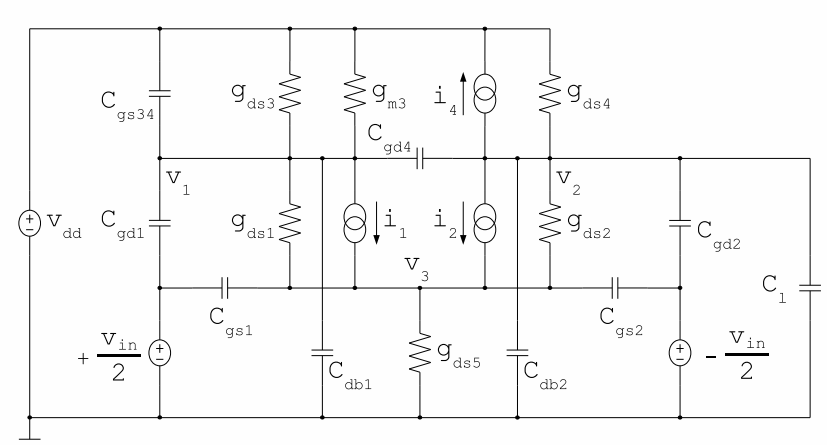


Figure 2.10: Small signal model of simple OTA with NMOS input.

The input devices M1 and M2 are modeled as a combination of voltage dependent current sources ($i_1 = g_{\mathrm{m1}}\, v_{\mathrm{gs1}}$ and $i_2 = g_{\mathrm{m2}}\, v_{\mathrm{gs2}}$) and output conductances (g_{ds1} and g_{ds2}). The current mirror which is composed of the PMOS transistors M3 and M4 is modeled in the same way. However the voltage-dependent current source of M3 is replaced by a conductance with a value of g_{m3}, since at M3, the gate-source and the drain-source voltage are the same. The current-sink composed of transistor M5 is modeled by its output conductance g_{ds5}. The transconductance g_{m5} is not included in the model, since the gate of M5 can be assumed as AC-ground.

The model includes the parasitic gate-source and gate-drain capacitances of the input devices ($C_{\mathrm{gs1}}/C_{\mathrm{gs2}}$ and $C_{\mathrm{gd1}}/C_{\mathrm{gd2}}$). The gate-source capacitances of M3 and M4 are summarized as C_{gs34}. The gate-drain capacitance of M4 is called C_{gd4} and models the capacitive coupling between nodes v_1 and v_2. The capacitive coupling of v_1/v_2 to ground is modeled by the drain-bulk capacitances C_{db1} and C_{db2}. At the output of the OTA which is node v_2, a load capacitance C_l is assumed.

As driving forces there are the supply voltage v_{dd} and the input voltage v_{in}. The latter is divided symmetrically into $+v_{\mathrm{in}}/2$ at the positive input and $-v_{\mathrm{in}}/2$ at the negative input. This way the differential gain of the OTA is observed, and the common mode gain is suppressed. In many applications however, one input of the OTA is at DC and the other input is driven by an AC signal. In this case the common mode gain would be relevant and the AC behavior is slightly different.

A mathematical formulation of the small signal model is given in the equations below. The right side of Equation 2.31, depends on the driving force that is considered. If the signal gain should be evaluated, $\underline{i_{\mathrm{src}}}$ is replaced by $\underline{i_{\mathrm{sig}}}$. If the power supply gain should be evaluated $\underline{i_{\mathrm{src}}}$ is replaced by $\underline{i_{\mathrm{pwr}}}$. The capacitances of nodes v_1 to v_3 are summarized in the variables C_{n1} to C_{n3}.

$$\begin{bmatrix} g_{\mathrm{m3}}+g_{\mathrm{ds3}}+g_{\mathrm{ds1}}+sC_{\mathrm{n1}} & -sC_{\mathrm{gd4}} & -(g_{\mathrm{m1}}+g_{\mathrm{ds1}}) \\ g_{\mathrm{m4}}-sC_{\mathrm{gd4}} & g_{\mathrm{ds2}}+g_{\mathrm{ds4}}+sC_{\mathrm{n2}} & -(g_{\mathrm{m2}}+g_{\mathrm{ds2}}) \\ -g_{\mathrm{ds1}} & -g_{\mathrm{ds2}} & g_{\mathrm{m1}}+g_{\mathrm{m2}}+g_{\mathrm{ds1}}+g_{\mathrm{ds2}}+g_{\mathrm{ds5}}+sC_{\mathrm{n3}} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \underline{i_{\mathrm{src}}} \tag{2.31}$$

$$\underline{i_{sig}} = \begin{bmatrix} \frac{1}{2}(-g_{\mathrm{m1}}+sC_{\mathrm{gd1}})v_{\mathrm{in}} \\ \frac{1}{2}(+g_{\mathrm{m2}}-sC_{\mathrm{gd2}})v_{\mathrm{in}} \\ \frac{1}{2}s(C_{\mathrm{gs1}}-C_{\mathrm{gs2}})v_{\mathrm{in}} \end{bmatrix} \qquad \underline{i_{\mathrm{pwr}}} = \begin{bmatrix} (g_{\mathrm{m3}}+g_{\mathrm{ds3}}+sC_{\mathrm{gs34}})v_{\mathrm{dd}} \\ (g_{\mathrm{m4}}+g_{\mathrm{ds4}})v_{\mathrm{dd}} \\ 0 \end{bmatrix} \tag{2.32}$$

$$C_{\mathrm{n1}} = C_{\mathrm{gs34}} + C_{\mathrm{db1}} + C_{\mathrm{gd1}} + C_{\mathrm{gd4}} \tag{2.33}$$

$$C_{\mathrm{n2}} = C_{\mathrm{l}} + C_{\mathrm{db2}} + C_{\mathrm{gd2}} + C_{\mathrm{gd4}} \tag{2.34}$$

$$C_{\mathrm{n3}} = C_{\mathrm{gs12}} \tag{2.35}$$

The mathematical model from above has been solved with the Symbolic Math Toolbox in Matlab, as it was described in the previous section. The resulting formulas for the gain and pole/zero locations are given in the equations below. The presented formulas are extensively simplified versions of the symbolic results. For simplification, the assumption $g_\mathrm{m} \gg g_\mathrm{ds}$ has been used. In Figure 2.11 the analytical formulas are compared to simulation results from Spectre. The plots are based on an OTA with input devices of size 2µm/1µm, all other devices of size 1µm/2µm, and a bias current of 10 nA.

Signal Gain

$$A_0 = \frac{g_\mathrm{m12}}{g_\mathrm{ds2} + g_\mathrm{ds4}} \tag{2.36}$$

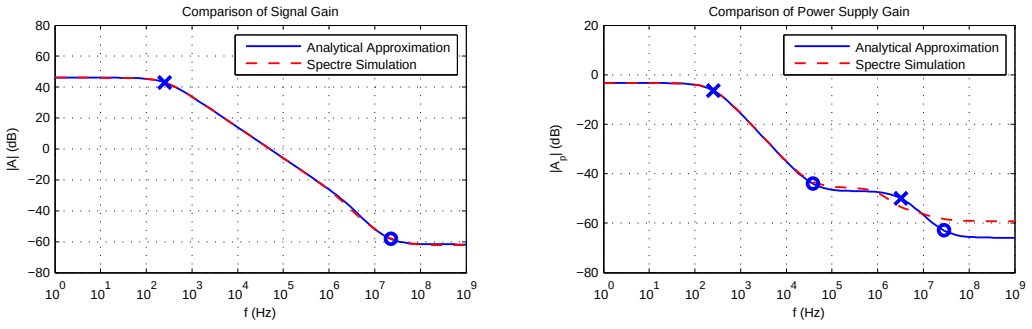
$$\omega_\mathrm{p1} = \frac{g_\mathrm{ds2} + g_\mathrm{ds4}}{C_\mathrm{n2}} \qquad \omega_\mathrm{z1} = \frac{g_\mathrm{m2}}{C_\mathrm{gd2}} + 2\,\frac{g_\mathrm{m3}}{C_\mathrm{n1}} \tag{2.37}$$

Power Supply Gain

$$A_0 = \frac{g_\mathrm{ds1} + g_\mathrm{ds4}}{g_\mathrm{ds2} + g_\mathrm{ds4}} \approx 1 \tag{2.38}$$

$$\omega_\mathrm{p1} = \frac{g_\mathrm{ds2} + g_\mathrm{ds4}}{C_\mathrm{n2}} \qquad \omega_\mathrm{z1} = \frac{g_\mathrm{ds1} + g_\mathrm{ds4}}{C_\mathrm{gd1} + C_\mathrm{db1} + 2\,C_\mathrm{gd4}} \tag{2.39}$$

$$\omega_\mathrm{p2} = \frac{g_\mathrm{m3}}{C_\mathrm{n1}} \qquad \omega_\mathrm{z2} = \frac{g_\mathrm{m3}}{C_\mathrm{gs34}}\,\frac{C_\mathrm{gd1} + C_\mathrm{db1} + 2\,C_\mathrm{gd4}}{C_\mathrm{gd4}} \tag{2.40}$$



(a) Signal gain. (b) Power supply gain.

Figure 2.11: Signal- and power supply gain of simple NMOS OTA.

2.3.3 Simple OTA with PMOS Input

Similar to the analysis of the OTA with NMOS input devices, also an OTA with PMOS input devices was analyzed. Figure 2.12 shows the associated small signal model. The mathematical formulation of the model is given below.

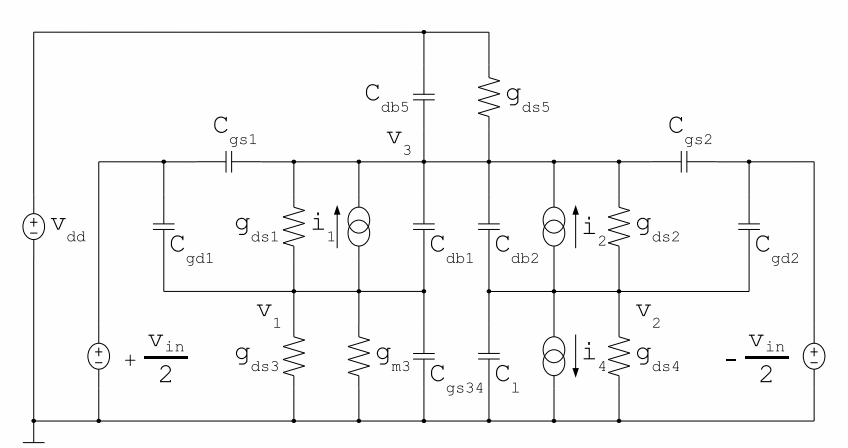


Figure 2.12: Small signal model of simple OTA with PMOS input.

$$\begin{bmatrix} g_{\mathrm{m3}}+g_{\mathrm{ds3}}+g_{\mathrm{ds1}}+sC_{\mathrm{n1}} & 0 & -(g_{\mathrm{m1}}+g_{\mathrm{ds1}}+sC_{\mathrm{db1}}) \\ g_{\mathrm{m4}} & g_{\mathrm{ds2}}+g_{\mathrm{ds4}}+sC_{\mathrm{n2}} & -(g_{\mathrm{m2}}+g_{\mathrm{ds2}}+sC_{\mathrm{db2}}) \\ -(g_{\mathrm{ds1}}+sC_{\mathrm{db1}}) & -(g_{\mathrm{ds2}}+sC_{\mathrm{db2}}) & g_{\mathrm{m1}}+g_{\mathrm{m2}}+g_{\mathrm{ds1}}+g_{\mathrm{ds2}}+g_{\mathrm{ds5}}+sC_{\mathrm{n3}} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \underline{i_{\mathrm{src}}} \tag{2.41}$$

$$\underline{i_{sig}} = \begin{bmatrix} \frac{1}{2}(-g_{\mathrm{m1}}+sC_{\mathrm{gd1}})v_{\mathrm{in}} \\ \frac{1}{2}(+g_{\mathrm{m2}}-sC_{\mathrm{gd2}})v_{\mathrm{in}} \\ \frac{1}{2}s(C_{\mathrm{gs1}}-C_{\mathrm{gs2}})v_{\mathrm{in}} \end{bmatrix} \qquad \underline{i_{\mathrm{pwr}}} = \begin{bmatrix} 0 \\ 0 \\ (g_{\mathrm{ds5}}+sC_{\mathrm{db5}})v_{\mathrm{dd}} \end{bmatrix} \tag{2.42}$$

$$C_{\mathrm{n1}} = C_{\mathrm{gs34}} + C_{\mathrm{db1}} + C_{\mathrm{gd1}} \tag{2.43}$$

$$C_{\mathrm{n2}} = C_{\mathrm{l}} + C_{\mathrm{db2}} + C_{\mathrm{gd2}} \tag{2.44}$$

$$C_{\mathrm{n3}} = C_{\mathrm{gs12}} + C_{\mathrm{db1}} + C_{\mathrm{db2}} + C_{\mathrm{db5}} \tag{2.45}$$

The evaluation of the mathematical model with respect to signal gain resulted in the same formulas as for the OTA with NMOS input devices. With respect to the power supply gain however a significantly different AC behavior can be observed. The obtained formulas for the gain and the pole/zero locations are given below.

Power Supply Gain

$$A_0 = \frac{1}{2}\frac{g_{\mathrm{ds5}}}{g_{\mathrm{ds2}}+g_{\mathrm{ds4}}}\left[\left(\frac{g_{\mathrm{m2}}+g_{\mathrm{ds2}}}{g_{\mathrm{m1}}+g_{\mathrm{ds1}}}-\frac{g_{\mathrm{m4}}}{g_{\mathrm{m3}}}\right)\frac{g_{\mathrm{m1}}+g_{\mathrm{ds1}}}{g_{\mathrm{m1}}}+\frac{g_{\mathrm{ds1}}+g_{\mathrm{ds3}}}{g_{\mathrm{m3}}}\right] \tag{2.46}$$

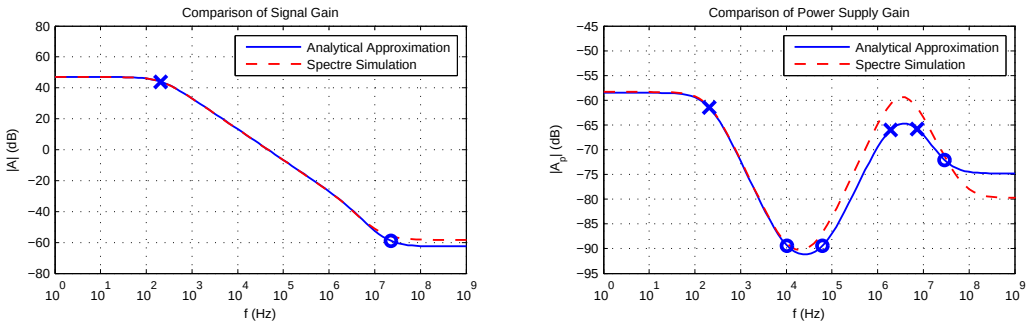
$$\omega_{\mathrm{p1}} = \frac{g_{\mathrm{ds2}}+g_{\mathrm{ds4}}}{C_{\mathrm{n2}}} \qquad \omega_{\mathrm{z1}} = \frac{g_{\mathrm{ds1}}+g_{\mathrm{ds3}}}{C_{\mathrm{n1}}} \tag{2.47}$$

$$\omega_{\mathrm{p2}} = \frac{g_{\mathrm{m1}}+g_{\mathrm{m2}}}{C_{\mathrm{n1}}+C_{\mathrm{n3}}} \qquad \omega_{\mathrm{z2}} = \frac{g_{\mathrm{ds1}}+g_{\mathrm{ds3}}}{C_{\mathrm{n1}}}+\frac{g_{\mathrm{ds5}}}{C_{\mathrm{db5}}} \tag{2.48}$$

$$\omega_{\mathrm{p3}} = \frac{g_{\mathrm{m1}}}{C_{\mathrm{n1}}}+\frac{g_{\mathrm{m2}}}{C_{\mathrm{n3}}} \qquad \omega_{\mathrm{z3}} = \frac{g_{\mathrm{m2}}}{C_{\mathrm{db2}}} \tag{2.49}$$

The formula for the DC-gain in Equation 2.46 involves many terms, but can be simplified extensively if perfect matching of the differential pair is assumed. With this assumption, the expression in the round brackets reduces to zero and the complete formula can be reduced to $A_0 = 1/2 \cdot g_{\mathrm{ds5}}/g_{\mathrm{m3}}$.

Figure 2.13 shows a comparison of the obtained model with simulations from Spectre. Again the same dimensioning as in the previous section was used. Compared to the OTA with NMOS input devices, the OTA with PMOS input devices shows a significantly higher power supply rejection.



(a) Signal gain.

(b) Power supply gain.

Figure 2.13: Signal- and power supply gain of simple PMOS OTA.

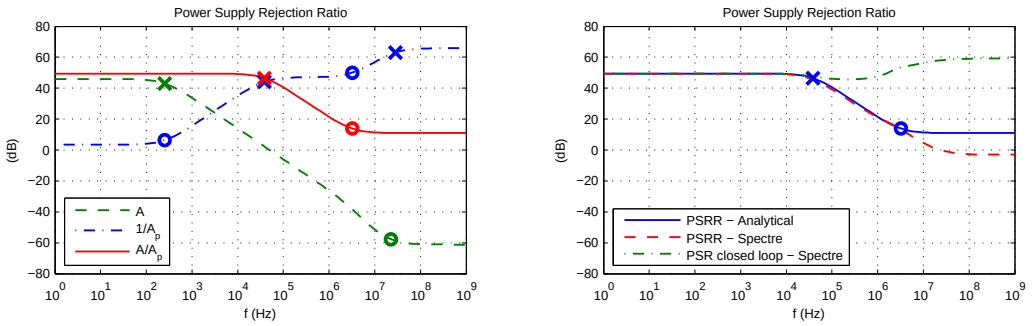
2.3.4 Power Supply Rejection Ratio

With the models for the signal gain and the power supply gain derived in the previous sections, also a model of the power supply rejection ratio can be set up. The PSRR is defined as the ratio of signal gain over power supply gain (A/A_p). In a logarithmic scale the division becomes a subtraction and the PSRR can be obtained by adding the magnitude of A with the negative magnitude of A_p.

OTA with NMOS input

In Figure 2.14a the construction of the PSRR is visualized for the case of the OTA with NMOS input devices. The dominating pole of the signal gain is compensated by the dominating pole of the power supply gain, which has become a zero because of the inversion. Also the high frequency zero of the signal gain is compensated by a zero of the power supply gain. What remains for the PSRR are the medium frequency pole and zero of the power supply gain.

In Figure 2.14b the model for the PSRR is compared to simulation results from Spectre. The PSRR from Spectre was obtained in the same way as the analytical PSRR, but based on AC simulations of A and A_p. The power supply rejection of the closed loop was obtained by a simulation of A_p in a closed loop configuration. At low frequencies, PSRR and the PSR of the closed loop are the same, since the closed loop power supply gain, calculated by $A_\mathrm{p}/(1+A)$ can be well approximated by A_p/A. Above the unity gain frequency of the amplifier this approximation is not valid any more, and the closed loop power supply gain approaches the open loop power supply gain A_p.



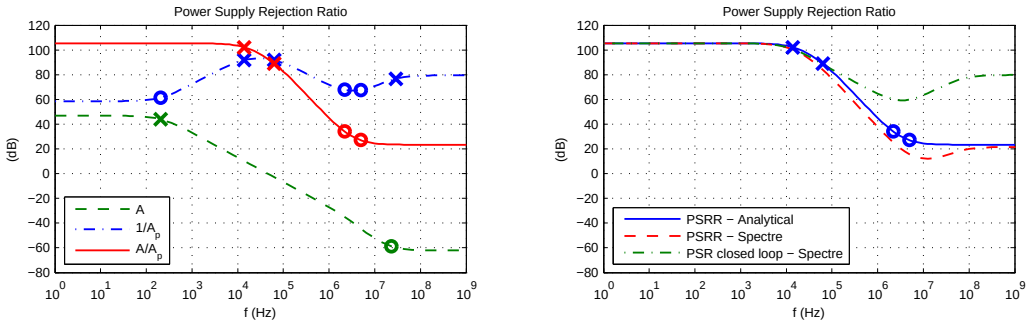
(a) Construction of PSRR.

(b) Comparison of PSRR.

Figure 2.14: Power supply rejection ratio of simple NMOS OTA.

OTA with PMOS input

In Figure 2.15a the PSRR of the OTA with PMOS input devices is presented. In this case, there are two medium frequency poles which lead to a steeper descent of the PSRR. In Figure (b) the analytical PSRR is compared to simulations from Spectre and the same observations can be made as in the previous case of the OTA with NMOS input.



(a) Construction of PSRR.

(b) Comparison of PSRR.

Figure 2.15: Power supply rejection ratio of simple PMOS OTA.

Comparison

A comparison of the OTAs with NMOS and PMOS input shows that the option with the PMOS input has a significantly better power supply rejection. The PSRR at low frequencies can be estimated by Equations 2.50 and 2.51.

$$\mathrm{PSRR}_{\mathrm{DC,NMOS}} = \frac{g_{\mathrm{m12}}}{g_{\mathrm{ds1}} + g_{\mathrm{ds4}}} \tag{2.50}$$

$$\mathrm{PSRR}_{\mathrm{DC,PMOS}} = \frac{g_{\mathrm{m12}}}{g_{\mathrm{ds2}} + g_{\mathrm{ds4}}} \frac{2\, g_{\mathrm{m3}}}{g_{\mathrm{ds5}}} \tag{2.51}$$

From these equations it can be seen that the DC-PSRR of the OTA with NMOS input is approximately equal to its DC-gain. At the OTA with PMOS input, it is equal to its DC-gain, multiplied with a factor of $2\, g_{\mathrm{m3}}/g_{\mathrm{ds5}}$. However it has to be noted that mismatch can reduce the good PSRR performance of the OTA with PMOS input.

3 Regulator Design

3.1 System Architecture

As it was shown in the previous chapter, the AC characteristics of an NMOS pass device are advantageous compared to a PMOS pass device. On the other hand, the NMOS pass device requires a more sophisticated biasing concept to allow low dropout operation. In general some kind of switched capacitor circuit is necessary with an NMOS pass device that generates a higher voltage level for the biasing of the pass device gate.

In the present work, the option with an NMOS pass device is chosen. This choice is motivated by the requirements of the application which are efficiency and robustness. The option with an NMOS pass device is advantageous for efficiency, because the high frequency output pole of the pass device allows a low quiescent current error amplifier. Of course it has to be assured, that this saving of current is not compensated by the additional circuitry which is needed for the biasing. The requirement of robustness is also easier achieved with an NMOS pass device, since a separation of the two open loop poles is facilitated by the high frequency pole of the pass device.

After the choice of the pass device, a biasing method has to be chosen. A simple approach is to use a charge pump for the generation of a higher voltage and to supply the error amplifier from this charge pump [10, 11]. The disadvantage of this biasing concept is that the charge pump decreases the overall efficiency. The more current is drained from the charge pump, the higher must be the clock frequency. A higher clock frequency increases the current consumption of the oscillator. To maximize the efficiency of the regulator, the current drained from the charge pump must be minimized.

As alternative, switched capacitor circuits have been proposed that do not require a charge pump for the biasing of the pass device [12, 13, 14]. In the following, two examples will be described. After that, the architecture that is proposed in this work will be presented.

Switched Capacitor Biasing

The biasing concept proposed in [13] avoids the use of a charge pump by the use of a capacitive level shifter. A schematic of this architecture is depicted in Figure 3.1. The biasing capacitor C_b is periodically recharged by a flying capacitor C_f which is in turn periodically recharged by a voltage source. The flying capacitor just needs to transport the charge that is lost because of leakage. Therefore it can be designed significantly smaller than the biasing capacitor. One difficulty with this architecture is that the switch between node v_2 and the flying capacitor must be able to operate at a level which is higher than the supply voltage. This requires the use of a boosted clock to drive the switch transistor.

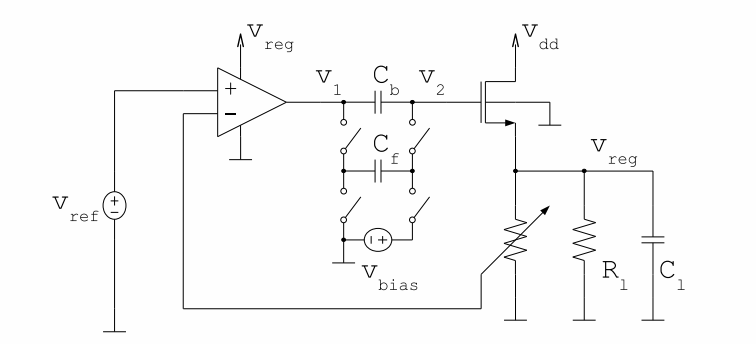


Figure 3.1: Architecture for NMOS LDO, proposed in [13].

A similar biasing concept is proposed in [14]. There, two equal sized biasing capacitors are used, as is shown in Figure 3.2. At each time instant one of them is used for biasing, and the other one is recharged by a voltage source. Periodically the roles of the two capacitors are interchanged. Basically this architecture would also require a boosted clock to drive the switches connected to the gate of the pass device. To avoid this, the switches are implemented by a special configuration of diode connected transistors.

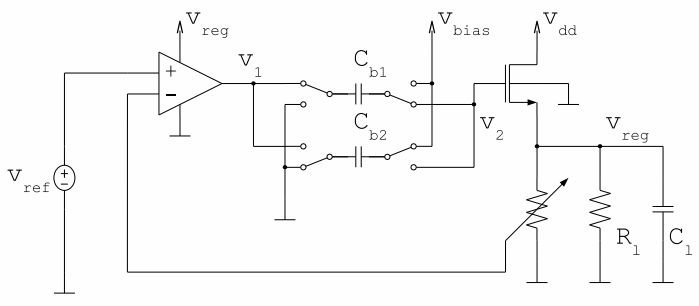


Figure 3.2: Architecture for NMOS LDO, proposed in [14].

Regulated Capacitor Biasing

In the present work a new architecture is proposed that combines a charge pump with a biasing capacitor. It is depicted in Figure 3.3 and uses a biasing amplifier to regulate the voltage across a capacitive level shifter.

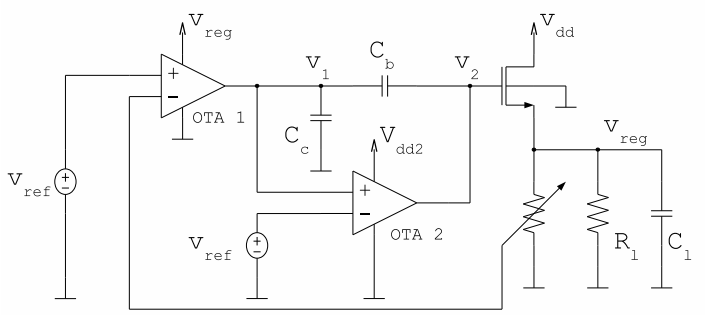


Figure 3.3: Architecture for NMOS LDO, proposed in this work.

The error amplifier (OTA 1) is supplied by the regulated voltage v_{reg} and the biasing amplifier (OTA 2) is supplied by a higher voltage v_{dd2} that is taken from a charge pump. The big advantage of this architecture, compared to the architecture were the charge pump supplies the error amplifier, is that here only a minimal amount of current is needed for the biasing amplifier. The majority of the quiescent current is used for the error amplifier and its efficiency is not decreased by the charge pump.

The output signal of the error amplifier is AC-coupled to the gate of the pass device by the biasing capacitor C_{b}. Additionally it is level-shifted by the voltage across the biasing capacitor. The amount of this level-shift is set by the biasing amplifier. This amplifier compares v_1 with v_{ref} and regulates v_2 to a level, such that the difference of the two input voltages becomes zero. If v_1 is larger than v_{ref}, the biasing amplifier increases v_2. This increases v_{reg} and also increases the feedback voltage that is sensed by the negative input of the error amplifier. As a result, v_1 becomes smaller.

In the present design, a reference voltage of approximately 600 mV is used. This voltage is taken from a low voltage bandgap reference like it is presented in [15]. For the feedback network a voltage divider consisting of three diode-connected PMOS transistors is used [16]. It implements a division ratio of 2/3 which leads to a regulated output voltage of approximately 900 mV.

For current saving reasons, both amplifiers are implemented as simple CMOS OTAs. The voltage levels require the use of NMOS input devices for both OTA 1 and OTA 2. At OTA 1 the input voltage headroom to the supply voltage is too small for PMOS input devices. At OTA 2 the output common mode level is higher than the input common mode level. This also requires NMOS input devices.

3.2 Control Loop

For the design of the control loop it is important to know, how the small signal parameters of the amplifiers and the dimensions of the capacitors are connected to the regulation performance. To investigate this connection, small signal models are set up and analytical formulas for the AC responses are derived. The following sections present results concerning the stability, the power supply rejection and the output impedance.

3.2.1 Stability

To analyze the stability of the control loop, a small signal model of the open loop was set up. It is depicted in Figure 3.4 and models the signal path from the negative input of the error amplifier $v_{\mathrm{in,n}}$ to the regulated output voltage $v_{\mathrm{reg}} = v_3$.

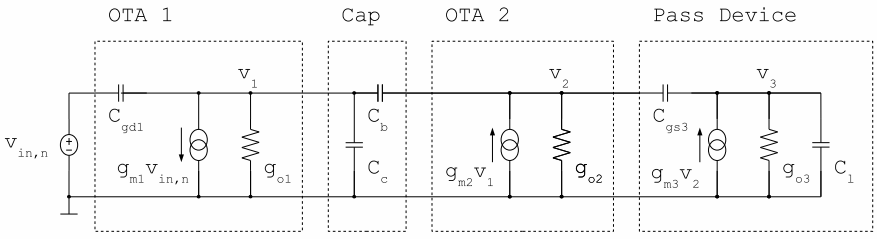


Figure 3.4: Small signal model of proposed LDO topology.

The differential amplifiers are modeled as combinations of voltage-dependent current sources and output conductances $g_{\mathrm{o}} = g_{\mathrm{ds2}} + g_{\mathrm{ds4}}$. The first amplifier shows an inverting gain of $-g_{\mathrm{m1}}/g_{\mathrm{o1}}$ and the second amplifier shows a non-inverting gain of $g_{\mathrm{m2}}/g_{\mathrm{o2}}$. The pass device is also modeled by a combination of voltage-dependent current source and output conductance, but here g_{o3} is composed by g_{m3}, g_{mbs3}, g_{ds3}, and g_{l}. This leads to a non-inverting gain which is smaller than one.

Equation 3.1 shows the mathematical model of the circuit from Figure 3.4. Like in the previous chapter, the Symbolic Math Toolbox from Matlab has been used to solve the equation system and to derive analytical formulas for the gain and pole/zero locations.

$$\begin{bmatrix} g_{\mathrm{o1}}+sC_{\mathrm{n1}} & -sC_{\mathrm{b}} & 0 \\ -(g_{\mathrm{m2}}+sC_{\mathrm{b}}) & g_{\mathrm{o2}}+sC_{\mathrm{n2}} & -sC_{\mathrm{gs3}} \\ 0 & -(g_{\mathrm{m3}}+sC_{\mathrm{gs3}}) & g_{\mathrm{o3}}+sC_{\mathrm{n3}} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} (-g_{\mathrm{m1}}+sC_{\mathrm{gd1}})\, v_{\mathrm{in,n}} \\ 0 \\ 0 \end{bmatrix} \tag{3.1}$$

The analytical model of the open loop response is given in Equations 3.2 to 3.5. Negative frequencies indicate that the pole/zero is located in the right half of the s-plane. Since there are two poles with negative frequencies, the open loop is not BIBO stable.

$$A_0 = \frac{g_{m1}}{g_{o1}} \frac{g_{m2}}{g_{o2}} \frac{g_{m3}}{g_{o3}} \tag{3.2}$$

$$\omega_{p1} = -\frac{g_{o1}}{C_b\, g_{m2}/g_{o2}} \qquad \omega_{z1} = \frac{g_{m2}}{C_b} \tag{3.3}$$

$$\omega_{p2} = -\frac{g_{m2}}{C_c + C_{gs3}(1 + C_c/C_b)} \qquad \omega_{z2} = -\frac{g_{m1}}{C_{gd1}} \tag{3.4}$$

$$\omega_{p3} = |\omega_{p2}| + \frac{g_{o3}}{C_{n3}} \qquad \omega_{z3} = \frac{g_{m3}}{C_{gs3}} \tag{3.5}$$

Figure 3.5 shows a comparison between the analytical model and simulations from Spectre. The graphical evaluation is based on a design with the parameters from Table 3.1. For the amplifiers and the pass device, the same dimensions as in chaper 2 were used.

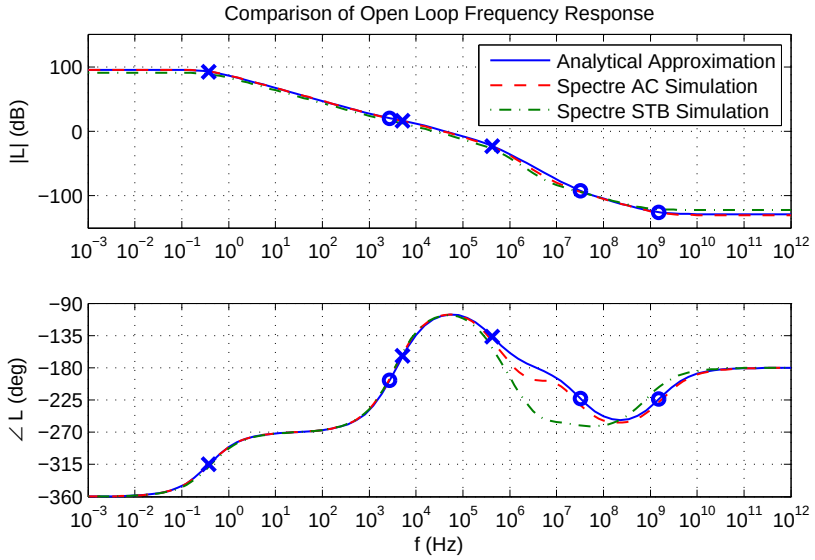


Figure 3.5: Bode plot of open loop.

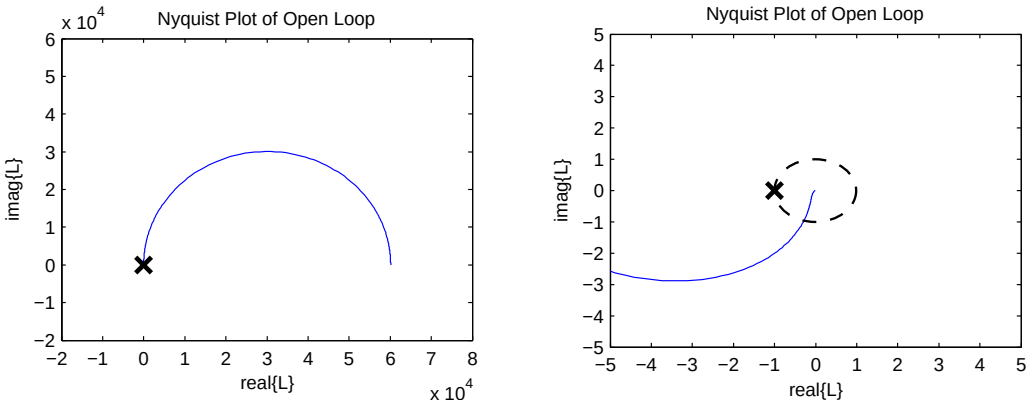
Parameter	C_b	C_c	C_l	$i_{b,OTA1}$	$i_{b,OTA2}$	i_l
Value	1 pF	500 fF	100pF	10 nA	1 nA	10 µA

Table 3.1: Parameters of control loop.

From the model above it can be seen that the open loop has a very high gain at low frequencies, caused by the cascade of the two differential amplifiers. The dominating pole ω_{p1} is located at a very low frequency, because the value of C_{b} is increased by the inner feedback loop of the biasing amplifier. Beside of the pole/zero doublet of ω_{p2} and ω_{z1}, the non-dominating pole ω_{p3} is of importance. The separation of this pole from the unity gain frequency mainly defines the phase margin.

For the simulation of the open loop in Spectre, attention must be paid to chose the correct simulation methodology. If an AC simulation is used, the loop is opened for AC signals and closed for DC signals. This results in the correct DC operating points, but neglects the changed load impedance at both sides of the opened loop. A further pitfall is the presence of additional feedback paths. In the present example, there is not only the main feedback path, but there is an additional path over the supply of the error amplifier. If the loop is closed, this additional path can be neglected, since the regulated voltage remains at a constant level. If the loop is opened however, the additional feedback path changes the AC response of the open loop. This means that for a correct simulation, the additional feedback path must be disabled by the use of an ideal voltage source for the error amplifier. A better choice for the determination of the open loop response is the stability analysis (STB) that is available in Spectre. With this simulation, the loop can remain closed and the open loop response is determined mathematically. For this purpose a current probe is added inside the loop. If multiple feedback loops exist, the probe must be added at a position where all loops are affected. In the present example, the probe has been added at the output of the error amplifier.

A question that remains to be answered is, if the simplified Nyquist criterion based on phase margin is applicable in the present case. To answer this questions, the Nyquist plot of the open loop is investigated. In Figure 3.6a, the complete Nyquist plot and in Figure 3.6b a zoomed version is shown.



(a) Complete area.

(b) Zoomed area.

Figure 3.6: Nyquist plot of open loop.

The open loop transfer function $L(s)$ has two poles on the right and no pole on the imaginary axis. The Nyquist criterion can therefore be formulated as $\Delta\text{arc}\{1+L(j\omega)\} = +2\pi$. At zero frequency, the open loop response starts at 100 dB and –360°. In the plot of Figure 3.6a this point is located at the far right. With increasing frequency, the magnitude decreases and the phase increases. This results in a curve in counter-clockwise direction. When it comes near to the origin, the zoomed plot of Figure 3.6b is used. It can be seen that the curve passes the point $-1+j\,0$ from below and finally moves to the origin at $0+j\,0$. If the complete curve is observed by an arrow that is drawn from the point $-1+j\,0$ to the curve, the arrow shows a full counter-clockwise rotation. This gives a continuous phase change of $+2\pi$ and the Nyquist criterion if fulfilled.

3.2.2 Power Supply Rejection

Beside the stability of the control loop, the performance with respect to power supply rejection is of importance. In the present section, analytical models of the power supply transfer functions from v_{dd} and v_{dd2} are derived.

Power Supply Gain from v_{dd}

For the evaluation of the power supply gain from v_{dd}, the small signal model of Figure 3.7 is used. This model represents the closed loop, including the feedback divider ratio β. The coupling from v_{dd} over the pass device is modeled by the output conductance g_{ds3} and the parasitic capacitor C_{gd3} of the pass device.

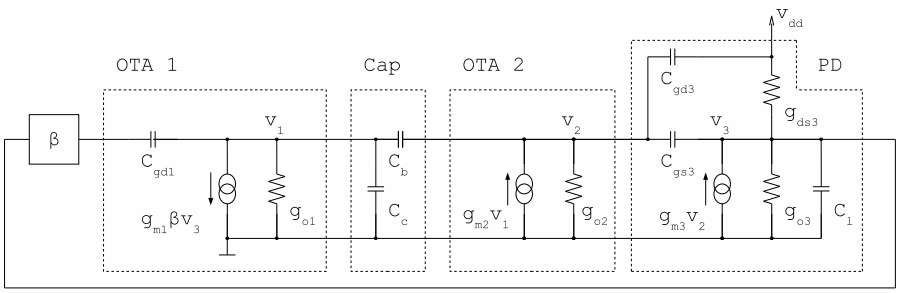


Figure 3.7: Small signal model of power supply gain from v_{dd}.

The mathematical model of the circuit from Figure 3.7 is given in Equation 3.6. As simplification, the feedback loop from v_3 to v_1 has been modeled unidirectional, i.e. there is no direct influence in the direction from v_1 to v_3.

$$\begin{bmatrix} g_{o1}+sC_{n1} & -sC_b & \beta(g_{m1}-sC_{gd1}) \\ -(g_{m2}+sC_b) & g_{o2}+sC_{n2} & -sC_{gs3} \\ 0 & -(g_{m3}+sC_{gs3}) & g_{o3}+sC_{n3} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ sC_{gd3}\, v_{dd} \\ g_{ds3}\, v_{dd} \end{bmatrix} \tag{3.6}$$

The analytical formulas for the power supply transfer function are given in Equations 3.7 to 3.9. The gain at DC is formed by the power supply gain of the pass device which is g_{ds3}/g_{m3}, divided by the open loop gain of the control loop. The first zero ω_{z1} is equal to the dominating pole of the open loop.

$$A_0 = \frac{g_{ds3}/g_{m3}}{\beta\, g_{m1}/g_{o1}\, g_{m2}/g_{o2}} \tag{3.7}$$

$$\omega_{p1} = \frac{\beta\, g_{m1} - g_{m2}}{C_c} \qquad \omega_{z1} = -\frac{g_{o1}}{C_b\, g_{m2}/g_{o2}} \tag{3.8}$$

$$\omega_{p2} = \frac{g_{o3}}{C_{n3}} \qquad \omega_{z2} = \frac{g_{m3}}{C_{gs3}} \tag{3.9}$$

Figure 3.8 shows a comparison of the analytical model with a simulation from Spectre. The worst case power supply rejection is observed at medium frequencies. For a better PSR performance at medium frequencies, the pole ω_{p1} could be moved to lower frequencies, either by decreasing g_{m1} or increasing C_c.

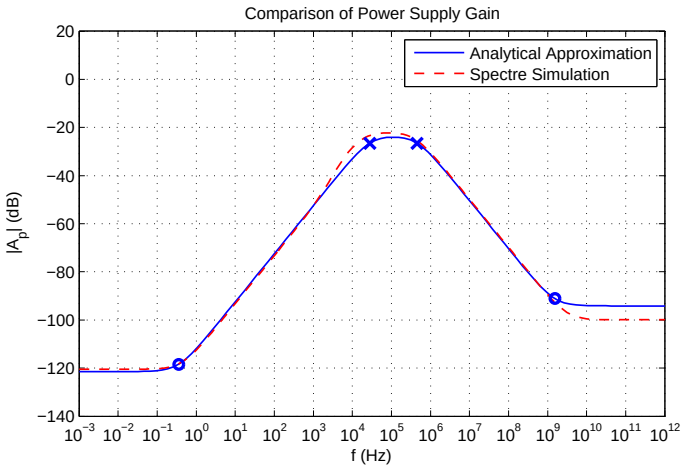


Figure 3.8: Closed loop power supply gain from v_{dd}.

Power Supply Gain from v_{dd2}

To model the power supply transfer function from v_{dd2}, a more complex model has to be used. The biasing amplifier OTA 2 which is supplied by v_{dd2} is used in a configuration, were the input signal is connected to the positive input of the OTA. This means that the input signal must cross the current mirror to get to the output of the OTA. The supply coupling from v_{dd2} also happens over the current mirror. To model these effects, OTA 2 has to be represented by a model consisting of two nodes.

As can be seen in Figure 3.9, the model of OTA 2 is similar to the OTA model that was used in Chapter 2. The only simplification that was used is that the current source at the bottom of the differential pair is considered as ideal and is therefore neglected.

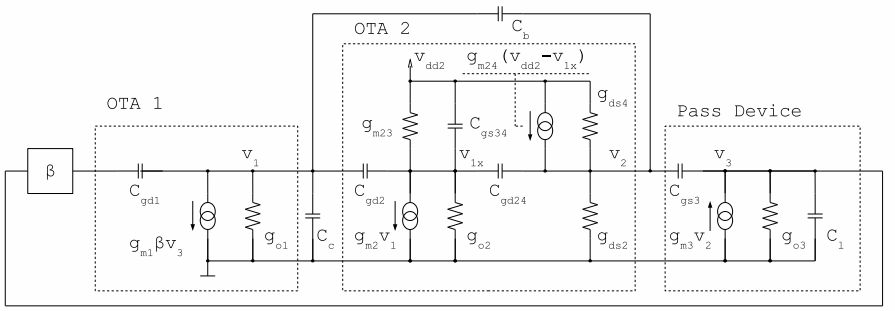


Figure 3.9: Small signal model of power supply gain from v_{dd2}.

Because of the additional node v_{1x}, the mathematical model in Equation 3.10 consists of four nodes. For the solution of the equation system this does not make a difference.

$$\begin{bmatrix} g_{\mathrm{o1}}+sC_{\mathrm{n1}} & -sC_{\mathrm{gd2}} & -sC_{\mathrm{b}} & \beta(g_{\mathrm{m1}}-sC_{\mathrm{gd1}}) \\ g_{\mathrm{m2}}-sC_{\mathrm{gd2}} & g_{\mathrm{m23}}+g_{\mathrm{o2}}+sC_{\mathrm{n1x}} & -sC_{\mathrm{gd24}} & 0 \\ -sC_{\mathrm{b}} & g_{\mathrm{m24}}-sC_{\mathrm{gd24}} & g_{\mathrm{o2}}+sC_{\mathrm{n2}} & -sC_{\mathrm{gs3}} \\ 0 & 0 & -g_{\mathrm{m3}}-sC_{\mathrm{gs3}} & g_{\mathrm{o3}}+sC_{\mathrm{n3}} \end{bmatrix} \begin{bmatrix} v_1 \\ v_{1x} \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ (g_{\mathrm{m23}}+sC_{\mathrm{gs34}})\, v_{\mathrm{dd2}} \\ g_{\mathrm{m24}}\, v_{\mathrm{dd2}} \\ 0 \end{bmatrix} \tag{3.10}$$

The model from above leads to four poles and four zeros. However two of them are pole/zero doublets and can be neglected. Equations 3.11 to 3.13 show the remaining model. The DC-gain consists of the power supply gain of OTA 2 which is approximately 1, divided by the open loop gain. The worst case PSR at medium frequencies is defined by the same poles as in the previous case of v_{dd}.

$$A_0 = \frac{1}{\beta\, g_{\mathrm{m1}}/g_{\mathrm{o1}}\; g_{\mathrm{m2}}/g_{\mathrm{o2}}} \tag{3.11}$$

$$\omega_{\mathrm{p1}} = \frac{\beta\, g_{\mathrm{m1}} - g_{\mathrm{m2}}}{C_{\mathrm{c}}} \qquad \omega_{\mathrm{z1}} = \frac{g_{\mathrm{o1}}}{C_{\mathrm{n1}}} \tag{3.12}$$

$$\omega_{\mathrm{p2}} = \frac{g_{\mathrm{o3}}}{C_{\mathrm{n3}} + C_{\mathrm{n1x}}\, g_{\mathrm{o3}}/g_{\mathrm{m23}}} \qquad \omega_{\mathrm{z2}} = \frac{g_{\mathrm{m3}}}{C_{\mathrm{gs3}}} \tag{3.13}$$

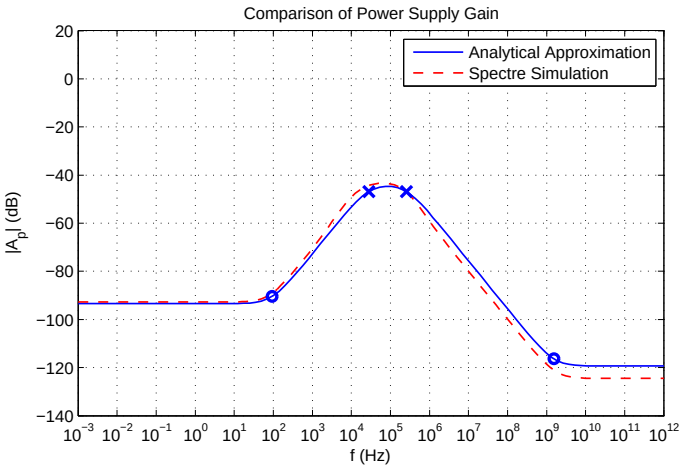


Figure 3.10: Closed loop power supply gain from v_{dd2}.

3.2.3 Output Impedance

For a derivation of the closed loop output impedance, the model of Figure 3.11 has been used. In this case, the transfer function from the load current i_{l} to the regulated voltage v_3 is derived. The respective equation system is given in Equation 3.14.

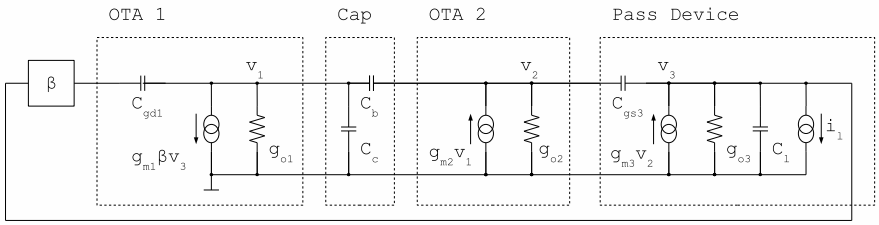


Figure 3.11: Small signal model of closed loop output impedance.

$$\begin{bmatrix} g_{o1}+sC_{n1} & -sC_b & \beta(g_{m1}-sC_{gd1}) \\ -(g_{m2}+sC_b) & g_{o2}+sC_{n2} & -sC_{gs3} \\ 0 & -(g_{m3}+sC_{gs3}) & g_{o3}+sC_{n3} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1\, i_l \end{bmatrix} \tag{3.14}$$

The DC-gain of the output impedance is equal to the output impedance of the pass device ($1/g_{m3}$), divided by the loop gain of the control loop. The worst case output impedance is observed at medium frequencies, similar to the plots of the power supply rejection. At higher frequencies, the output impedance decreases because of the output capacitor C_l. In a real circuit however, C_l would have also an equivalent series resistance (ESR) which would define a minimum output resistance at high frequencies.

$$A_0 = \frac{1}{g_{m3}} \frac{1}{\beta\, g_{m1}/g_{o1}\; g_{m2}/g_{o2}} \tag{3.15}$$

$$\omega_{p1} = \frac{\beta\, g_{m1} - g_{m2}\, g_{o2}/g_{o3}}{C_b} \qquad \omega_{z1} = -\frac{g_{o1}}{C_b\, g_{m2}/g_{o2}} \tag{3.16}$$

$$\omega_{p2} = \frac{g_{o3}}{C_{n3}} \tag{3.17}$$

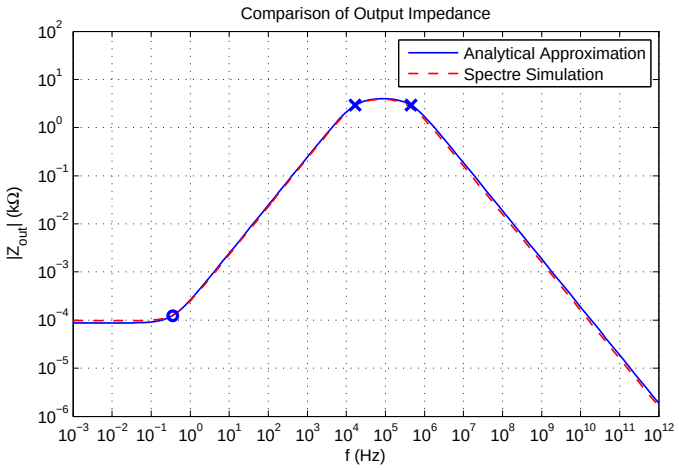


Figure 3.12: Closed loop output impedance.

3.3 System Design

The next step after the investigation of the control loop is the actual implementation of the regulator system. Beside the control loop, the chosen regulator architecture needs an oscillator and a charge pump for the generation of the additional supply voltage v_{dd2}. Also a startup circuit is necessary to ensure that the circuit reaches the correct operating point, after the main supply voltage v_{dd} is ramped up.

An important constraint for the design of the peripheral blocks and the integration of the complete system is the power supply concept. In Figure 3.13, a block diagram of the system and the needed supply voltages is shown.

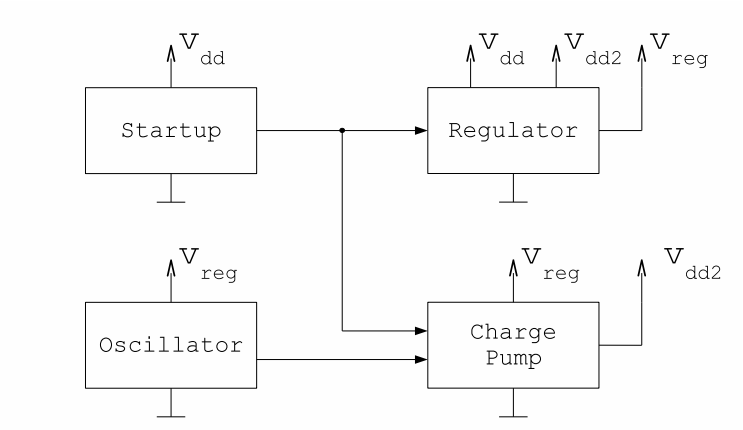


Figure 3.13: Block diagram of power supply concept.

The regulator block in the top right of Figure 3.13 contains the control loop. For correct operation it needs both v_{dd} and v_{dd2}. Additionally the regulated supply voltage v_{reg} is used internally of the regulator block to supply the error amplifier OTA 1.

For the supply of the oscillator and the charge pump, there are basically two possibilities: v_{dd} or v_{reg}. Like in the case of the error amplifier, the regulated supply voltage v_{reg} was chosen. This requires a more complicated startup procedure, but on the other hand the advantage of a fixed supply voltage predominates. With v_{dd}, an additional regulation mechanism would be necessary to ensure that v_{dd2} maintains a defined voltage level. Additionally the power dissipation of the oscillator strongly depends on the voltage level and is therefore minimized if the low-voltage v_{reg} is used as supply.

In the next sections the circuit level design of the oscillator and the charge pump is described. After that, the startup procedure and its implementation are presented.

3.3.1 Oscillator

The oscillator is an important block with respect to power dissipation. To ensure good overall efficiency, an ultra low power oscillator is needed. Simulations have shown that the charge pump requires an operation frequency of only a few kilohertz, to supply the biasing amplifier with the ultra low bias current of approximately 1 nA. On the other hand, such a low operation frequency leads to an unreasonable long startup time. To address both issues, namely ultra low current consumption during operation and a relatively fast startup, an adjustable frequency is needed. Figure 3.14 shows the architecture of a relaxation oscillator that was designed to meet the requirements.

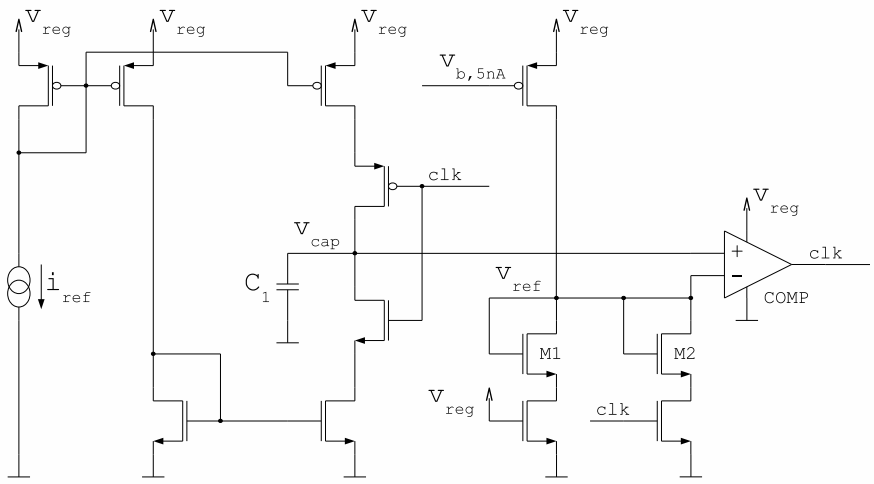


Figure 3.14: Schematic of the relaxation oscillator.

The main component of the oscillator is a continuous-time comparator which compares the voltages v_{cap} and v_{ref}. The voltage v_{cap} is stored in the capacitor C_1 and can be increased or decreased by the reference current i_{ref}. The voltage v_{ref} is derived from diode-connected transistors and can be switched between two states: If *clk* is low, it is derived from the gate-source voltage of M1. If *clk* is high, it is derived from the gate-source voltage of M1 and M2 in parallel. In the first case, the voltage v_{ref} shows a higher value of $v_{\text{ref,high}}$ and in the second case, it shows a lower value of $v_{\text{ref,low}}$.

The circuit is initialized by a reset-signal that disables the comparator, sets the signal *clk* to low and connects the nodes v_{cap} and v_{ref} such that both reach a voltage of $v_{\text{ref,high}}$. If the reset-signal is released, the connection of v_{cap} and v_{ref} is opened and v_{cap} starts to rise. As a result, the comparator changes to high, the reference voltage v_{ref} changes to $v_{\text{ref,low}}$ and the capacitor voltage v_{cap} decreases again. This decrease continues, until v_{cap} reaches $v_{\text{ref,low}}$. After that, the comparator changes to low, v_{ref} becomes $v_{\text{ref,high}}$ and the capacitor voltage v_{cap} rises again. By continuous repetition of this process, the rectangular signal *clk* is generated.

The continuous-time comparator that is needed for the oscillator is implemented by the circuit in Figure 3.15. It consists of a current-mirror OTA, cascaded by two inverters. In the present case, it is important to use an OTA with large output swing, since the first inverter should be driven by a rail-to-rail signal. With a simple OTA, the output swing would be limited by the input common mode level. This would lead to an undesired feedback of the output voltage to the input voltage, when the comparator changes its state. To prevent this feedback, the current-mirror topology is used, even if it needs more current than the simple OTA architecture. The two inverters behind the OTA are designed with non-minimal, but decreasing transistor lengths: The first inverter uses a length of 2 µm and the second inverter uses a length of 1 µm. This limits the inverter-quiescent-current that is caused by the limited slew-rate of the OTA.

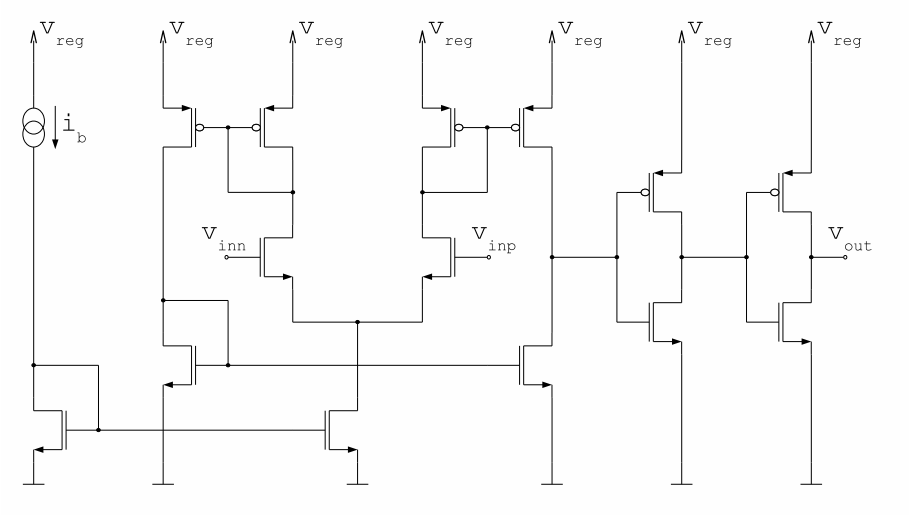


Figure 3.15: Schematic of comparator.

The resulting frequency of the oscillator can be estimated by Equation 3.18. The real oscillation frequency will be slightly smaller, since the formula neglects the delay of the comparator. As can be seen from the formula, the oscillation frequency can be modified either by the size of the capacitor, by the reference current, or by the reference voltage difference. In the present design, a variable reference current is used to realize two modes: A normal mode with 20 kHz and a fast mode with 200 kHz.

$$f = \frac{1}{2\,C_1}\,\frac{i_{\text{ref}}}{v_{\text{ref,high}} - v_{\text{ref,low}}} \tag{3.18}$$

An interesting property of the proposed oscillator is the possibility of temperature compensation. As can be seen from Equation 3.18, the oscillation frequency depends on the ratio of the reference current and the reference voltage difference. If these terms are matched, the frequency of the oscillator can be made temperature independent.

The difference of the gate-source voltage for two MOS transistors that are biased by the same current, can be expressed by Equation 3.19, if weak inversion is assumed.

$$\Delta V_{\mathrm{GS}} = n\,\Phi_{\mathrm{T}} \ln\left(\frac{(W/L)_1}{(W/L)_2}\right) = n\frac{k\,T}{e}\ln(k) \tag{3.19}$$

It depends on the subthreshold slope factor n, the thermal potential Φ_{T}, and the transistor dimensions W/L. The thermal potential Φ_{T} is composed of the Boltzmann constant $k = 1.38 \times 10^{-23}$ J/K, the elementary charge $e = 1.6 \times 10^{-19}$ As, and the absolute temperature T in Kelvin. For matching reasons it is the best choice to use equal sized transistors. A difference in gate-source voltage is then created by the connection of an integer number of these transistors in parallel, compared to an other integer number of these transistors in parallel. The ratio of these integers is defined as k in Equation 3.19.

If a device ratio of $k = 2$ and a subthreshold slope factor of $n = 1.5$ is assumed, the difference of gate-source voltage can be approximated by Φ_{T}. This term has an absolute value of 25.9 mV at a temperature of 300 K and a temperature coefficient of +86.25 µV/K.

The low temperature coefficient of a gate-source voltage difference is often used for the generation of bias currents. The circuit in Figure 3.16 shows a PTAT reference, where PTAT stands for "proportional to absolute temperature". It consists of two current mirrors formed by unit-sized devices. This means that all transistors have the same dimensions, except for M2 which is formed by k devices in parallel. At the source of M2, a resistor is connected. By this configuration, the voltage across the resistor is equal to the difference of the gate-source voltages of M1 and M2. As a consequence, also the current that flows through the resistor is proportional to this voltage difference.

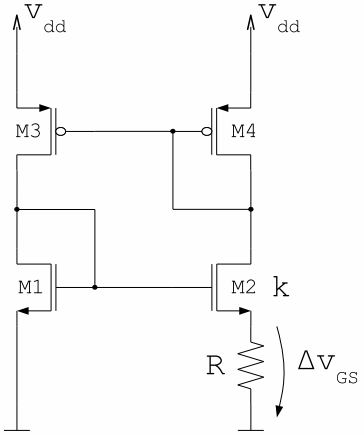


Figure 3.16: Schematic of PTAT reference generator.

After the basic considerations concerning temperature behavior from above, it will now be investigated how the reference current and the reference voltage of the oscillator must be designed to get a temperature compensated oscillation frequency. In Equation 3.20 the temperature dependency of the oscillation frequency is given.

$$f(T) = \frac{1}{2\,C_1}\,\frac{i_{\mathrm{ref}}(T)}{\Delta v_{\mathrm{ref}}(T)} \tag{3.20}$$

If the formula is derived with respect to temperature, by use of the quotient rule of differentiation, the term of Equation 3.21 is obtained.

$$\frac{\partial f}{\partial T} = \frac{1}{2\,C_1\,\Delta v_{\mathrm{ref}}}\left[\frac{\partial i_{\mathrm{ref}}}{\partial T} - \frac{i_{\mathrm{ref}}}{\Delta v_{\mathrm{ref}}}\frac{\partial \Delta v_{\mathrm{ref}}}{\partial T}\right] \tag{3.21}$$

To get a temperature independent oscillation frequency, the term in the square brackets must be equal to zero. This is achieved, if Equation 3.22 is fulfilled.

$$\frac{\partial i_{\mathrm{ref}}}{\partial T} = \frac{i_{\mathrm{ref}}}{\Delta v_{\mathrm{ref}}}\frac{\partial \Delta v_{\mathrm{ref}}}{\partial T} \tag{3.22}$$

It requires that the temperature coefficient of i_{ref} is proportional to the temperature coefficient of Δv_{ref}, with a factor of proportionality that is formed by the absolute values of the respective quantities. It can also be formulated by Equation 3.23.

$$\frac{\mathrm{TC}_{i_{\mathrm{ref}}}}{i_{\mathrm{ref}}} = \frac{\mathrm{TC}_{\Delta v_{\mathrm{ref}}}}{\Delta v_{\mathrm{ref}}} \tag{3.23}$$

The ratio of temperature coefficient over absolute value must be the same for both reference current and reference voltage difference. In the case of the reference voltage difference it can be concluded from Equation 3.19 that this ratio is equal to $1/T$. In the case of a reference current that is generated by a PTAT cell like it was shown in Figure 3.16, the ratio is also $1/T$, if the temperature coefficient of the resistor is neglected.

From a practical perspective, the criterion from Equation 3.23 can be used as basis for simulations. A good matching was achieved, when the same device ratio $k = 2$ was used for both the PTAT cell and the reference voltage generation. As bias current for the reference voltage generation, the same current as in the PTAT cell was used.

3.3.2 Charge Pump

For the charge pump, a voltage doubler architecture was chosen. The circuit is depicted in Figure 3.17 and is mainly taken from [17]. It consists of a differential pumping stage on the left and a pair of output switches on the right.

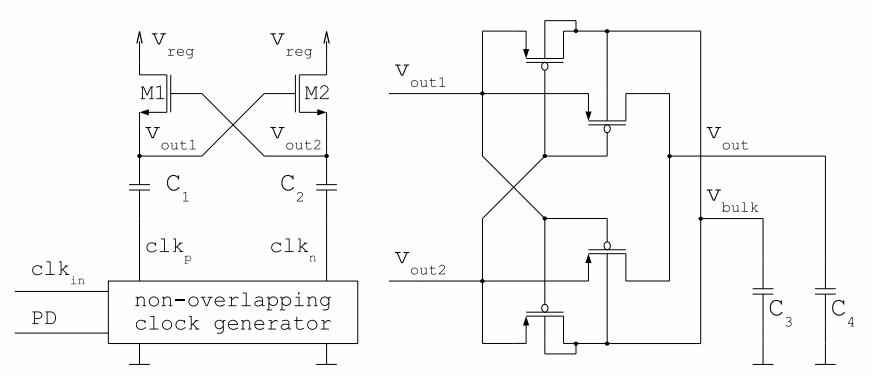


Figure 3.17: Schematic of voltage doubler charge pump.

At the bottom plate of the capacitors C_1/C_2, a differential clock with a magnitude of v_{reg} is applied. With the assumption that the capacitors are initially charged to v_{reg}, this leads to a differential signal at the top plates that changes between v_{reg} and $2\,v_{\mathrm{reg}}$. The cross coupled pair of M1/M2 ensures that the capacitors remain charged, because it connects them alternately to the supply voltage.

The output switches on the right are also composed of cross coupled transistors. They alternately connect the higher voltage node of v_{out1} and v_{out2} to the output node v_{out} and the bulk node v_{bulk}. If the node v_{out1} is at the higher potential of $2\,v_{\mathrm{reg}}$, the other node v_{out2} is at the lower potential of v_{reg}. Because of this, the upper transistors conduct and $2\,v_{\mathrm{reg}}$ is connected to the output. In the next clock phase, the potentials are interchanged and the lower transistors conduct. The reason why a separate capacitor for the bulk node is used, is to ensure that this node always stays at the highest potential, even if the voltage at v_{out} is decreased by loading.

The differential clock that is needed for the operation of the charge pump is generated by the circuit in Figure 3.18. It takes the clock generated by the oscillator as input and generates a differential clock with non-overlapping high states at the output. To enable a pre-charging of the charge pump capacitors, also a power-down mode is implemented that pulls both $\mathrm{clk_p}$ and $\mathrm{clk_n}$ to ground. All logic gates are supplied by v_{reg}.

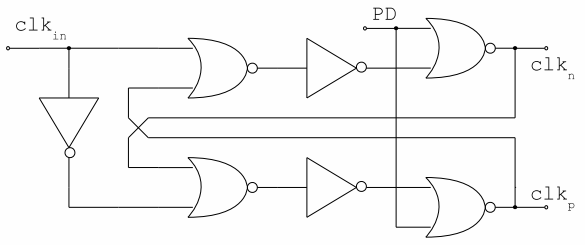


Figure 3.18: Schematic of non-overlapping clock generator.

3.3.3 Startup

With the proposed regulator topology, the regulated voltage v_{reg} must be initially ramped up, such that the error amplifier, the oscillator and the charge pump can start working. This ramp-up requires an initial charging of the capacitor C_{b} in Figure 3.19. It is accomplished by an external circuit that drives a current in node v_2 and pulls down node v_1 by transistor M1. The initial charging of C_{b} will be called the precharge phase.

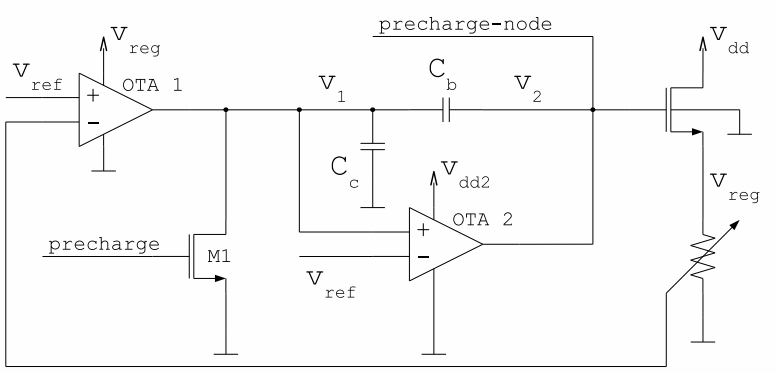


Figure 3.19: Control loop with startup circuit.

During the precharge-phase, OTA 1 is in a power-down-mode, and the voltage across C_{b} is increased, until it reaches a level of approximately 1 V. After that, the signal *precharge* changes to low, the error amplifier OTA 1 is enabled and the external circuit at node v_2 changes to a high impedance state. Since the regulated voltage v_{reg} has not reached its final value yet, OTA 1 drives a current into node v_1. This increases the voltage at nodes v_1 and v_2. Because of the initial precharge of C_{b}, the voltage at v_2 can now reach a level that is higher than the supply voltage, without the use of the charge pump.

For a fixed time interval after the precharge-phase, the oscillator is operated at the high clock frequency of 200 kHz to ramp up the charge pump voltage v_{dd2}. This activates the biasing amplifier OTA 2 which regulates the voltage at v_1 to a level of v_{ref}.

Precharge

The circuit in Figure 3.20 is used for the initial precharge of node v_2 and the generation of the *precharge* signal. Additionally it supports two additional precharge-nodes that are used to precharge the capacitors C_3 and C_4 of the charge pump. The circuit can be divided in the three functional blocks: current source, flip flop and switches.

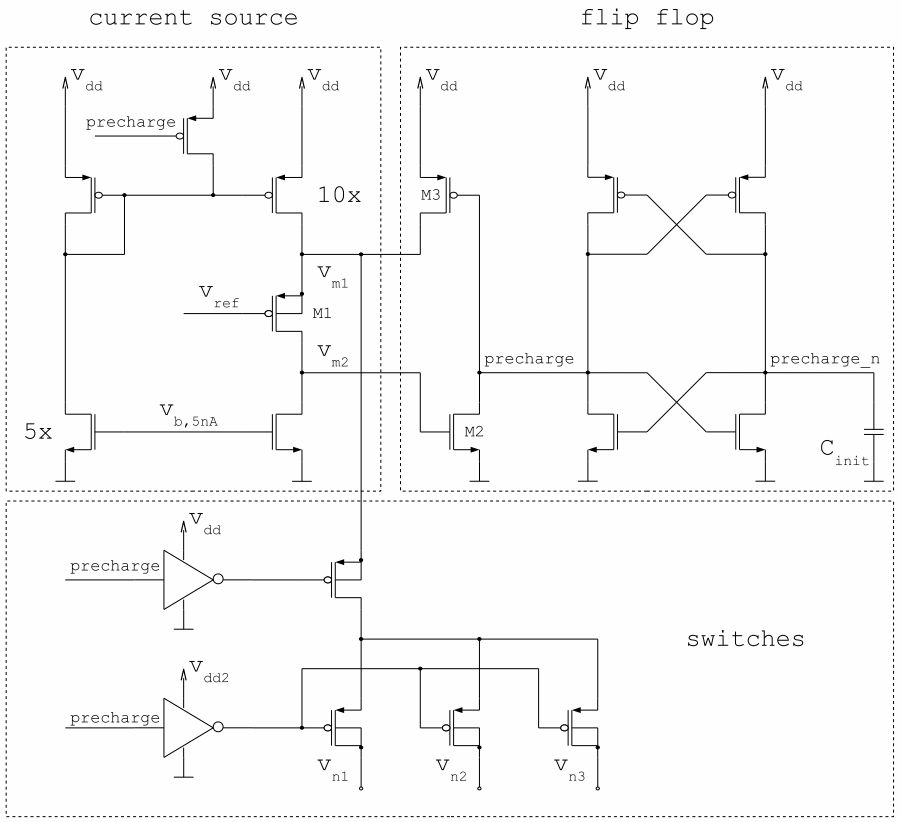


Figure 3.20: Schematic of precharge circuit.

The NMOS transistors of the current source are biased with a voltage that generates a current of 5 nA. At the right NMOS transistor this current is used to keep v_{m2} initially at low. The left NMOS transistor is composed of five unit-sized devices and multiplies the current by a factor of 5. Over the PMOS current mirror it is then multiplied by a factor of 10, which leads to a current of 250 nA. Through the switches at the lower side of the precharge circuit, this current is forwarded to the precharge nodes v_{n1} to v_{n3}, which are connected to the node v_2 of the control loop and the capacitors of the charge pump.

The flip flop of the precharge circuit is initialized with the signal *precharge* at high. This is caused by the capacitor C_{init} which is initially uncharged. The transistor M1 is initially in a non-conducting state, since both v_{m1} and v_{m2} are at a low potential. During the precharge phase the voltage at v_{m1} rises. When it reaches a level that is sufficiently above v_{ref}, the transistor M1 starts to conduct. As a result, v_{m2} is pulled up and the transistor M2 starts to conduct. This pulls the *precharge* node to low and the flip flop changes its state. The transition of the *precharge* signal to low disables the current source and the switches. The node v_{m1} is now held at a high potential by transistor M3.

The switches are implemented as a serial combination of two PMOS transistors. The reason why two transistors are used, is to ensure that a high impedance state can be reached when the *precharge* signal is low. At the side of v_{m1} a voltage of v_{dd} is present. At the side of v_{n1} to v_{n3}, the maximum voltage is v_{dd2}. Since the main supply can have a voltage between 1 V and 3 V, and the charge pump reaches a voltage of approximately 1.8 V, it is not known in advance whether the higher voltage is present at v_{m1} or v_{n1} to v_{n3}. To address both cases, the node v_{m1} is isolated by a switch that is supplied by v_{dd} and the nodes v_{n1} to v_{n3} are isolated by switches that are supplied by v_{dd2}.

Startup-Delay

A fixed time after the *precharge* signal has changed to low, the frequency of the oscillator is decreased to the normal level of 20 kHz. Additionally a control signal $v_{\text{reg,OK}}$ is generated that signalizes that the regulated voltage v_{reg} has reached its final value. To implement the needed time delay, the circuit of Figure 3.21 is used.

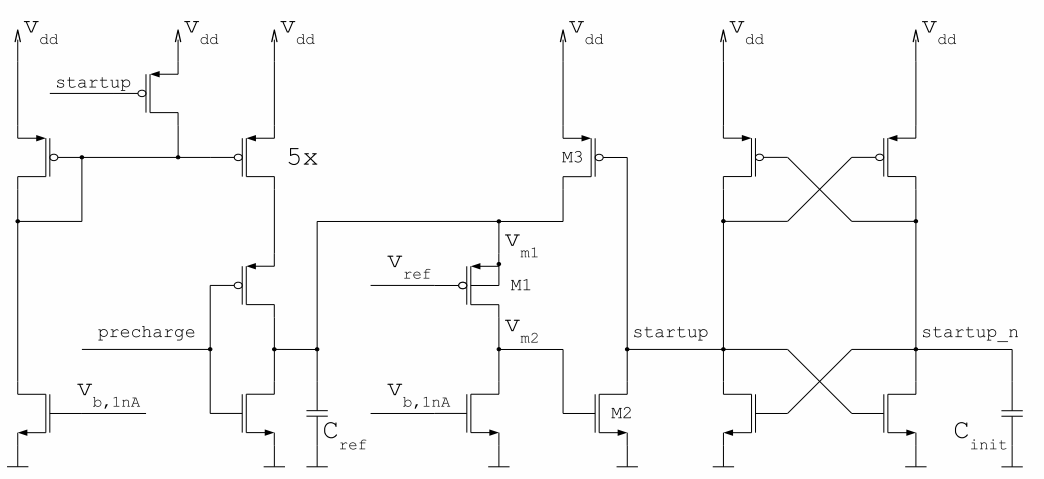


Figure 3.21: Schematic of circuit for startup-delay.

The circuit of Figure 3.21 is based on similar principles as the precharge circuit. It uses a current source that is connected to the top node of an inverter. As long as the *precharge* signal is high, the PMOS transistor of the inverter is in a high impedance state and the NMOS transistor keeps the voltage at C_{ref} at ground. If the *precharge* signal changes to low, a current flow from the current source to C_{ref} is enabled. This leads to a linear increase of the voltage at node v_{m1}. When it reaches a level that is sufficiently above v_{ref}, the transistor M1 starts to conduct and pulls v_{m2} up. This switches on M2, the signal *startup* is pulled to low and the flip flop changes its state. By the transistor M3 the node v_{m1} is permanently pulled to v_{dd} and the current source is disabled.

The time delay that is realized by this circuit can be estimated by Equation 3.24. For C_{ref}, a value of 500 fF is used, i_{ref} is equal to 5 nA, and v_{ref} is approximately 600 mV. The term $v_{\text{GS,1nA}}$ specifies the gate-source voltage were the PMOS transistor M1 shows a drain-current of 1 nA. This drain current is reached at a gate-source voltage of approximately 300 mV. The resulting time delay is approximately 90 µs.

$$\Delta t = \frac{C_{\text{ref}}}{i_{\text{ref}}}(v_{\text{ref}} + v_{\text{GS,1nA}}) \tag{3.24}$$

Power-Down

For the implementation of a power-down mode additional transistors were added to the system. In the power-down mode, the flip flops of the precharge and the startup-delay circuits are reset and the capacitors at the precharge nodes are discharged. This is accomplished by power-down transistors that are added to specific nodes and pull the nodes to ground, if a power down signal PD goes to high.

At the precharge and the startup-delay circuits, transistors were added to the nodes v_{m2} and precharge_n/startup_n. At the control loop a transistor was added to node v_2 and in the charge pump, transistors were added to the nodes $v_{\text{out}}/v_{\text{bulk}}$.

Power-down transistors were also added to the current mirrors that distribute the bias currents. This way zero current consumption is assured during power-down mode.

4 Simulation Results

In the following chapter, simulation results are presented that were obtained with the circuit simulator Spectre, based on an Infineon 130 nm CMOS process. In Section 4.1, nominal transistor models, a temperature of 27°C and a supply voltage of 1.2 V were used. In Section 4.2 the influence of parameter variations is investigated.

4.1 Transient Simulations

Figure 4.1 shows the startup behavior of the circuit, after a fast ramp-up of v_{dd}. In the precharge phase the voltages v_2 and v_{dd2} are ramped up to approximately 1 V. After that, the oscillator is enabled and operates at the high frequency. Simultaneously also the error amplifier is enabled which increases the voltage at v_1 and v_2. After the startup signal has changed to low, the final operating point is reached and the oscillator operates at the low frequency. The complete startup procedure takes approximately 100 µs.

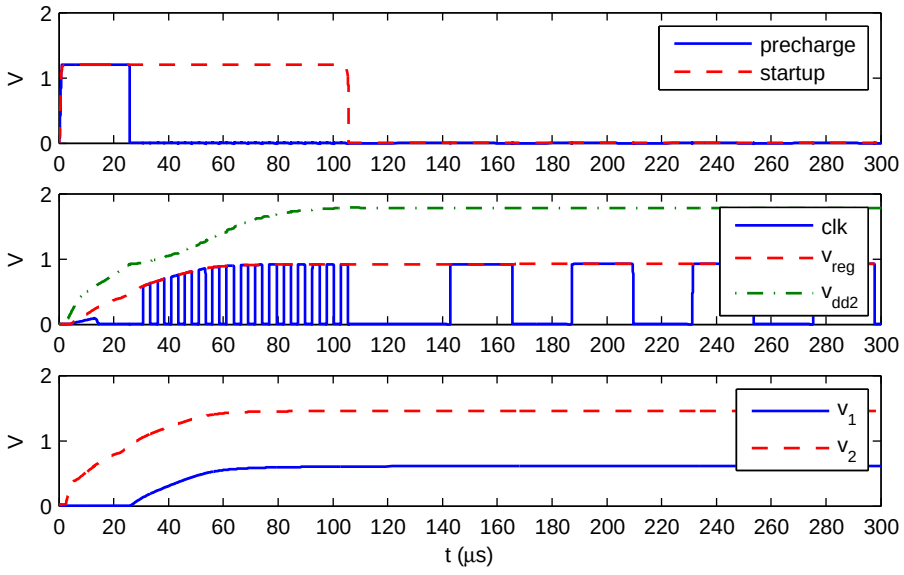


Figure 4.1: Transient simulation of startup behavior.

Oscillator

In Figure 4.2, the startup behavior is considered with respect to internal signals of the oscillator. In the upper plot the input signals of the comparator v_{cap} and v_{ref} are depicted. It can be seen that the reference voltage v_{ref} changes between two states and the voltage at the capacitor v_{cap} linearly increases and decreases to follow the behavior of v_{ref}. In the lower plot, the signals clk and v_{comp} are depicted. The signal v_{comp} represents the output signal of the OTA inside the comparator. The signal clk is the output signal of the comparator that is generated after the two inverter stages.

Charge Pump

In Figure 4.3, internal signals of the charge pump are shown. In this plot, the time axis is zoomed in compared to the previous plots to focus on the phase were the oscillator operates at the higher frequency. It can be seen that during the precharge phase, the differential clock signals $\mathrm{clk}_1/\mathrm{clk}_2$ are low and the nodes $\mathrm{v}_{\mathrm{out}}/\mathrm{v}_{\mathrm{bulk}}$ are charged by the precharge circuit. As a consequence also the voltage at $\mathrm{v}_{\mathrm{out1}}/\mathrm{v}_{\mathrm{out2}}$ rises. When the differential clock is enabled, the signals $\mathrm{v}_{\mathrm{out1}}/\mathrm{v}_{\mathrm{out2}}$ are alternately pumped up to $2\,v_{\mathrm{reg}}$ and the higher voltage node is connected to the output nodes $\mathrm{v}_{\mathrm{out}}/\mathrm{v}_{\mathrm{bulk}}$.

Power-Down

Beside the startup that is caused by a ramp-up of the main supply, also the startup after a power-down must be guaranteed. In Figure 4.4 the power-down signal PD is initially high when the main supply is ramped up. At 100 µs the signal PD goes to low and a startup procedure is initiated. At 400 µs, when the system has reached the final operating point, the system is sent to power-down mode again. To verify a correct startup after a power-down mode, the system is woken up again at 600 µs.

Main Supply Drop

A further scenario that was simulated is the startup after a drop of the main supply. In Figure 4.5 the system starts up regularly, but at 250 µs there is a fast drop of v_{dd} to 0 V. After 500 µs, the main supply is ramped up again, and the system is able to start again. It can be seen from this simulation, that the drop of v_{dd} leads to a drop of v_{reg}. This causes a drop of v_1 because OTA 1 is supplied by v_{reg}. However the voltage stored in C_{b} and in the charge pump capacitors nearly stays the same.

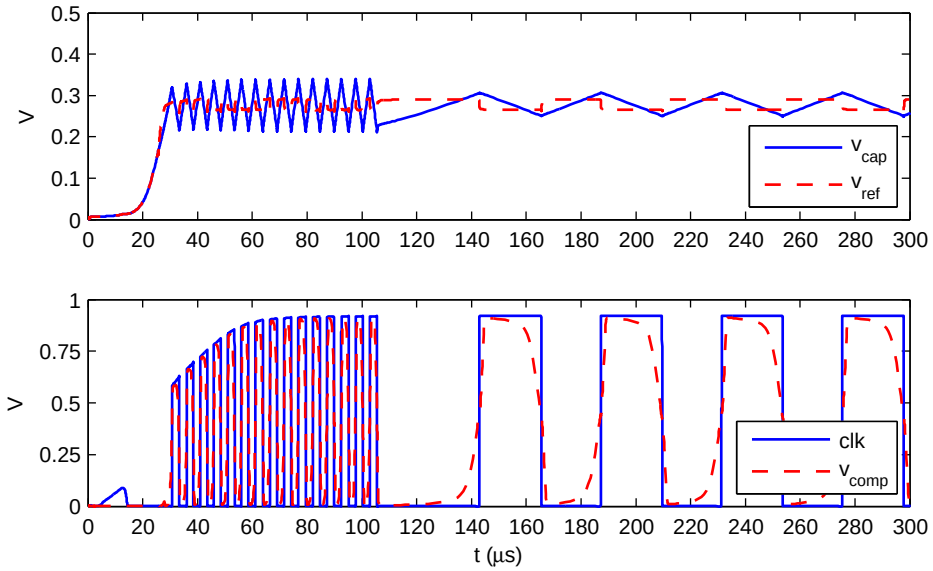


Figure 4.2: Oscillator signals during startup.

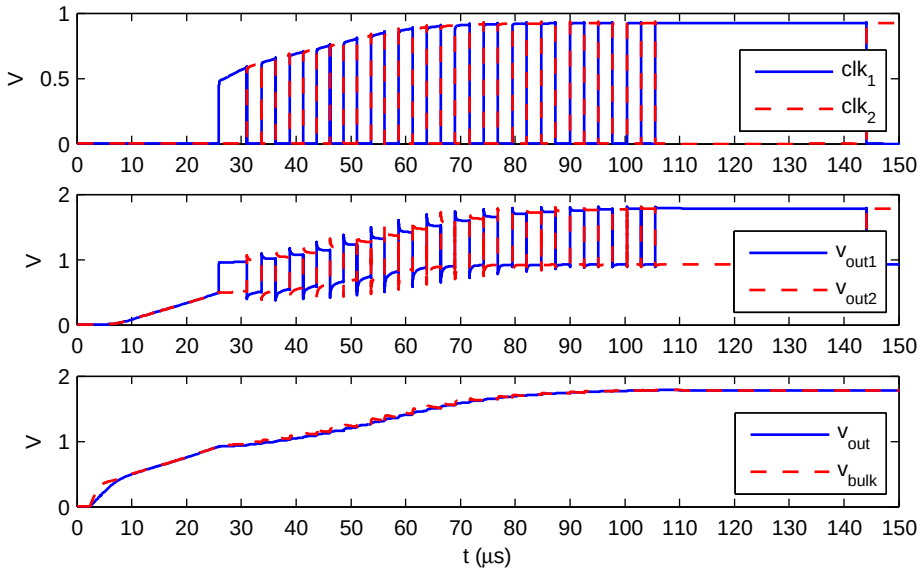


Figure 4.3: Charge pump signals during startup.

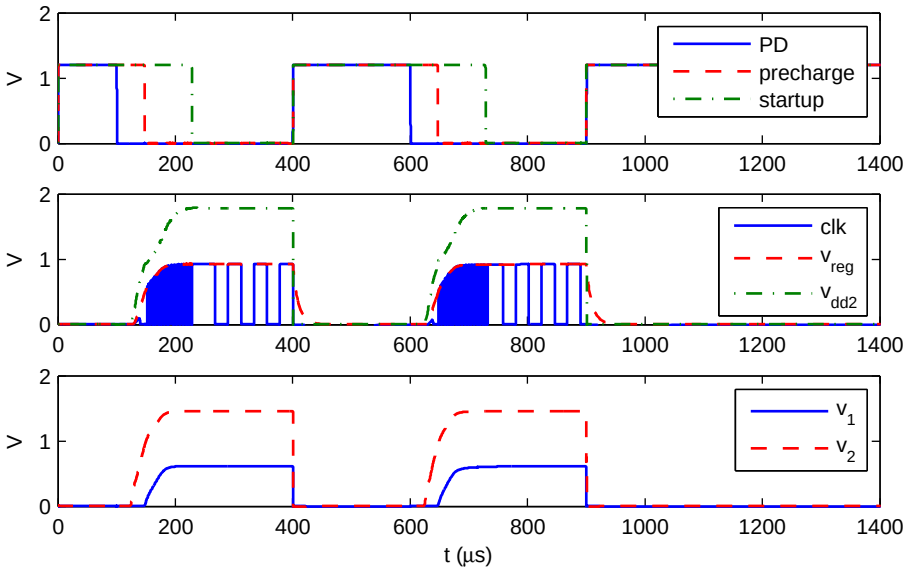


Figure 4.4: Transient simulation of power down mode.

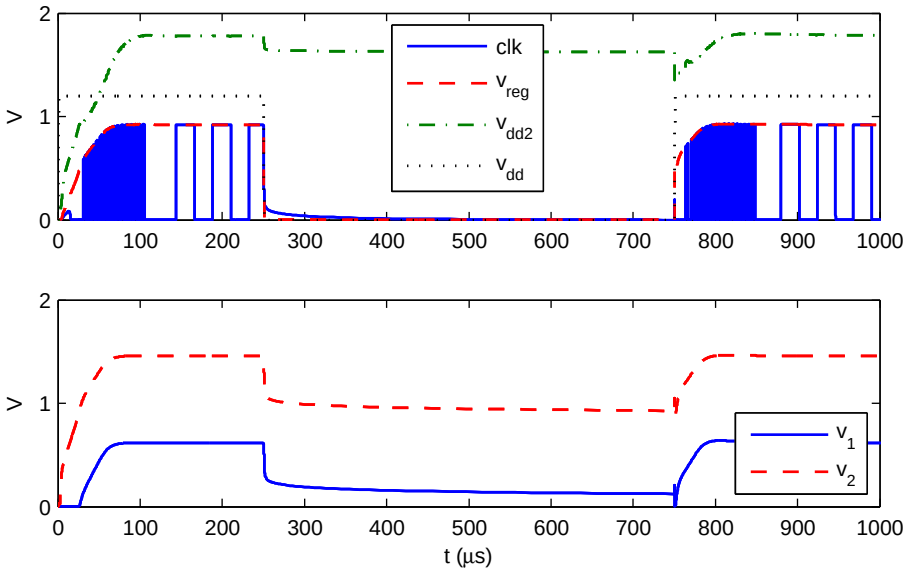


Figure 4.5: Transient simulation of main supply drop.

Slow Main Supply Ramp-Up

In Figure 4.6, a slow ramp-up of the main supply is shown. It can be seen that also at a slow ramp-up, the flip flops of the startup circuits initialize in the correct state. The signals precharge and startup rise therefore linearly with v_{dd}. When v_{dd} reaches approximately 500 mV, the current source of the precharge circuit gets active and the nodes v_2 and v_{dd2} rise also with v_{dd}. When v_{dd} reaches approximately 1 V, the precharge signal changes to low and the startup procedure is initialized.

Slow Main Supply Ramp-Down

In Figure 4.7, the simulation from before is continued with a slow ramp-down of the main supply. It can be seen that the regulated voltage is kept constant, up to the point were v_{dd} reaches v_{reg}. After that, the regulator tries to keep v_{reg} constant by an increase of v_1 and v_2. As a consequence the pass device changes into the linear region and behaves like a closed switch. This leads to a linear ramp-down of v_{reg} and v_1/v_2. The biasing capacitor C_{b} and the charge pump capacitors remain charged.

Current Consumption

In Figure 4.8, the quiescent current during the startup process is shown. When the power-down signal PD is high, the regulator does not drain any current. When the signal PD goes to low, the precharge circuit gets active and the static current consumption rises to approximately 350 nA. The delay between the change of PD and the rise of quiescent current is caused by buffer capacitors at the current mirrors. These capacitors need to be charged by the low bias current of 5 nA, before a higher precharge current can be generated. After the precharge phase, the system is operated with the high clock frequency and dissipates approximately 200 nA. At normal operation, the current consumption is reduced to approximately 100 nA.

Line- and Load-Regulation

As a measure of line- and load-regulation, the response of the regulated voltage to supply voltage steps and load current steps was simulated. In Figure 4.9, a supply voltage step of 200 mV and a load current step of 5 µA is applied. It can be seen that the regulated output voltage shows a maximum deviation of approximately 10 mV. The feed-through of the clock frequency from the charge pump causes peaks of approximately 1 mV.

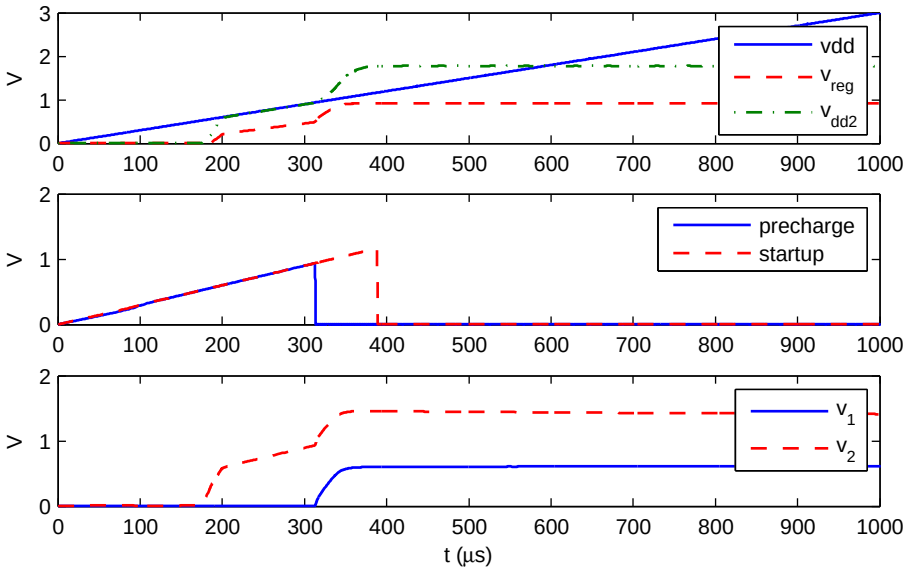


Figure 4.6: Transient simulation of slow main supply ramp-up.

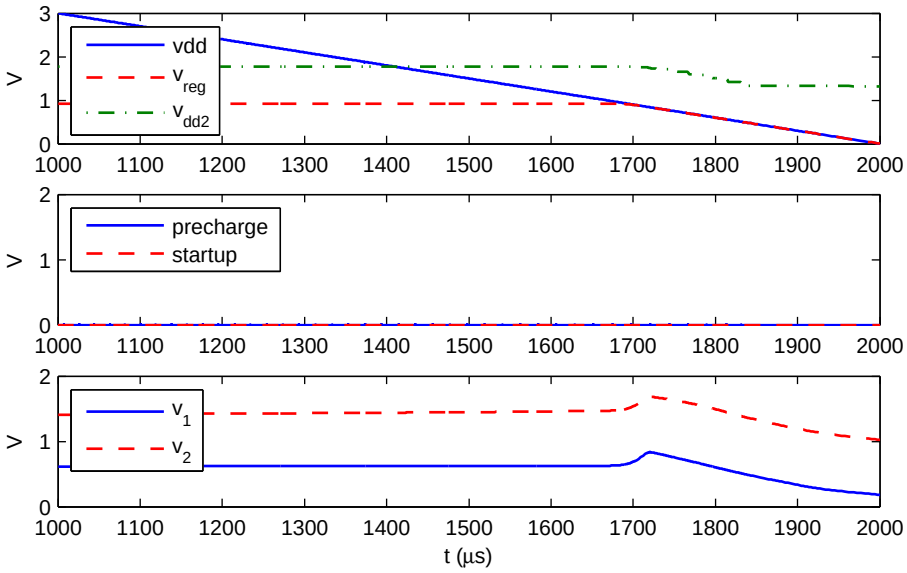


Figure 4.7: Transient simulation of slow main supply ramp-down.

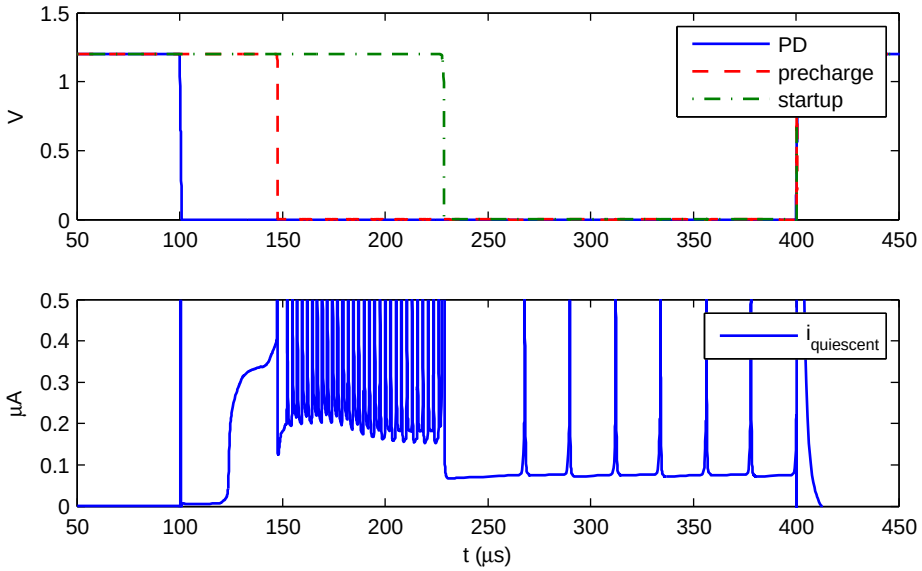


Figure 4.8: Transient simulation of current consumption.

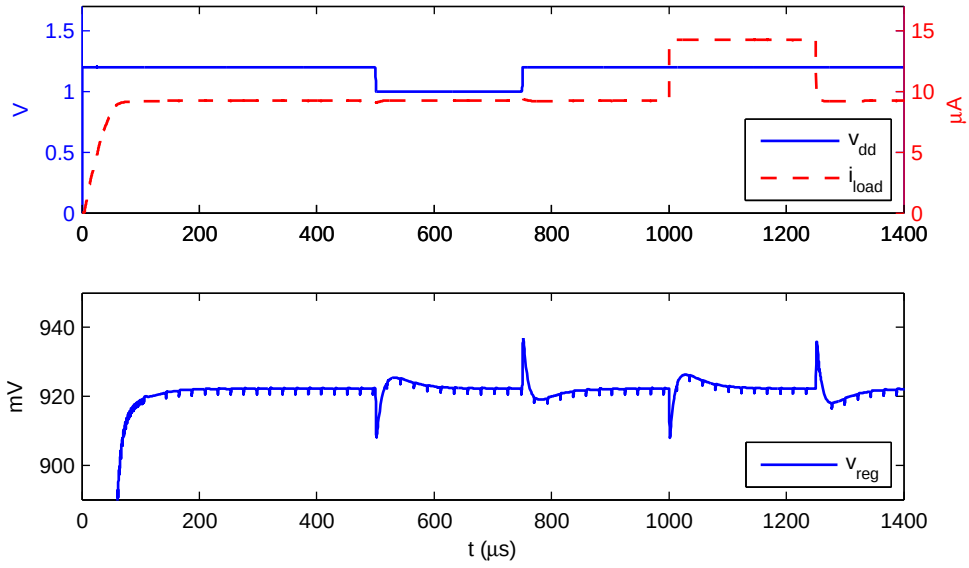


Figure 4.9: Transient simulation of line- and load-regulation.

4.2 Parameter Variations

The operation of the circuit must be verified for varying process parameters, temperatures, supply conditions and load conditions. For the simulation of such parameter variations there are corner simulations, Monte Carlo simulations and parametric sweeps.

Corner Simulations

At corner simulations the extremal parameters of device models are combined in a number of ways such that the worst case operation conditions are checked. If the circuit operates under these conditions, it can be assumed that it will also operate under more moderate conditions. The advantage of corner simulations is that only a few simulation runs are needed. On the other hand, the most extreme operation conditions will most likely never occur in practice. Furthermore the influence of mismatch is not seen in corner simulations, since all devices parameters in one group are modified in the same way.

A central parameter of transistor models is the threshold voltage. This parameter is influenced by both process and temperature variations. The lowest threshold voltage is observed at the fast process corner and the highest temperature. The highest threshold voltage is observed at the slow process corner and the lowest temperature. In Figure 4.10 the startup behavior is shown for both cases.

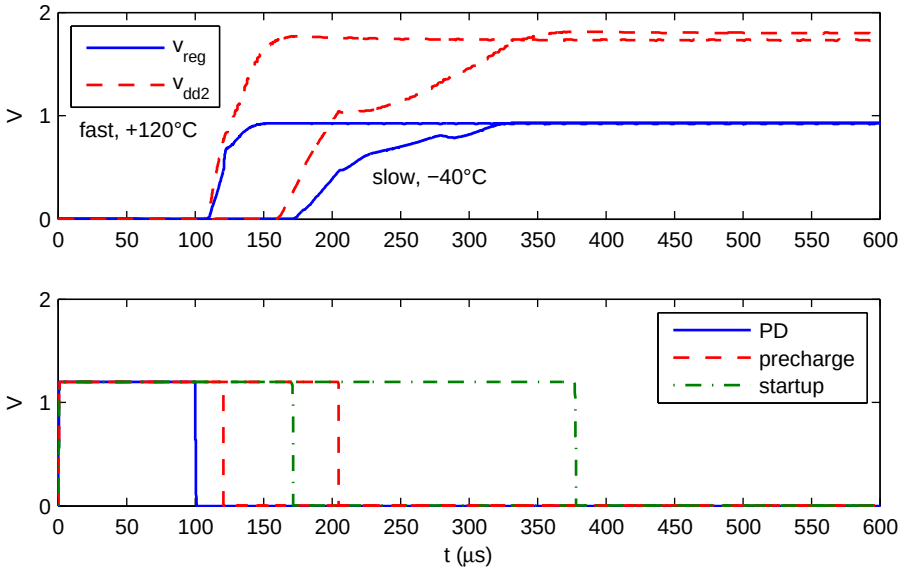


Figure 4.10: Corner simulation of startup.

Depending on the process corner and temperature, the startup procedure is initialized at different levels of v_{dd}. To determine the minimal supply voltage that is needed for the startup, a slow ramp-up of v_{dd} was simulated. In Figure 4.11 the results for the most extreme corners are shown. At the fast process corner and the high temperature, the signal precharge changes to low, if v_{dd} reaches a level of approximately 750 mV. At the slow process corner and the low temperature, a level slightly above 1 V is needed.

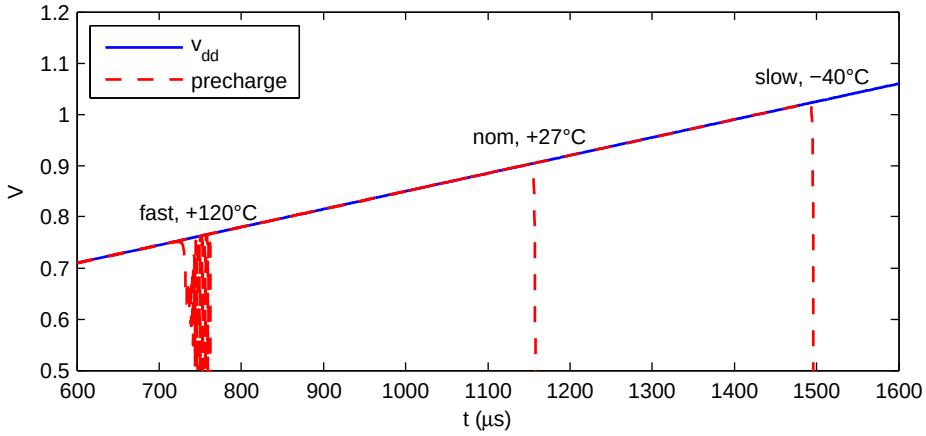


Figure 4.11: Corner simulation of minimal supply voltage.

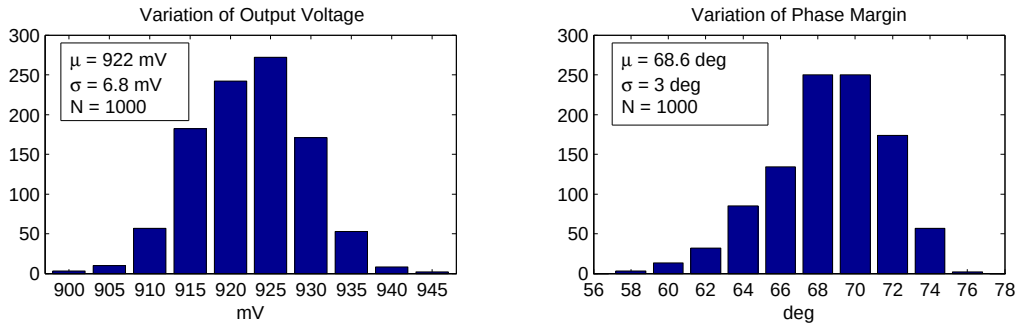
Monte Carlo Simulations

At Monte Carlo simulations a statistical distribution of the parameters is assumed and for each simulation run, the individual device parameters are sampled from this distribution. Monte Carlo simulations are more realistic and can also include mismatch. On the other hand, a large number of simulation runs are needed to get reliable results.

In Figure 4.12 the results of Monte Carlo simulations are depicted in the form of histogram plots. Both represent the distribution of a parameter, based on 1000 simulation runs.

In Figure 4.12a the DC level of the regulated output voltage was simulated. For the simulation a constant reference voltage of 615 mV and a feedback divider ratio of 2/3, composed of diode connected PMOS transistors were used. This leads to the mean value of 922 mV. The variation of the output voltage is mainly caused by the offset of the error amplifier and the mismatch of the feedback divider.

In Figure 4.12b the distribution of the phase margin is depicted. The variability is mainly caused by parameter variations between the error amplifier and the biasing amplifier.



(a) Output Voltage. (b) Phase Margin.

Figure 4.12: Results of Monte Carlo simulations.

Parametric Sweeps

At parametric sweeps a number of simulations is executed with varying parameter values. The simulation results are then evaluated to show dependencies between the swept parameters and circuit characteristics.

For the plot in Figure 4.13, the open loop frequency response was simulated at different load conditions. The resulting phase margin is depicted as a function of load current. It can be seen that for good phase margin, the load current should not be below 1 µA.

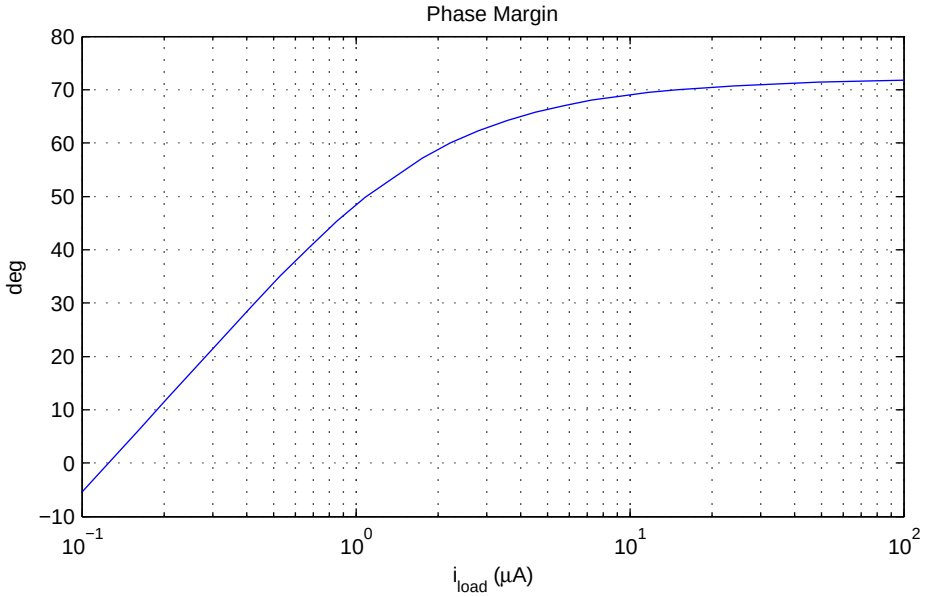


Figure 4.13: Phase margin as a function of load current.

5 Conclusion

The present work described the development of an integrated ultra low power voltage regulator. The regulator was designed with an NMOS pass device and a new control loop topology that allows the biasing for low dropout operation. Integrated into the regulator system are an ultra low power oscillator and a charge pump.

Based on detailed simulations, the functionality of the circuit was verified and the performance with respect to efficiency and regulation behavior was investigated. An implementation on silicon is planned, but is not yet available at the date of this publication.

The current consumption of the complete system at nominal operation conditions is approximately 100 nA. With a load current of 1 µA to 100 µA and the ability for low dropout operation this allows highly efficient operation.

The regulator features a power supply rejection of 120 dB at DC and and a worst case power supply rejection of 20 dB at 100 kHz. The transient response to a supply drop of 200 mV and a load current step of 5 µA shows a maximum deviation of 10 mV.

Additionally to the development of the circuit, theoretical foundations have been presented and a methodology for symbolic circuit analysis was applied. These analytical results might also be useful in the context of other circuit design tasks.

Bibliography

[1] B. Lee, "Technical review of low dropout voltage regulator operation and performance," *Application Note of Texas Instruments*, 1999.

[2] P. Gray, P. Hurst, S. Lewis, and R. Meyer, *Analysis and Design of Analog Integrated Circuits*. Wiley, 2001.

[3] M. Horn and N. Dourdoumas, *Regelungstechnik, Rechnerunterstützter Entwurf zeitkontinuierlicher und zeitdiskreter Regelkreise*. Pearson, 2004.

[4] M. Tian, V. Visvanathan, J. Hantgan, and K. Kundert, "Striving for small-signal stability," *Circuits and Devices Magazine, IEEE*, vol. 17, no. 1, pp. 31–41, Jan. 2001.

[5] M. Steyaert and W. Sansen, "Power supply rejection ratio in operational transconductance amplifiers," *Circuits and Systems, IEEE Transactions on*, vol. 37, no. 9, pp. 1077–1084, Sept. 1990.

[6] J. Teel, "Understanding power supply ripple rejection in linear regulators," *Texas Instruments Analog Applications Journal*, 2005/2Q.

[7] V. Gupta, G. Rincon-Mora, and P. Raha, "Analysis and design of monolithic, high PSR, linear regulators for SoC applications," in *SOC Conference, 2004. Proceedings. IEEE International*, Sept. 2004, pp. 311–315.

[8] P. Wambacq, F. Fernandez, G. Gielen, W. Sansen, and A. Rodriguez-Vazquez, "Efficient symbolic computation of approximated small-signal characteristics of analog integrated circuits," *Solid-State Circuits, IEEE Journal of*, vol. 30, no. 3, pp. 327–330, Mar. 1995.

[9] L. Schmidt, G. Schaller, and S. Martius, *Grundlagen der Elektrotechnik, Band 3, Netzwerke*. Pearson, 2006.

[10] S. Gruber, H. Reinisch, H. Unterassinger, M. Wiessflecker, G. Hofer, W. Pribyl, and G. Holweg, "An ultra-low power voltage regulator for wireless sensor nodes," in *Microelectronics (ICM), 2010 International Conference on*, Dec. 2010, pp. 216–219.

[11] G. Nebel, T. Baglin, I. San Sebastian, H. Sedlak, and U. Weder, "A very low drop voltage regulator using an NMOS output transistor," in *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*, vol. 4, May 2005, pp. 3857–3860.

[12] W. Kruiskamp and R. Beumer, "Low drop-out voltage regulator with full on-chip capacitance for slot-based operation," in *Solid-State Circuits Conference, 2008. ESSCIRC 2008. 34th European*, Sept. 2008, pp. 346–349.

[13] G. Giustolisi, G. Palumbo, C. Falconi, and A. D'Amico, "NMOS low drop-out regulator with dynamic biasing," in *Electronics, Circuits and Systems, 2006. ICECS '06. 13th IEEE International Conference on*, Dec. 2006, pp. 204–207.

[14] D. Camacho, P. Gui, and P. Moreira, "An NMOS low dropout voltage regulator with switched floating capacitor gate overdrive," in *Circuits and Systems, 2009. MWSCAS '09. 52nd IEEE International Midwest Symposium on*, Aug. 2009, pp. 808–811.

[15] M. Wiessflecker, G. Hofer, G. Holweg, H. Reinisch, and W. Pribyl, "A sub 1V self clocked switched capacitor bandgap reference with a current consumption of 180nA," in *Circuits and Systems, 2012. ISCAS '12. IEEE International Conference on*, May 2012, pp. 2841–2844.

[16] J. Heidrich, D. Brenk, J. Essel, M. Heinrich, M. Jung, G. Hofer, G. Holweg, R. Weigel, and G. Fischer, "Design of a low-power voltage regulator for RFID applications," in *Computational Technologies in Electrical and Electronics Engineering (SIBIRCON), 2010 IEEE Region 8 International Conference on*, July 2010, pp. 552–557.

[17] P. Favrat, P. Deval, and M. J. Declercq, "A high-efficiency CMOS voltage doubler," *Solid-State Circuits, IEEE Journal of*, vol. 33, no. 3, pp. 410–416, Mar. 1998.