Low Power CMOS Interface Circuitry for Sensors and Actuators

Nielsen, Jannik Hammel; Lehmann, Torsten; Bruun, Erik

Publication date:
2004

Document Version
Publisher's PDF, also known as Version of record

LOW POWER CMOS INTERFACE CIRCUITRY FOR SENSORS AND ACTUATORS

Jannik Hammel Nielsen

This thesis is submitted in partial fulfillment
of the requirements for obtaining the Ph.D. degree at:

Ørsted•DTU
Technical University of Denmark (DTU)

9 December, 2003

ABSTRACT

This thesis is concerned with the design of integrated CMOS circuits applicable for transducer interfacing. The project entry-point is the development of circuitry suitable for biomedical implantable sensors/actuators for neural prosthesis.

Implantable sensor outputs are characterized by very weak signals at a relatively low bandwidth. Typically for the nerve-cuff electrode, these signals have a magnitude of only a few μV over a 4 kHz bandwidth, making pre-amplification an absolute necessity prior to any further processing. To this end, we have developed two prototype nerve signal instrumentation amplifiers characterized by high, well-controlled gain and low noise performance. The first version achieves low noise performance, comparable to bipolar designs, by careful dimensioning of the input stage. The second version addresses some drawbacks of the first design and employs more advanced design solutions for improved noise performance (chopper modulation of the amplifier input stage).

Prior to further processing of the nerve signals, quantization is done. To accommodate quantization, a data converter of the sigma-delta type has been analyzed, optimized and implemented according to the specifications put forth by the recorded nerve signals. The main concern in the optimization task was the required low power consumption of the data converter.

For nerve stimulation, current pulses are used. For this purpose, current-steering digital to analog converters are useful due to their ability to directly convert a digital input word to an output current for stimulation. The D/A circuit implemented in this work, was however not aimed at implanted actuator applications, but rather high-speed, high-accuracy demonstration purposes. Although not intended for implantable circuits, the current-steering D/A principle is easily scaled for other requirements and the developed design techniques remain valid.

RESUMÉ

Afhandlingen beskriver konstruktion af integrerede kredsløb i CMOS teknologi til anvendelse som transducer interfaces. Der er primært fokuseret på interfaces til transducere, som implanteres i det menneskelige legeme med henblik på måling af nervesignaler eller elektrisk stimulering af nerver. De væsentligste nye resultater i relation hertil er opnået inden for design af støjsvage forstærkere til måling af nervesignaler.

Signaler fra nerver er meget svage, typisk nogle få μV, og har en båndbredde på omkring 4kHz. Der er derfor behov for en forstærkning af signalerne, inden en analog til digital konvertering og videre signalbehandling kan finde sted. I dette projekt er der udviklet to nye CMOS forstærkertyper til forstærkning af nervesignaler. I den første er der ved hjælp af en omhyggelig dimensionering af indgangstrinnet opnået specifikationer mht. støjegenskaber, som hidtil kun har været rapporteret i forstærkere baseret på bipolær teknologi. I den anden forstærker er støjegenskaberne yderligere forbedret ved anvendelse af andre kredsløbsteknikker (chopperkobling anvendt i forstærkerens indgangstrin).

Efter forstærkeren er der brug for en datakonvertering, og i dette projekt er en sigma-delta A/D konvertertype analyseret, optimeret og implementeret ud fra de specifikationer, der kræves i forbindelse med nervesignaler, herunder især et lavt effektforbrug.

Med henblik på generering af signaler til stimulering af nerver anvendes digital til analog konvertere, som kan levere en strømimpuls som udgangssignal. Der er i projektet foretaget design af en current steering D/A konverter. Den konkrete konverter er optimeret mod højere konverteringshastigheder end hvad der er behov for i forbindelse med nervesignaler, men de udviklede designmetoder er generelt anvendelige.

ACKNOWLEDGEMENTS

This research was sponsored by the Danish Medical Research Council and the Grosserer Alfred Nielsen og Hustru Foundation.

I would like to thank my supervisor, Professor Erik Bruun, for making this study possible. Thanks are due to Professor Quiting Huang, ETH Zürich, for accepting me as a guest ph.d. student for my external studies. I am indebted for the technical support ranging from OS setup, IC design to measurement setup, provided by engineer Allan Jørgensen, without whom many tasks would have been arduous.

For the proof reading effort of this thesis, my gratitude is extended to Søren Gaardestrup, Ulrik Wismar and Pier-Andrea Francese.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations and Acronyms

AA	Anti aliasing
AC	Alternating current
A/D	Analog to digital
ADC	Analog to digital converter
BW	Bandwidth
CMFB	Common mode feedback
CMOS	Complementary metal oxide semiconductor
CMRR	Common mode rejection ratio
CMS	Common mode signal
CT	Continuous time
D/A	Digital to analog
DAC	Digital to analog converter
DC	Direct current
DNL	Differential non linearity
DR	Dynamic range
DT	Discrete time
ETF	Error transfer function
FES	Functional electrical stimulation
FFT	Fast fourier transform
FIR	Finite impulse response
GBW	Gain bandwidth
HP	High pass
IC	Inversion coefficient
INL	Integral non linearity
ITF	Integrator transfer function

LLSB	Lower least significant bits
LP	Low pass
LSB	Least significant bit
MASH	Multistage noise shaping
MOST	Metal oxide semiconductor transistor
MSA	Maximum stable amplitude
MSB	Most significant bit
MS/s	Megasamples per second
NEF	Noise excess factor
NMOST	N-channel MOST
NRZ	Non return to zero
NTF	Noise transfer function
OCC	Offset compensation circuit
Opamp	Operational amplifier
OSR	Oversampling ratio
OTA	Operational transconductance amplifier
PDF	Probability density function
PMOST	P-channel MOST
PSD	Power spectral density
PSRR	Power supply rejection ratio
RF	Radio frequency
RZ	Return to zero
SC	Switched capacitor
SFDR	Spurious free dynamic range
SI	Strong inversion
SNR	Signal to noise ratio
SNDR	Signal to noise and distortion ratio

SQNR	Signal to quantization noise ratio
SR	Slew rate
STF	Signal transfer function
SUFR	Signal to update frequency ratio
THD	Total harmonic distortion
ULSB	Upper least significant bits
WI	Weak inversion

CHAPTER 1

INTRODUCTION

In all information processing systems dealing with signals originating in the physical world, sensors form a crucial system component as they serve to collect data at the system front end. As many physical signals are not available as electrical signals, sensors often serve as transducers in order to transform the measured quantity into a voltage or current. The electrical signal post-processing then conditions the sensed signal into a form suitable for monitoring. In systems where feedback is desired in response to a measured signal, a response is evoked through the use of an actuator.

In the recent decades, a steady increase of sensor/actuator systems has been observed as both sensor- and signal processing technology has matured. Examples of sensor systems with feedback can be found in many areas of daily life. Applications found in the automotive industry for safety precautions, such as ABS brakes and airbags, can be mentioned. Many workplaces employ movement detectors used in a feedback manner for illumination control or for after-hours intrusion detection. In the biomedical industry, the continuing advances in processing technology has led to the highly configurable modern day digital hearing aids and cochlear implants, where direct stimulation of the inner ear hearing nerves is done through electrodes. One of the first biomedical applications which was targeted for complete human implantation, is the pacemaker. Pacemakers have evolved from devices performing fixed frequency electrical stimulation of the cardiac sinus node, to the modern-day advanced rate-responsive pacemakers, which senses changes in body movements and sets the heart pace accordingly.

The driver for the development of all these applications, is the electronic revolution the world has witnessed throughout the past 40 years. The continuous miniaturization of transistors has led to the ability to synthesize increasingly complicated systems with still fewer discrete components, leading to complete systems integration on a single chip. One of the biggest success stories in the ongoing miniaturization, is the CMOS technology. Although generally inferior to bipolar technology in terms of analog circuit performance, CMOS is the technology of choice for digital systems. This has led to the availability of inexpensive batch fabrication of systems implemented in standard CMOS processes, sparking high interest in the implementation of high-performance analog systems for integration of analog and digital subsystems on the same chip. Research into integrating CMOS technology with sensor technology has led to the emergence of silicon based microsensors enabling so-called “smart” sensors. These systems typically incorporate not only analog and digital systems, but also the sensor itself. An example of which is the silicon thermal microsensor [1], which converts IR radiation to a voltage through a polysilicon/aluminum thermocouple.

Some sensors are also applicable as actuators. One such example is a biomedical sensor/actuator

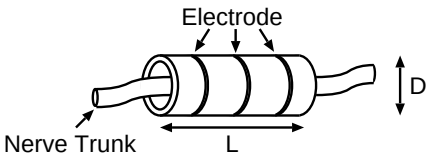


Figure 1.1: Nerve cuff electrode.

type, the so-called nerve cuff electrode, the interfacing of which is the subject of this thesis. The nerve cuff is a sensor/actuator type for neural prosthesis, i.e. for sensing and stimulating nerve bundles. Thousands of individuals sustain damages to the nervous systems each year through, e.g. accidents or stroke. Whereas the limbs of the body may be paralyzed, the nerves and muscles typically remain healthy. The last 20 years, increasing attention has been given to revitalize the paralyzed limbs of patients. One such approach, is to use the natural sensors and muscles of the body by artificial nerve interaction, for which the nerve cuff electrode has proven useful.

A figure depicting a nerve cuff is given in fig. 1.1. The cuff is loosely fitted around the nerve trunk in order not to cause any damage to the nerve. The cuff itself is manufactured from silicone and typically has a length *L*, of 2-4 cm. The diameter *D*, of the cuff is at least 20 % larger than the nerve diameter to avoid pressure damage to the nerve. For signal pickup a number of electrodes are fitted on the inside of the cuff. These electrodes are typically made from stainless steel wire or platinum. The impedance of the cuff electrode is typically in the range of a few $\mathrm{k}\Omega$. A further detailed description of cuff electrodes can be found in [2].

When applied as a sensor the signal pickup results in extremely small voltage signal magnitudes, on the order of $\pm 10\,\mu\mathrm{V}$. The main signal energy lies in a band ranging from 400 Hz to 4kHz [2].

The actuator action of the nerve cuff, stems from neural stimulation through functional electrical stimulation (FES). By applying a signal current to the cuff electrodes, the nerve will be excited and a muscle contraction signal will be evoked. Hence, the actuator action is done by stimulating the paralyzed muscles with the cuff as the artificial nerve signal conveyor.

Some of the successful trials of closed loop systems incorporating cuff electrodes, include bladder control [3] and hand grasp restoration [4]. A neural prosthesis company NeuroDan, a spin-off from Aalborg University, is planning the introduction of a foot drop correction system, ActiGait®, in 2004. In this system, a cuff electrode is fitted around the peroneal nerve enabling flexing of the foot during gait. For foot contact reference, a heel switch is used. The goal of such implanted systems, is to minimize the number of external component. A system proposal for implanted sensor and actuator is given in [5], here the system power is supplied from the outside world through an RF inductive link, through which communication is also done. A system outline of the implant is shown in fig. 1.2. The design of such telemetry powered systems and power management of these, is the subject of [6]. Here, the sensor and actuator both receive their power supply from a central transceiver chip placed just below the skin for good power transfer efficiency. The transceiver also conditions the sensor and actuator control signals for transmission over the RF channel. By partitioning the system, the sensor can be placed sufficiently far from the RF field thus minimizing the risk of desensitizing the preamplifier.

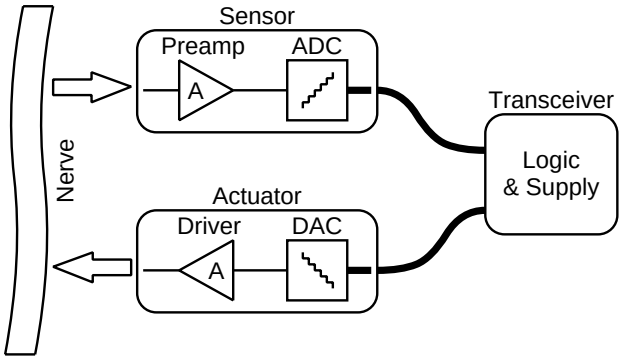


Figure 1.2: Partitioned implantable system for FES. The actuator and sensor are not necessarily connected to the same nerve.

A proposal for such a system poses several design challenges. This thesis treats some of the challenges associated with the realization of the sensor and actuator system components. I.e. the design of preamplifiers is discussed and analog to digital (A/D) conversion and digital to analog (D/A) conversion is treated. In the sensor, a preamplifier should be realized with a sufficient gain to be able to supply the A/D converter (ADC) with a signal of sufficient amplitude. As the input signal has an extremely small magnitude, the preamplifier should have very low internal noise levels in order not to corrupt the signal, whereas exact control of the amplification factor is important to a lesser degree. For all such implanted systems, low power consumption is crucial due to the limits posed by the power transmission method. The maximum available power for the implanted system is typically in the range of a few tens of milliwatts [7, 8].

This thesis is divided into three parts. The first part treats the necessary theory for the circuits implemented during the course of this study. In the second part, the implemented circuits are presented. The design considerations and measured performance are given. The appendices are contained in part three, where the last appendix contains the papers published during the course of this study.

Chapter 2 deals with the theory for employing the chopper modulation technique in low-noise amplifiers. Chopper modulation can be used to eliminate low frequency noise which is prevalent in CMOS circuits. Often for low frequency systems, CMOS low frequency noise will prove to be the bottleneck for noise performance, justifying an investigation of methods to eliminate this problem.

The theory of Nyquist rate D/A conversion is considered in chapter 3. The common specifications used for data converters are introduced. The current steering DAC architecture is thereafter considered. As the cuff electrode is employed as an actuator by delivering current pulses, current steering DACs are well suited for this purpose. Their power efficiency can also be high, as the DAC output current can be delivered directly to the DAC load. The prevalent types of current steering DACs are introduced along with their advantages and drawbacks. The differing techniques employed for achieving high precision is discussed and finally, architecture level causes for DAC dynamic nonlinearities is presented.

A special type of A/D converters are the oversampling converters. These are the subject of chapter 4. Although the concepts introduced in chapter 3 also remain valid for A/D converters, the properties of oversampling converters rely heavily on signal processing concepts and are therefore described

separately. Modeling of these types of converters is discussed and equivalence between discrete-time and continuous-time oversampling converters is covered. Finally, some error effects for oversampling converters with continuous-time loopfilters is discussed.

Two different versions of neural preamplifiers are the subject of chapter 5. The first of which relies on a simple, but power efficient architecture which however does not actively combat the dominant low frequency noise. The second preamplifier presented has a slightly more complex architecture, which however deals with some of the problems of the first design and uses chopper modulation for low frequency noise reduction. For both designs, the design considerations, measurements and derived advantages and drawbacks are presented.

An implemented $\Sigma\Delta$ A/D converter using a continuous-time loopfilter is described in chapter 6. Considerations for power optimization for a set of given design specifications are given for the chosen ADC architecture. The measurements of the implemented circuit is presented and commented.

The final chapter on implementations is chapter 7, where a current steering D/A converter is presented. This DAC accomplishes very high precision while maintaining very high speed, i.e. 200 MS/s. Although the specifications for the implemented DAC do not match with those of a DAC dedicated for implantable circuitry, the current steering DAC is easily scaled for other specifications and the underlying design principles remain valid for such scaling.

Conclusion for this study are drawn in chapter 8.

Part I

Theory

CHAPTER 2

THE CHOPPER MODULATION TECHNIQUE

The chopper modulation technique is historically termed chopper stabilization and has it's roots in applications for vacuum tube amplifier design. The ideas of chopper modulation however, remain valid to this day and are applicable to modern integrated circuit design. The basic idea of chopper modulation is to eliminate low frequency noise and DC offset by modulating these components out of the signal band by the use of an AC-carrier. Whereas sampled techniques such as the auto-zero technique causes downfolding of wideband thermal noise, chopper modulation does not alter wideband noise as no sampling process takes place [9].

This chapter present the mathematical foundation for the chopper modulation technique and illustrates the advantages gained with regards to low frequency noise and DC offset.

2.1 Chopper Modulation Principle

As shown in section A.3, MOS transistor noise has two principal components: Low frequency flicker (1/f) noise and wideband thermal white noise. The noise corner f_{corner}, where the two noise types have equal magnitude typically lies at several kHz, making 1/f-noise the prime contributor to low-frequency noise. Furthermore, DC-offset will be present due to the non-perfect matching of transistors [10]. For recruitment and amplification of very weak sensor signals, the control of these non-idealities are crucial issues.

A method for reducing the effect of 1/f-noise and DC-offset is to employ chopper modulation. The chopper modulation principle is illustrated in both the time- and frequency domain in fig. 2.1. In the time domain the input signal $x(t)$, is multiplied by a square wave signal $c(t)$, which alternates between +1 and -1, prior to being amplified. The thereby resulting chopped signal is effectively frequency shifted to the odd harmonics of the square wave signal fundamental frequency $f_{\text{chop}} = 1/T_{\text{chop}}$. After amplification, the chopped amplifier output signal $x_{\text{chop}}(t)$ is again multiplied by the square signal $c(t)$, effectively "un-inverting" the amplified chopped signal, rendering the amplified output signal $y(t) = Ax(t)$.

In the frequency domain, the input signal $X(f)$ is folded by the square wave function $C(f)$, shifting it's spectrum to the odd harmonics of the square signal. The amplifier with transfer function $A(f)$ is then applied to the frequency shifted input signal. The amplified output signal is then shifted back to the original frequency by folding with $C(f)$ once more, yielding the amplified output signal $Y(f)$. The inherent nonlinear behavior of the chopping technique however, leaves spectral images of the input

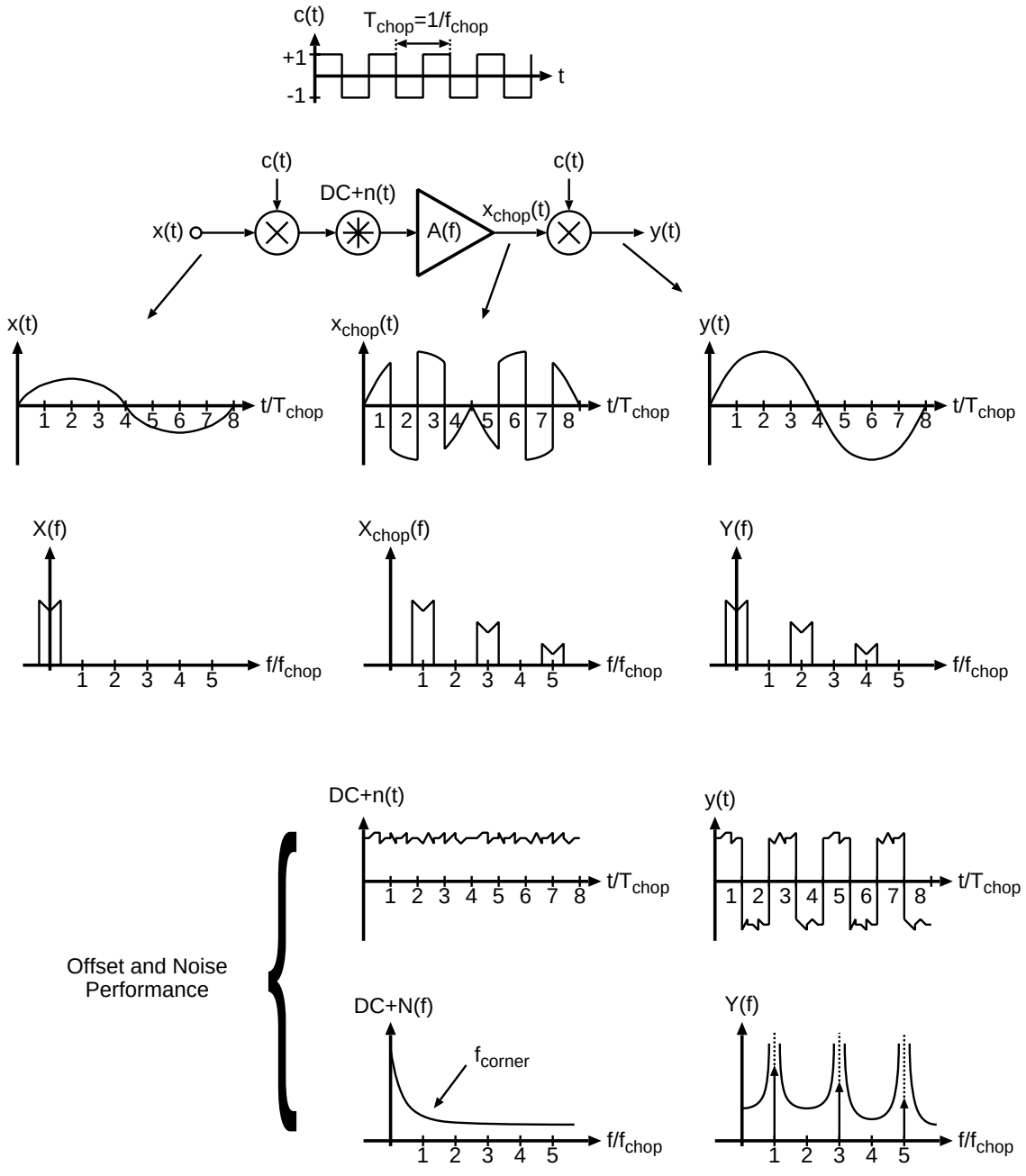


Figure 2.1: Chopper modulation principle illustrated in both the time- and frequency domain.

signal at the even harmonics of the chopping signal fundamental frequency. Hence for faithful signal recreation, a low-pass filter should be employed for suppression of the spectral replicas.

So, the successful amplification of the input signal relies on double modulation of the input signal. By contrast, the noise and DC-offset source is located after the input modulator and is thus only modulated once at the output. Hence, any DC-offset and 1/f-noise will be shifted to the odd harmonics of $c(t)$, leaving the signal baseband free of these components. Only the thermal noise remains in the baseband.

2.2 Chopper Amplifier Transfer Function

In this section we will examine the overall transfer function of the chopper modulated amplifier, in order to resolve the critical issues for employing this technique.

As the chopping signal is periodic with the period T_{chop}, it can be expressed with a Fourier series:

$$c(t) = \sum_{m=-\infty}^{m=+\infty} C(m) e^{j2\pi \frac{m}{T_{\text{chop}}} t} \tag{2.1}$$

In the frequency domain, we have equivalently:

$$C(f) = \sum_{m=-\infty}^{m=+\infty} C(m) \delta\left(f - m f_{\text{chop}}\right) \tag{2.2}$$

where $\delta\left(f\right)$ is the Dirac delta function Assuming a 50% duty cycle for the chopping signal and as the signal can be considered even-sided, it can be shown that the Fourier coefficients are given by:

$$\begin{aligned} C(m) &= \frac{2\sin\left(m\pi/2\right)}{m\pi} \\ \Rightarrow |C(m)| &= \begin{cases} \frac{2}{m\pi} & , m\,\text{odd} \\ 0 & , m\,\text{even} \end{cases} \end{aligned} \tag{2.3}$$

The choice of a 50% duty cycle is optimal as the even order harmonics of the chopping signal are zero in this case, as seen from eq. (2.3).

In the outline presentation of the overall chopping principle given in section 2.1, it was implicitly assumed that the amplifier employed was ideal, i.e. infinite bandwidth and no phase shift. However, we will in the following use the frequency- and time dependent notation for the amplifier, $A(f)$ and $a(t)$ respectively, to be able to model the amplifier non-idealities.

From fig. 2.1, we see that the overall chopped amplifier signal transfer function in the time domain is given by:

$$y(t) = \left[x(t)c(t) * a(t)\right] c(t) \tag{2.4}$$

where the '$*$' operator denotes the convolution operation. The equivalent frequency domain expression is:

$$Y(f) = \left[X(f) * C(f)\right] A(f) * C(f) \tag{2.5}$$

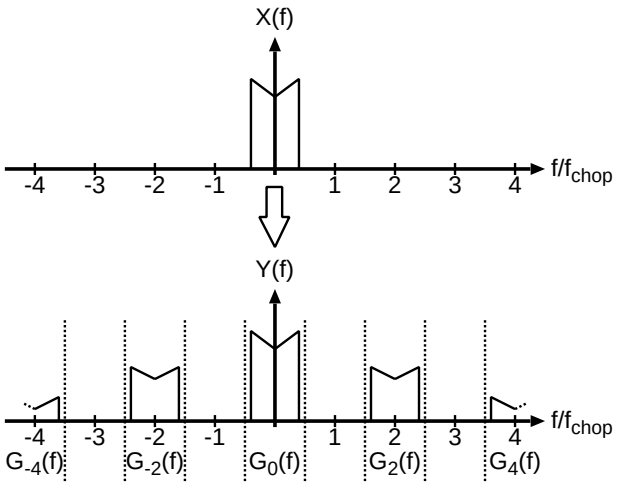


Figure 2.2: Overall chopper amplifier transfer function.

Inserting the Fourier coefficients of eq. (2.3) in the frequency domain representation of the chopping signal given in eq. (2.2) and solving for the resulting signal output given in eq. (2.5) yields:

$$Y(f) = \left(\frac{2}{\pi}\right)^2 \sum_{\substack{n=-\infty \\ n\,\mathrm{odd}}}^{n=+\infty} \frac{1}{n} A\left(f - nf_{\mathrm{chop}}\right) \sum_{\substack{m=-\infty \\ m\,\mathrm{odd}}}^{m=+\infty} \frac{1}{m} X\left(f - mf_{\mathrm{chop}} - nf_{\mathrm{chop}}\right) \tag{2.6}$$

A general description of the chopper amplifier can be obtained by substituting $k = m + n$ into eq. (2.6). Hereby an infinite sum of linear transfer functions can be obtained:

$$Y(f) = \sum_{\substack{k=-\infty \\ k\,\mathrm{even}}}^{k=+\infty} \underbrace{\left(\frac{2}{\pi}\right)^2 \sum_{\substack{n=-\infty \\ n\,\mathrm{odd}}}^{n=+\infty} \frac{1}{n}\left(\frac{1}{k-n}\right) A\left(f - nf_{\mathrm{chop}}\right)}_{G_{\mathrm{k}}(f)} \quad X\left(f - kf_{\mathrm{chop}}\right) \tag{2.7}$$

The resulting expression describing the overall chopper output is thus given by:

$$Y(f) = \sum_{\substack{k=-\infty \\ k\,\mathrm{even}}}^{k=+\infty} G_{\mathrm{k}}\left(f\right) X\left(f - kf_{\mathrm{chop}}\right)$$

In fig. 2.2, the overall chopper output is illustrated showing the contributions for each of the spectral replicas produced by $G_{\mathrm{k}}(f)$, $k \neq 0$, for $|k| \leq 4$. The 50% duty cycle renders all the odd order transfer

functions equal to zero. In the figure, it is assumed that the input signal has a bandwidth restricted to less than half the chopping frequency. If the input signal has a bandwidth exceeding the chopping frequency, aliasing would occur and correct signal reconstruction would not be achievable.

Of the linear combination of transfer functions, $G_0(f)$ is the transfer function of interest as it conveys the baseband signal. From eq. (2.7) we have:

$$G_0(f) = \left(\frac{2}{\pi}\right)^2 \sum_{\substack{n=-\infty \\ n\,\text{odd}}}^{n=+\infty} \left(\frac{1}{n}\right)^2 A\left(f - nf_{\text{chop}}\right) \tag{2.8}$$

Any signal delay that exists between the choppers will cause the chopper amplifier output signal to be reconstructed non-optimally. The delay effect can be examined by applying an amplifier transfer function with linear phase response:

$$A(f) = A_0 e^{-j2\pi f \Delta t} \tag{2.9}$$

where A_0 is the amplifier DC-gain. Inserting eq. (2.9) in the baseband transfer function $G_0(f)$ and simplifying yields:

$$\begin{aligned} G_0(f) &= \underbrace{A_0 e^{-j2\pi f \Delta t}}_{A(f)} \underbrace{\left(\frac{2}{\pi}\right)^2 \sum_{\substack{n=-\infty \\ n\,\text{odd}}}^{n=+\infty} \left(\frac{1}{n}\right)^2 e^{-j2\pi \frac{\Delta t}{T_{\text{chop}}} n}}_{\text{tri}\left(\frac{\Delta t}{T_{\text{chop}}}\right)} \end{aligned} \tag{2.10}$$

So we see that by inserting an amplifier with linear phase delay, the resulting chopper amplifier transfer function becomes the amplifier transfer function $A(f)$ modified by a function denoted $\text{tri}\left(\frac{\Delta t}{T_{\text{chop}}}\right)$ in the above expression. It can be shown that $\text{tri}\left(\frac{\Delta t}{T_{\text{chop}}}\right)$ is in fact the Fourier series of a triangular function defined by [1]:

$$\begin{aligned} \text{tri}(x) &= 1 - 4\,|x| \quad , |x| \leq 1/2 \\ \text{tri}(x+n) &= \text{tri}(x) \quad , n \in \mathbb{N} \end{aligned} \tag{2.11}$$

Hence for $\frac{\Delta t}{T_{\text{chop}}} \leq 1/2$, eq. (2.10) can be simplified to:

$$G_0(f) = A(f)\left(1 - 4\left|\frac{\Delta t}{T_{\text{chop}}}\right|\right) \tag{2.12}$$

From the obtained expression, we see that the effective gain of the chopper modulated amplifier is sensitive to any phase shift induced by the amplifier. So a parameter in the design process is to choose a chopping frequency sufficiently below the amplifier 3-dB frequency in order to ensure that the overall gain does not deteriorate significantly.

2.3 Chopper Amplifier Noise Performance

All noise sources of the amplifier including DC offset can be referred to the amplifier input, represented by the sole noise source given in fig. 2.1. From the schematic in fig.. 2.1, we see that the chopper modulated amplifier noise output is given by:

$$y(t) = (a(t) * n(t))\, c(t) \tag{2.13}$$

For an ideal amplifier, i.e. having infinite bandwidth, eq. (2.13) simplifies to:

$$y(t) = A_0 n(t) c(t) \tag{2.14}$$

The resulting chopper amplifier output noise power spectral density (PSD) can be found from the Fourier transform of the output autocorrelation function. Assuming a stochastic stationary process for the noise function, the output autocorrelation function $R_{\mathrm{y}}(\tau)$, can be resolved to:

$$\begin{aligned} R_{\mathrm{y}}(\tau) &= E\{y(t)y(t+\tau)\} \\ &= E\{A_0 n(t)c(t)A_0 n(t+\tau)c(t+\tau)\} \\ &= A_0^2 E\{n(t)n(t+\tau)\}\, E\{c(t)c(t+\tau)\} \\ &= A_0^2 R_{\mathrm{n}}(\tau) R_{\mathrm{c}}(\tau) \end{aligned} \tag{2.15}$$

where $E\{\cdot\}$ is the statistical expectation operator and we have used the fact that $c(t)$ and $n(t)$ are orthogonal. The equivalent PSD function $S_{\mathrm{y}}(f)$, for the frequency domain can be found from the Fourier transform of eq. (2.15):

$$\begin{aligned} S_{\mathrm{y}}(f) &= \mathcal{F}\{R_{\mathrm{y}}(\tau)\} \\ &= A_0^2 \mathcal{F}\{R_{\mathrm{c}}(\tau)\}\, \mathcal{F}\{R_{\mathrm{n}}(\tau)\} \\ &= A_0^2 \sum_{\substack{n=-\infty \\ n\,\mathrm{odd}}}^{n=+\infty} |C(m)|^2\, \delta(f - n f_{\mathrm{chop}})\, S_{\mathrm{n}}(f) \end{aligned} \tag{2.16}$$

Inserting the Fourier coefficient as defined in eq. (2.3) in eq. (2.16) we arrive at the noise PSD function for the chopper amplifier:

$$S_{\mathrm{y}}(f) = A_0^2 \left(\frac{2}{\pi}\right)^2 \sum_{\substack{n=-\infty \\ n\,\mathrm{odd}}}^{n=+\infty} \frac{1}{n^2} S_{\mathrm{n}}(f - n f_{\mathrm{chop}}) \tag{2.17}$$

In the following we will discuss the impact of chopper modulation on the two noise types presented in section A.3.

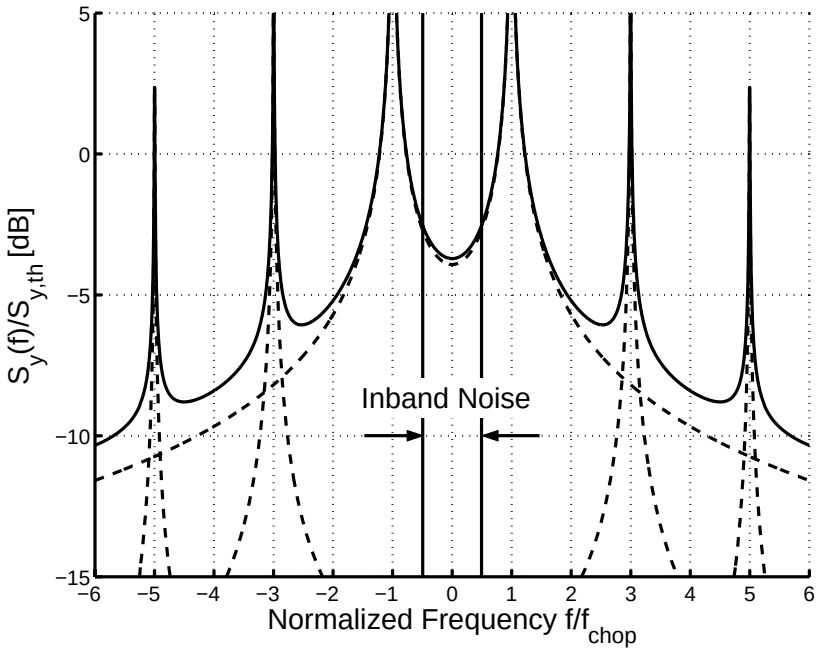


Figure 2.3: Chopper amplifier output 1/f-noise PSD.

2.3.1 Thermal Noise

The impact of chopping on thermal noise is most easily shown using autocorrelation functions. Using eq. (2.15) the chopper amplifier output noise power can be found at $\tau = 0$ using Parseval's theorem [11]:

$$P_{\mathrm{y}} = A_0^2 P_{\mathrm{n}} P_{\mathrm{c}} = A_0^2 P_{\mathrm{n}} \tag{2.18}$$

Since the chopping signal alternates between ± 1, the chopping signal power is: $P_{\mathrm{c}} = 1$, giving us the last identity in eq. (2.18). Examining the amplifier noise output power without applying any chopper modulation gives us:

$$P_{\mathrm{y}} = E\left\{A_0 n(t) A_0 n(t+\tau)\right\}|_{\tau=0} = A_0^2 P_{\mathrm{n}} \tag{2.19}$$

The result given in eq. (2.19) for the amplifier employing no chopping, is exactly the same as the result in eq. (2.18) for the chopper modulated amplifier. This result allows us to conclude that chopper modulation *does not* alter the thermal noise. The resulting chopper amplifier output thermal noise in the frequency domain can accordingly be written: $S_{\mathrm{y,\,th}}(f) = A_0^2 S_0$.

2.3.2 Flicker Noise

In section A.3, an expression for flicker (1/f) noise is given. For the coming analysis a useful expression for the 1/f-noise is obtained by referring the 1/f-noise to the thermal noise floor and the corner frequency:

$$S_{1/\mathrm{f}}(f) = S_0 \frac{f_{\mathrm{c}}}{|f|} \tag{2.20}$$

where f_c is the 1/f-noise corner frequency and S_0 is the thermal noise floor. The chopper amplifier output PSD for 1/f-noise can be found by inserting eq. (2.20) in the expression for the resulting output PSD found in eq. (2.17):

$$S_{y,\,1/f}(f) = \left(\frac{2A_0}{\pi}\right)^2 \sum_{\substack{n=-\infty \\ n\,\mathrm{odd}}}^{n=+\infty} \frac{1}{n^2} \frac{S_0 f_c}{|f - n f_{chop}|} \tag{2.21}$$

This result is illustrated in fig. 2.3 including the foldover components for *n*=1, 3 and 5. The individual foldover component are indicated by the dashed lines whereas the total summed noise is represented by the solid line. The output noise PSD is normalized to the output thermal noise floor given by $S_{y,\,th}(f) = A_0^2 S_0$. In the figure we have set the noise corner f_c, equal to half the chopping frequency f_{chop} and set the inband frequency range to half the chopping frequency as indicated by the solid-line delimiters.

We see from the figure, that chopper modulation in the case of 1/f-noise causes aliasing resulting in an increase of the effective thermal noise $S_{th,\,eff}(f)$, defined as the inband output thermal noise and the aliased 1/f-noise components which are in excess:

$$S_{th,\,eff}(f) = S_{y,\,th}(f) + S_{y,\,1/f}(f)\big|_{inband} \tag{2.22}$$

However the largest contributor by far of fold-over noise components stems from the first order terms $n = \pm 1$. Solving eq. (2.21) numerically reveals that in excess of 95% of the aliased noise can be accounted for by the first order terms. The third order terms $n = \pm 3$ account for approx. 3% rendering all higher order terms negligible.

An approximate expression for the resulting chopper amplifier output 1/f-noise floor within the baseband can therefore be extracted from eq. (2.21) by including only the first order terms and taking their magnitude at DC assuming a flat 1/f-noise contribution:

$$S_{y,\,1/f}(f) \approx A_0^2 S_0 \frac{8}{\pi^2} \frac{f_c}{f_{chop}} \tag{2.23}$$

which is seen to result in a frequency independent noise component. Collecting the two terms constituting the resultant chopper amplifier output effective thermal noise, we have:

$$S_{th,\,eff}(f) \approx A_0^2 S_0 \left(1 + \frac{8}{\pi^2} \frac{f_c}{f_{chop}}\right)$$

In fig. 2.4 the output excess noise due to downfolded 1/f-noise relative to the output thermal noise is depicted in percent versus chopping frequency relative to the 1/f-noise corner f_c. The solid line is the excess noise numerically evaluated applying eq. (2.21), where up till the ninth term $n = \pm 9$ has been included. The dashed line is the approximate expression given in eq. (2.23). It is seen that the approximate expression given in eq. (2.23) gives us a result which is precise within a few percent,

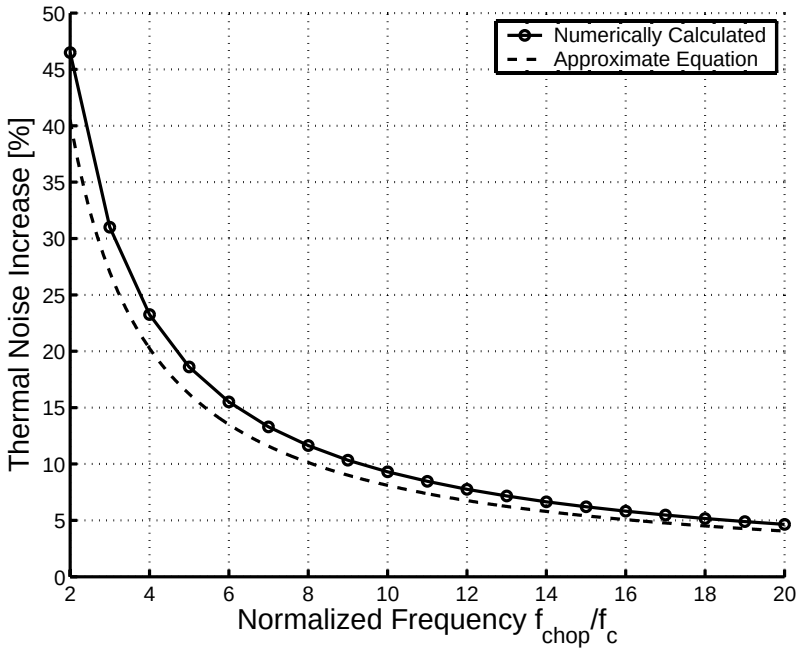


Figure 2.4: Excess noise relative to thermal noise floor added by downfolded 1/f-noise.

especially for larger $f_{\mathrm{chop}}/f_{\mathrm{c}}$ - ratios where the downfolded 1/f-noise attains a more flat shape inband.

With regards to chopping frequency versus noise corner frequency, fig. 2.4 reveals that less than 10% excess noise can be expected at chopping frequencies $f_{\mathrm{chop}} > 10f_{\mathrm{c}}$, which in most cases would be in the tens of kHz range, i.e. relatively low frequencies.

2.4 Summary

In this chapter we have investigated the effect of chopping on both the effective transfer function of the overall chopped amplifier and the effect on thermal- and 1/f-noise respectively.

For the transfer function, a general expression was found which revealed that scaled spectral replicas will be present at the output at all even multiples of the chopping frequency. Inserting an amplifier transfer function with linear phase, showed that the baseband transfer function is sensitive toward any delay induced between the modulators having a gain linearly reduced by the delay.

Finally, investigating the noise properties revealed that the chopped amplifier does not alter the level of the thermal noise floor. The 1/f-noise however, will be aliased by the chopping process, resulting in increased noise in the baseband. The term effective thermal noise was used to quantify the excess noise assuming a flat noise floor across the baseband. An approximate formula for predicting the effective thermal noise was shown to give a good estimate of the expected performance. Furthermore, numerical analysis showed that less than 10% increase in effective thermal noise, as compared to the un-chopped amplifier, can be expected for chopping frequencies in excess of ten times the noise corner f_{c}.

CHAPTER 3

NYQUIST RATE D/A CONVERSION

In this chapter, the fundamental principles of data conversion is presented. The type of converters considered in this chapter have sampling frequencies f_{s}, limited by the Nyquist rate, i.e. $f_{\mathrm{s}} = 2f_{\mathrm{b}}$, where f_{b} is the signal band frequency. The common measures used for specifying D/A converters in terms of their static and dynamic performance are explained. These measures are also valid for A/D converters. However the ADCs considered in this thesis utilize oversampling, the principle of which differs significantly from Nyquist rate data conversion and is the subject of chapter 4.

A specific DAC architecture is presented, the current steering DAC. This DAC type is appropriate for cuff electrodes due to their ability to deliver the stimulation current directly to the cuff. Two different design approaches for obtaining high precision in the current-steering DAC architecture is presented: Statistical design techniques, where knowledge of the distribution of die imperfections is used for the cancellation of these and calibration techniques, where current sources are trimmed against a reference for high precision.

Finally, some of the dynamic non-idealities are presented and techniques for alleviating the impact of these are presented.

3.1 Ideal D/A-Converter

A block diagram of an ideal D/A converter is shown in fig. 3.1 alongside with the desired output voltage characteristic.

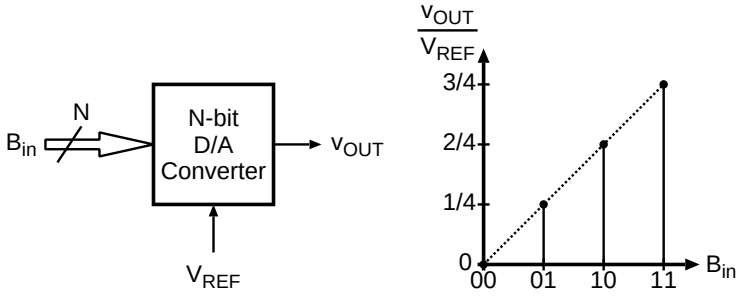


Figure 3.1: DAC block diagram and ideal transfer function.

The DAC input is an N-bit digital word B_{in}, given by:

$$B_{\mathrm{in}} = b_1 2^{-1} + b_2 2^{-2} + \cdots + b_{\mathrm{N}} 2^{-N} \tag{3.1}$$

We define the first bit in the input word b_1, to be the most significant bit (MSB) and the last bit b_{N}, as the least significant bit (LSB). The output voltage of the DAC v_{OUT}, is referred to a voltage reference V_{REF}, in such a manner that:

$$v_{\mathrm{OUT}} = V_{\mathrm{REF}} \left(b_1 2^{-1} + b_2 2^{-2} + \cdots + b_{\mathrm{N}} 2^{-N} \right) = V_{\mathrm{REF}} B_{\mathrm{in}} \tag{3.2}$$

In these definitions we employ voltage as the signal medium, however this could be any other quantity, e.g. current or charge. An example ideal output for a two-bit DAC is shown alongside the block diagram in fig. 3.1. The minimum voltage change in the output when one LSB changes is:

$$V_{\mathrm{LSB}} = \frac{V_{\mathrm{REF}}}{2^N - 1} \tag{3.3}$$

Commonly, the performance of a DAC is measured in terms of the number of LSB steps independent of the physical signal quantity. Hence a common unit is simply the unitless: $1\,\mathrm{LSB} = 1/2^N$. The '-1' is often omitted for convenience in calculations for high resolution converters.

3.2 D/A Converter Specifications

In order to be able to characterize a DAC in terms of it's static and dynamic performance, a set of metrics are common. These are the subject of this section.

3.2.1 Static Specifications

The ideal output transfer function as depicted in fig. 3.1 is not achievable when implementing a DAC. Due to mismatches inherent to the technology employed, some deviances will always be present [10]. Mismatches which causes deviances in the parameters of the individual transistors on chip, typically do not change significantly over time and are thus termed static. Some of the common parameters for static characterization are:

Gain error and offset

If we define the DAC output in terms of LSBs, we find that the ideal DAC has a transfer function slope of one. If we define a best fit linear curve for describing the DAC output points, the deviances of this slope compared to the ideal output slope are termed gain error. Similarly, if the best fit linear curve does not have the same origin as the ideal output curve, the difference is termed offset. These effects are however linear and are usually not crucial for most applications.

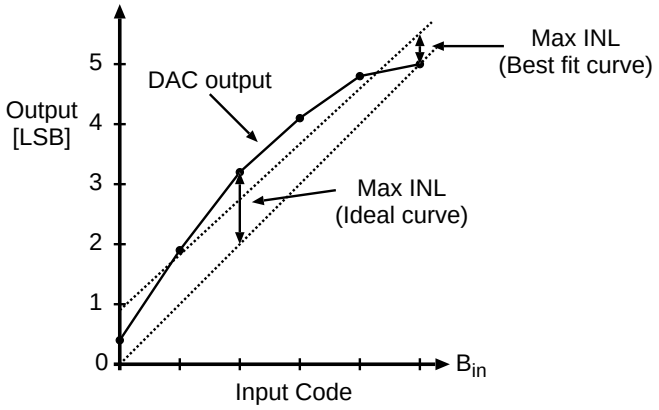


Figure 3.2: Integral nonlinearity error.

INL

The integral nonlinearity error (INL) is defined as the true output curve deviance from the ideal output curve for each input code. The INL for a given input code n, in terms of LSBs can thus be found:

$$\mathrm{INL}(n) = \frac{v_{\mathrm{OUT}}(n) - v_{\mathrm{OUT,ideal}}(n)}{V_{\mathrm{LSB}}} \tag{3.4}$$

Using the ideal DAC output curve is a conservative measure of the INL, and typically a best fit DAC transfer function is used instead as shown in fig. 3.2, where the DAC output transfer characteristic displays bowing. Here the conservative INL measure gives us an $\mathrm{INL}_{\max} > 1\,\mathrm{LSB}$, whereas using the best fit transfer function we have $\mathrm{INL}_{\max} \leq 1/2\,\mathrm{LSB}$. Though the best fit curve has both offset and gain error.

DNL

For an ideal DAC, each analog output step is equal to 1 LSB. A measure for deviance from the ideal step value is the differential nonlinearity. The DNL at input code n, is defined as:

$$\mathrm{DNL}(n) = \frac{v_{\mathrm{OUT}}(n) - v_{\mathrm{OUT}}(n-1)}{V_{\mathrm{LSB}}} - 1$$

and thus expresses the difference from the ideal step size in terms of LSBs

Monotonicity

A DAC can be said to be monotonic if the output step size Δ, fulfills $\Delta \geq 0$, for each increase in the DAC input code by a single step, over the entire DAC input range. In terms of the INL, we see that a DAC is guaranteed to be monotonic if we have: $\mathrm{INL}_{\max} \leq 1/2$ [12].

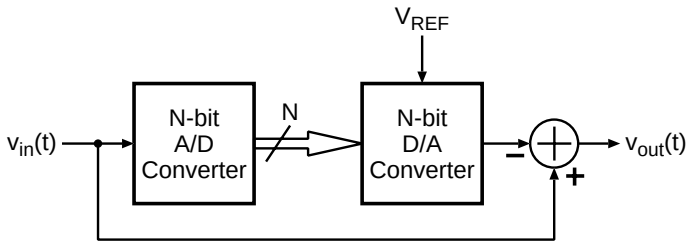


Figure 3.3: Quantization noise extraction circuit.

3.2.2 Dynamic Specifications

As the static specifications for a DAC were constant with time, a set of dynamic specifications are employed when the transient behavior between changing states is to be evaluated. Since the DAC output is continuous in time, the transient behavior is vital for the overall performance of a DAC. A brief introduction to some common dynamic measures is the subject of this section.

SNR

Any quantized signal will always be an imperfect representation of the corresponding analog signal. The resulting difference signal can be extracted using the circuit shown in fig. 3.3, for some arbitrary number of bits N. The resulting error signal is termed the quantization noise. The quantization noise is commonly modeled as a white noise source with a uniform probability density function (PDF). For a given LSB step size, the power of the quantization noise can be shown to be [13]:

$$P_{\mathrm{q}} = \frac{V_{\mathrm{LSB}}^2}{12} \tag{3.5}$$

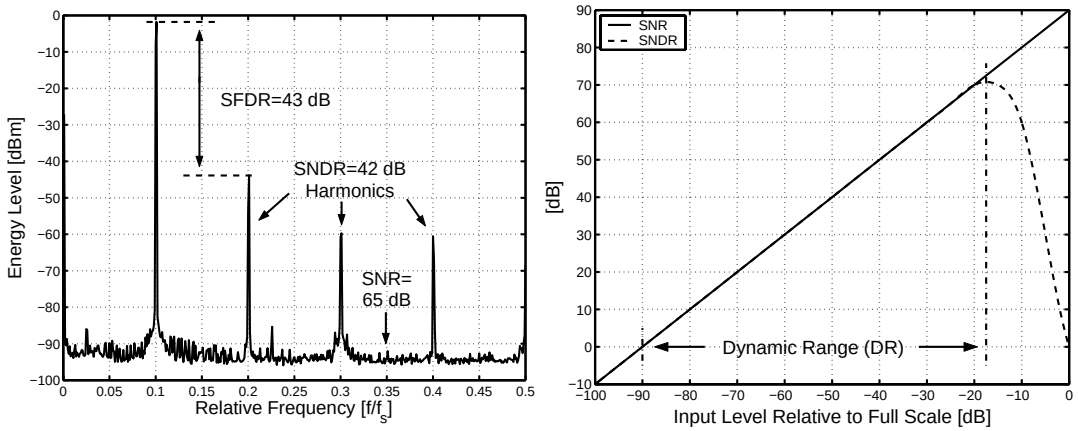
The signal to quantization noise ratio (SQNR) is then defined by the ratio of the signal and the corresponding quantization noise:

$$\mathrm{SQNR} = 10 \log \left(\frac{P_{\mathrm{s}}}{P_{\mathrm{q}}} \right) \tag{3.6}$$

Assume that we have a sinusoid signal with a peak-peak voltage V_{ref}. The corresponding SQNR for N bits of resolution is found to be:

$$\mathrm{SQNR} = 10 \log \left(\frac{\frac{1}{8} V_{\mathrm{ref}}^2}{\frac{1}{12} \left(\frac{V_{\mathrm{ref}}}{2^N} \right)^2} \right) = 6.02N + 1.76 \,[\mathrm{dB}] \tag{3.7}$$

which shows us that the for each bit added to the DAC (ADC), an increase of approx. 6 dB can be expected in the SQNR. The term, signal to noise ratio (SNR), is loosely used for SQNR, but actually encompasses all noise types and is not specific to data converter systems. Whereas the SQNR term is a metric used strictly for systems employing quantization.



(a) Example DAC output spectrum.

(b) SNR, SNDR and DR.

Figure 3.4: Common dynamic measures for data converters

SNDR

Any signal processing system may exhibit nonlinearity. E.g. the example DAC transfer function in fig. 3.2 which showed bowing of the transfer function. Any such nonlinearity will result in harmonic distortion, i.e. frequency components appearing at the multiples of the input frequency. The signal to noise and distortion ratio (SNDR) is then simply calculated as the SNR including the harmonic power in the noise power. This also implies that the SNR is typically calculated excluding the harmonic components.

An example output spectrum from an actual realized DAC is shown in fig. 3.4 (a). Here, the nonlinearities of the system result in strong harmonics of both even and odd order and thus sharply deteriorate the SNDR compared to the SNR. The harmonics are usually dependent on the input level of the data converter and thus the SNDR will track the SNR until a compression point is reached, where the harmonic power exceeds the noise level as a function of the input level. This is shown qualitatively in fig. 3.4 (b).

SFDR

A metric often used for DACs employed in communication systems is the spurious free dynamic range (SFDR). The SFDR is defined as the difference in level between the signal output energy and the highest output spur energy when the maximum input signal is applied to the system. Typically, one of the low order harmonics will determine the SFDR as shown for the example spectrum in fig. 3.4 (a). In communication systems, the bandwidth of the DAC is often split into channels, which stresses the importance of a good SFDR in these systems as signals relating to one channel otherwise could leak into adjacent channels.

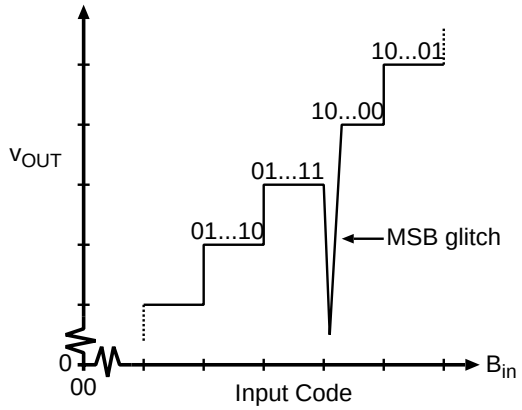


Figure 3.5: Major carry glitch for MSB transition.

DR

The dynamic range (DR) is defined from the SNR. The DR is the magnitude ratio between the largest sinusoidal input which does not overload or cause severe nonlinear effects in the converter to the the magnitude which provides a SNR of zero. The DR is shown qualitatively in fig. 3.4 (b).

Glitches

Whereas the other dynamic benchmarks have been determined in the frequency domain, the glitch energy is a time domain dynamic measure. Glitches appear during a carry transition when the output signal corresponding to the new input value appears either before or after the previous input code output value. The largest glitch usually appears for the carry transition of the MSB bit as shown qualitatively in fig. 3.5, where a glitch is generated at mid-code for a DAC. Thus any slight skew in the input signals can potentially result in glitches.

The glitches will result in an increased noise floor [12], thus deteriorating the SNR. If the glitch energy is non-proportional to the input signal (which is usually the case), harmonic distortion will be generated from the glitches, degenerating the SFDR.

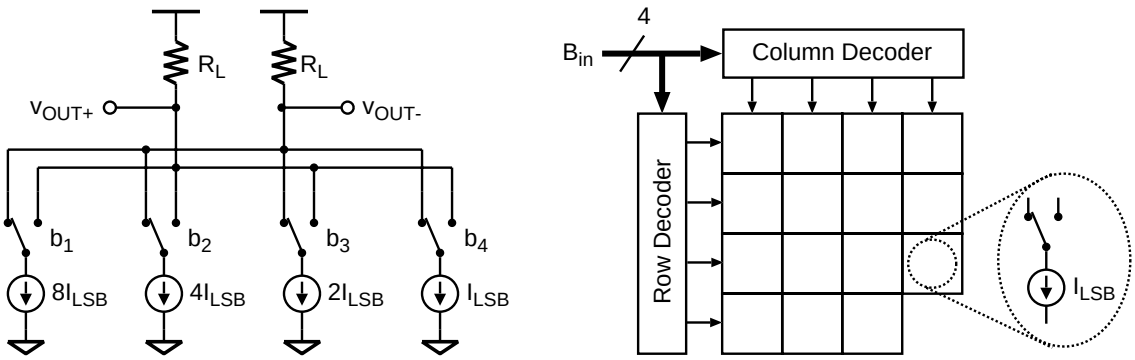
A method for evaluating glitch performance for a DAC is proposed in [12]. Here, the energy of 1 LSB: $E_{\mathrm{LSB}} = V_{\mathrm{REF}}/2^N \cdot T$, is compared to the glitch energy E_{glitch}. The ratio between the two is then used as a measure for glitch performance:

$$G = \frac{E_{\mathrm{glitch}}}{E_{\mathrm{LSB}}} \tag{3.8}$$

As a rule of thumb, the ratio G should remain below $1/2$ [12].

3.3 High-Precision D/A Converters

Numerous different architectures are available for D/A conversion. Of these can be mentioned the voltage division DAC based on resistor ladders, the charge redistribution DAC and the current steering



(a) Binary 4-bit Current Steering DAC

(b) Thermometer 4-bit Current Steering DAC

Figure 3.6: Basic current steering DAC architectures, fully binary and fully segmented.

DAC. Of the multiple different architectures, the current steering DAC is popular due to it's ability to obtain both high accuracy and high speed. In this thesis we are concerned with the implementation DACs based on the current steering technique, which will be the subject henceforth. Discussion of other DAC types can be found in many standard electronics textbooks, e.g. [13].

3.3.1 Current Steering Converters

The basic principle behind the current steering DAC is shown in fig. 3.6 (a). Here, a series of binary scaled current sources are fed to a resistive load, the voltage of which constitute either the positive or the negative voltage. The digital input word is in the binary current steering DAC fed directly to the corresponding current source. One of the advantages of the current steering architecture, is it's ability to drive a resistive load directly without necessitating an output buffer. This fact coupled with it's simple structure, enables the current steering DAC to operate at very high update frequencies.

The basic binary current steering DAC suffers from glitch problems though. Consider the mid-code transition 0111-1000 for the 4-bit example binary DAC: We see that in order to avoid glitching, bits $b_2 - b_4$ have to switch *exactly* at same time instant as b_1. To avoid any skew between the digital control signals is extremely difficult, especially at high update frequencies.

A method for alleviating this weakness, is to employ what is know as a fully segmented design, an example of which is shown in fig. 3.6 (b). Here, the DAC current sources are split into $2^{\mathrm{N}} - 1$ unit current source of magnitude I_{LSB}. The digital binary input word is decoded into what is known as thermometer code prior to selecting the switch position for any of the unit current sources. The thermometer code corresponding to a three-bit binary input word is shown for all values in table 3.1.

From the resulting thermometer decoded word D, it is seen that for all transitions from one input word to the adjacent, only a single bit will change. Hence no critical major carry glitches will occur for this scheme. Also, as the current is always increased with increasing input code, monotonicity is guaranteed for thermometer DACs.

The drawback of using thermometer decoded DACs is the increase in decoding complexity and area since for binary input words of N bits, $2^N - 1$ unit current sources are needed. Typically, DACs with

Decimal value	Binary code $b_1b_2b_3$	Thermometer code $d_1d_2d_3d_4d_5d_6d_7$
0	000	0000000
1	001	0000001
2	010	0000011
3	011	0000111
4	100	0001111
5	101	0011111
6	110	0111111
7	111	1111111

Table 3.1: Thermometer encoding of digital words.

high number of bits (>10) only employ thermometer decoding for the upper MSBs, whereas a binary sub-DAC is used for the lower LSBs. This type of partitioning is termed segmentation, i.e. a fully thermometer decoded DAC is called fully segmented. Using segmentation decreases circuit area and complexity while the price paid is increased glitch sensitivity for the binary decoded sub-DAC, which however is not so severe for the lower LSBs and thus deemed acceptable.

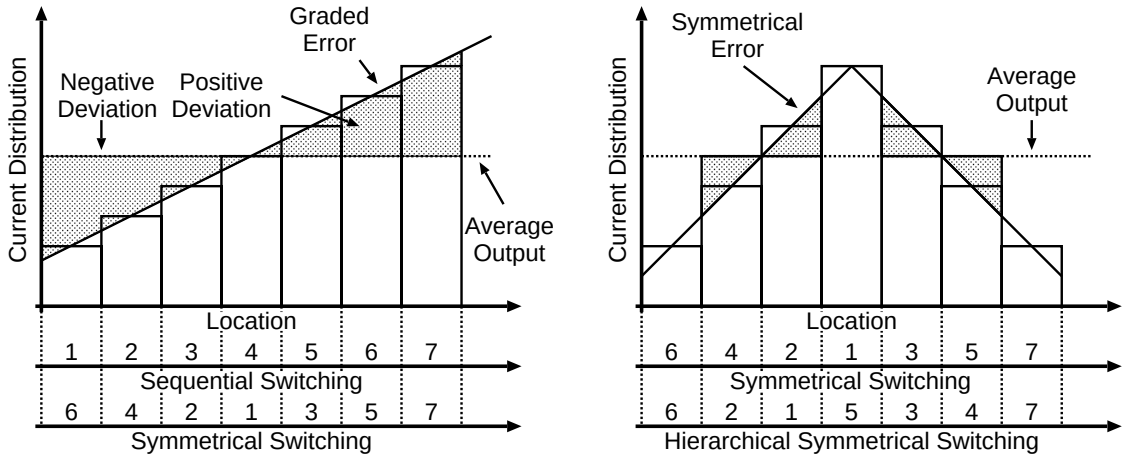
Achieving accuracy of DACs beyond 10 bits has proven to be a difficult task [14, 15, 16]. The intrinsic mismatch in the technology [10], causes errors across the implemented matrix of current sources and is thus one of the main considerations when implementing high-resolution DACs. Numerous different methods to circumvent intrinsic mismatches exist, but can roughly be classified into two different categories which are the subject of the next two sections.

3.3.2 Statistical Design Techniques

From the knowledge of the statistical properties of the mismatch error distribution across the die, special design techniques can be utilized to cancel these. The subject of this section is a brief introduction to some of the schemes used in the statistical design of DACs.

Although mismatch errors are distributed in a two-dimensional manner across the die surface, a one-dimensional model for the errors suffices for our discussion. Consider the exaggerated error distribution of current as a function of location across an array of sources as shown in fig. 3.7 (a). Here, a simple case is considered for a three bit current source array. The array has some average output current value, from which a negative and a positive deviation can be defined. If we apply a digital ramp input to a DAC suffering from graded error and if the individual cells in the array are switched in a sequential manner, it is clearly seen that the error will accumulate, giving a poor INL.

To alleviate this effect, a scheme known as symmetrical switching is introduced in [17], which uses a switching order as shown at the bottom of fig. 3.7 (a). By switching the cells symmetrically around the graded error mid-point, the positive and negative deviations effectively cancel each other other. This technique is used in [18], to realize a 16-bit DAC in a 6-10 segmented design, i.e. the 6 MSBs are thermometer decoded whereas the remaining 10 bits are binary weighted and thus rely on intrinsic matching.



(a) Graded error.

(b) Symmetrical error.

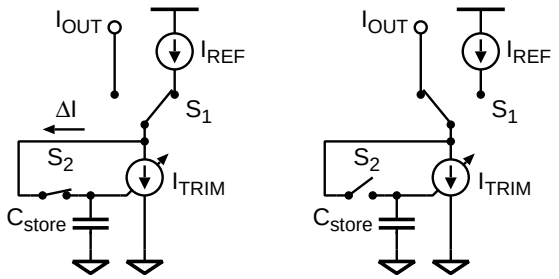
Figure 3.7: Graded and symmetrical errors across current source arrays.

Another mismatch error type typically present in current source arrays is symmetrical error as depicted in fig. 3.7 (b). Using symmetrical switching will *not* eliminate this error source. To mitigate this issue, an alternative switching scheme proposed in [19], known as hierarchical symmetrical switching, can be employed. At the bottom of fig. 3.7 (b), this switching scheme is shown for our simple example case. Rather than switching solely symmetrically around the global input code midpoint, this technique also employs local symmetrical switching as seen for sources 1-2 and 3-4. In this manner the symmetrical error is diminished. The graded error performance is however still retained as the locally switched pairs are symmetrically matched with opposing switched pairs, e.g. current source pairs 1-2 are matched with pairs 3-4.

Many other switching schemes have since been proposed, employing increasingly complex algorithms for the switching sequence. Of the more recent high-performance switching DACs, the DAC presented in [20] can be mentioned. This paper presents an 8-6 segmented 14-bit DAC with a maximum update rate of $150\,\mathrm{MS/s}$. The static measurements report that both the INL and the DNL are below 1/2 LSB. The dynamic performance is given as a SFDR of 84 dB for a signal to update frequency ratio (SUFR) of 1/300, dropping to 61 dB at a SUFR of 1/30. Both measurements are done at an update rate of 150 MS/s. The drawback in this design is the increased logic complexity necessary for decoding, which is however enabled for high speeds with modern day standard digital libraries.

3.3.3 Calibration/Trimming Design Techniques

An altogether different approach for obtaining the prescribed precision, is to employ a scheme which utilizes calibration of the individual current sources. The basic concept of calibrated current sources is shown in fig. 3.8, where the trimmable current source is voltage controlled. During the calibration phase of the current source, a reference current is fed to the trimmable current source. Any deviance $\Delta I = I_{\mathrm{REF}} - I_{\mathrm{TRIM}}$, between the reference current and the trimmed output current is integrated on the control voltage holding capacitor C_{store}, hence regulating the trimmable current source till it equals



(a) Calibration (b) Operation

Figure 3.8: Calibrated current source principle.

the reference current. For such a scheme, the trimmed current source can be implemented by a single transistor.

As intrinsic matching can be obtained for resolutions equivalent to 9-10 bits [21], only the MSBs exceeding the intrinsic matching accuracy are typically calibrated, e.g. for a 14-bit DAC, only the top 4-5 MSBs would typically be calibrated. The remaining LSBs would then be implemented as a sub-DAC implementing the necessary current division of one of the calibrated MSB. A schematic illustrating such a scheme is depicted in fig. 3.9 (trim circuit and storage capacitor omitted). As the trimmed MSB sources need to be of unit size for trimmability against a single reference, they are well suited of thermometer decoding. The sub-DAC could be implemented as any of the previously discussed current-steering DAC types and thus constitute a segmented DAC itself.

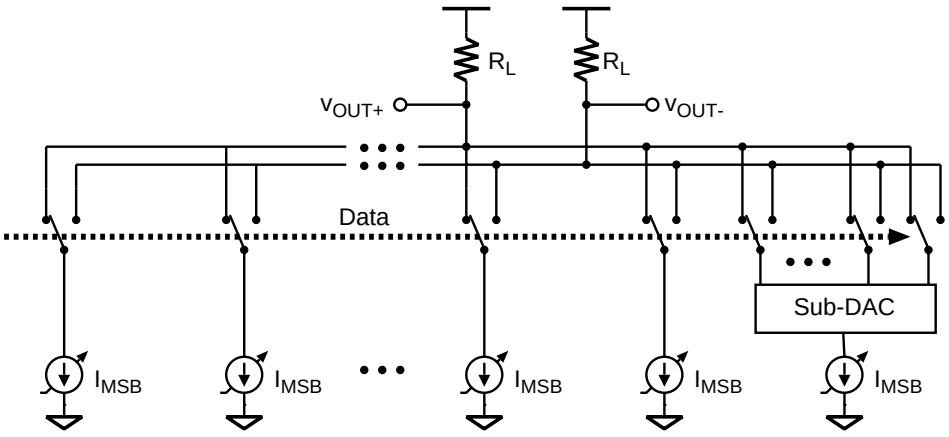


Figure 3.9: Basic trimmed current source DAC architecture.

Due to technology imperfections, the control voltage stored on C_{store}, will droop over time as the stored charge gradually leak from the capacitor. This necessitates some calibration frequency of the entire trimmable current source array. If continuous operation of the DAC is not required, calibration can be done during a special calibration phase where all current source are trimmed and henceforth put into operation. Continuous operation can be obtained if an extra trimmable current source is included for replacement of the current source being trimmed. The trimmed current source can be implemented

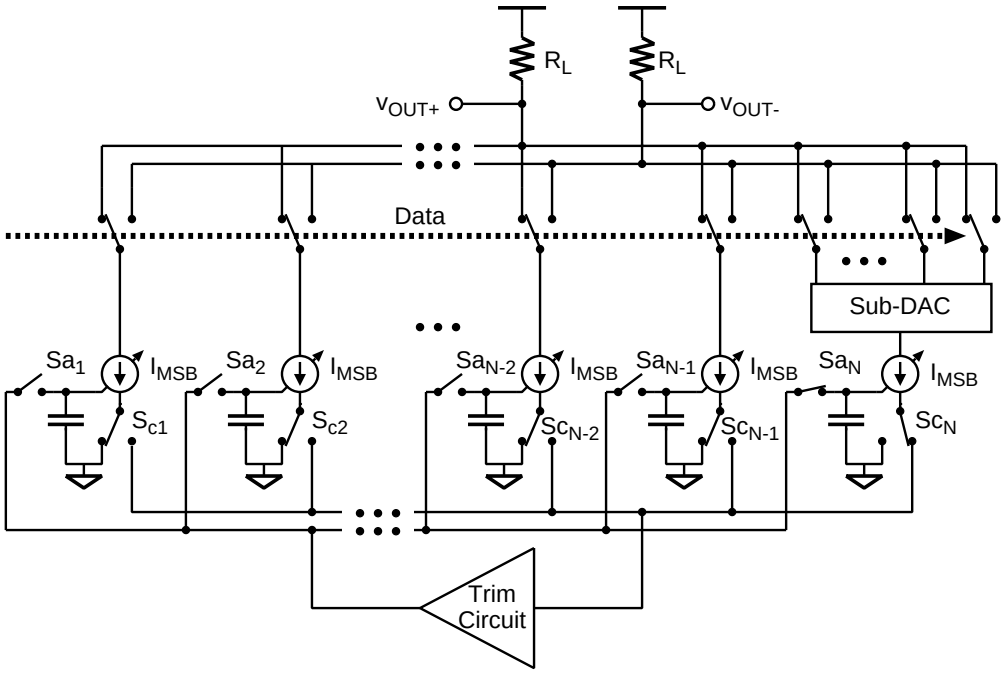


Figure 3.10: Basic background trimmed current source DAC architecture.

as a single MOST, but is normally implemented using two MOSTs as only a fraction of the total source current needs to be trimmed to match the reference current. The self-calibration technique was introduced in [22], where it was used for implementing a 16-bit DAC for audio range signals. It has since been the basis of numerous DAC implementations, of which one of the more recent is a 14-bit 100 MS/s DAC in CMOS only is reported in [23], which furthermore uses a special output stage for dynamic linearity enhancement. Another example where a classical calibrated DAC has been used is given in [24]. Here the sub-DAC has been omitted and a 6-bit trimmed current steering DAC has instead been used in a $\Sigma\Delta$-modulator scheme to obtain 14-bits performance. The drawback of using the $\Sigma\Delta$-modulator configuration is that the signal band is limited due to the required oversampling. In the case of [24] the effective output update rate was 10 MS/s.

Some of the drawbacks of using the simple trimmed current source scheme include:

- Increased power consumption as an extra current source has to implemented for replacement of the current source currently being calibrated
- Possible decreased SFDR due to the fact that an extra current source has to be switched in and out of operation thus introducing spurs in the output spectrum at the calibration frequency of the system .

This brings us to a solution which does away with these difficulties. By using a two-terminal, i.e. floating current source for each of the trimmed sources, the calibration of the current sources can be done seamlessly in the background of the DAC operation. This solution does away with the extra needed current source and dynamic impact of the calibration to the DAC output is minimal, since the current

source terminal employed for the DAC output is not affected by the calibration. The fundamental architecture for this scheme is shown in fig. 3.10. Each of the individual MSB current sources have their tail currents connected to a switch Sc, and their trim terminal connected to Sa. The individual current source being trimmed is thus part of a closed loop consisting of the current source and a trim circuit when both these switches are closed as shown for the last MSB cell feeding the sub-DAC in the figure. When not being trimmed, the Sc switch is positioned to dump the tail current, and the control voltage is stored on the holding capacitor.

The trim circuit can be implemented in both the digital and the analog domain. In [25], a background self trimming DAC is proposed using floating current sources. In this implementation, a $\Sigma\Delta$-modulator is used for quantizing the measured output current of the MSB cell being calibrated and the MSB cell current calibration is thus done in the digital domain. The resulting analog trim-voltage is then obtained by using an auxiliary DAC dedicated to the calibration circuit. This method seems unnecessarily complex for the task given as it need to employ data converters at both the input and output of the trim circuit to fulfill it's function.

Instead a pure analog approach could be considered. If the sub-DAC has a resolution of N bits, then the loop gain of the trim loop for the MSB cells need to have a gain:

$$A_{\text{loop}} > 2^{N+1} \tag{3.9}$$

for a settling error: $\varepsilon_{\text{settle}} < 1/2\,\text{LSB}$. If we assume that the sub-DAC has a resolution of 10 bits, which is approximately the maximum realizable intrinsic matching resolution, the loop gain is required to fulfill: $A_{\text{loop}} > 66\,\text{dB}$, which does not pose a severe design challenge.

3.4 D/A Converter Non-idealities

At higher input signal frequencies, the performance of DACs usually deteriorate due to the impact of dynamic error effects, which are the subject of this section.

3.4.1 Glitch Induced Nonlinearity

As we saw in section 3.2.2, major carry glitches could occur if binary weighted current sources were employed in the DAC. The fully segmented DAC presented a way to remedy the poor glitch performance.

Improved harmonic performance can also be expected from this scheme. Consider the partial sine wave output of a thermometer DAC qualitatively shown in fig. 3.11. For each time-step n, the DAC output at the given sample-time can be written as a non-switching part, i.e. the bits which remain unchanged, and a switching part which is glitch-infected. In the figure the shaded area represents the non-switching part and the non-shaded area is the switching part. For a fully thermometer decoded DAC, the switching glitch can be assumed to be equal for all unit LSB current sources [26]. Let $g(t)$, denote the glitch-infected step for 1 LSB change. Thus for input word $B_{\text{in}}(n)$, at sample-time n, we

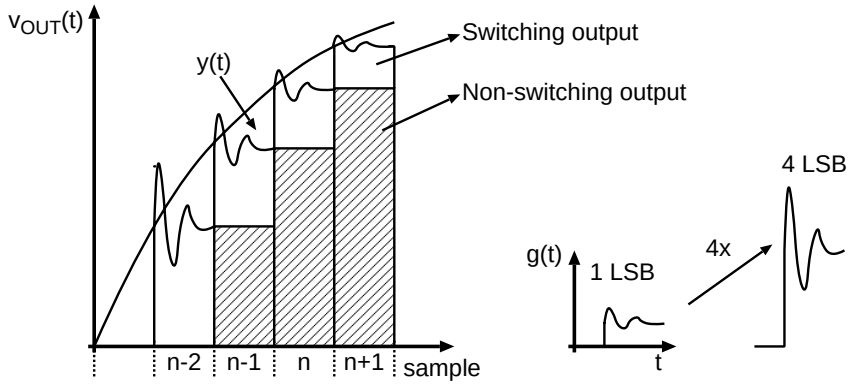


Figure 3.11: Output glitch behavior for a thermometer decoded DAC.

have the DAC output voltage:

$$v_{\mathrm{OUT}}(t) = g(t) \sum_{m=1}^{N} 2^m \left[b_m(n) - b_m(n-1) \right] + V_{\mathrm{LSB}} \sum_{m=1}^{N} b_m(n-1) 2^m \tag{3.10}$$

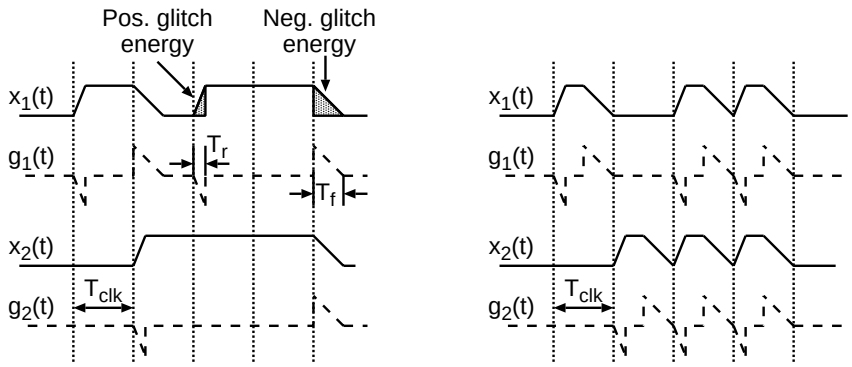
From this expression, we see that the DAC output is linearly dependent on $B_{\mathrm{in}}(n)$. Assuming that the glitchy step $g(t)$, is independent of B_{in}, the glitching will not produce any harmonic distortion [26]. This can also be explained as since the glitch is code-independent, multiple switching cells will only result in a multiplied glitch-signal which is not code-modulated and thus doesn't cause any nonlinearity.

However, as previously discussed, fully segmented DACs become impractical for high resolution DACs (number of bits > 10) and thus typically a binary decoded sub-DAC would be used for the lower LSBs. In the case of calibrated DACs, where only the thermometer decoded MSBs are calibrated, the sub-DAC could itself be segmented such that the upper LSBs (ULSBs) are thermometer decoded and only leaving a few lower LSBs (LLSBs) to be binary decoded. This presents a trade-off between circuit area and glitch performance which is application specific.

3.4.2 Code-Dependent Switch Feedthrough and Charge Injection

The individual switches steering the currents of the DAC are commonly implemented as CMOS transmission gates or simply as a single MOST.

Unless a fully segmented design is used, the individual current cells will carry a different amount of current. If only a standard size switch is realized, the amount of charge injected from the channel during switching will to a first order be the same for all switches as will the clock feedthrough due to parasitic overlap capacitance. This results in the charge injection and clock feedthrough to be code-dependent, i.e. causing nonlinearity. In order to match these non-idealities, all switches should be dimensioned such that their sizes are proportional to the current level switched. This will to a first order cancel the switch induced imperfections.



(a) Single source output, no RZ.

(b) Single source output, with RZ.

Figure 3.12: Intersymbol interference with and without using a RZ scheme.

3.4.3 Intersymbol Interference

Another effect of mismatches in the production of integrated circuits is non-equal rise- and fall-time for the switching waveforms. Consider the exaggerated waveforms shown in fig. 3.12(a). Here, the switching output for a single unit current source is shown for the input sequences: {0,1,0,1,1,0} and {0,0,1,1,1,0} for $x_1(t)$ and $x_2(t)$ respectively. Associated with each transition, we have a resulting error glitch resulting from the nonzero transition time. The unequal rise- and fall-time results in the glitch energy for the positive and the negative transitions to be unequal. The error signals corresponding to each of the input sequences are shown as $g_1(t)$ and $g_2(t)$ respectively.

In the case shown in fig. 3.12(a), both input sequences has the same mean value, i.e. 1/2. The error signals however, have unequal mean values, i.e. $m\{g_1(t)\} = 2m\{g_1(t)\}$, where $m\{\cdot\}$ is the mean value operator. This error effect is usually termed intersymbol interference. As the error signals are in fact code-dependent, though the code inputs may have the same mean value, nonlinearity will arise from this effect.

A countermeasure for this malady is to include a return-to-zero (RZ) scheme, which effectively clamps the output to zero for some fraction of the clock period T_{clk}. In fig. 3.12(b), the RZ scheme is utilized. Since we in this case always have the same number of positive and negative transitions for equal average input sequences, the resulting glitch error signals are will have the same average value as well as canceling the intersymbol interference. The price for including this scheme, is that some fraction of the output signal energy, proportional to the RZ nulling phase, will be lost and can thus deteriorate the resulting SNR.

3.5 Summary

In this chapter the basic functionality of the current steering DAC was presented. This architecture combines power efficiency, scalability with the ability to obtain very high update rates. Two possible implementations, binary and thermometer, were covered from which different pros and cons arise. In the case of the binary architecture, high glitch sensitivity was seen, whereas for the thermometer DAC

high complexity arises often leading to a design compromise in DACs, where segmentation is used such that the upper part of the DAC is thermometer decoded and the lower LSBs are binary decoded. The main challenge for this architecture is to obtain high resolution, for which numerous methods exist. We reviewed a DAC design principle based on knowledge of the statistical errors present on-die, from which different switching schemes for the individual current sources could be employed to combat the mismatch errors. Another approach covered was the calibrated DAC principle where the upper MSBs exceeding 9-10 bits, are subject to calibration against a reference to achieve the needed resolution.

Some of the major DAC non-idealities were covered and their impact on design decisions was discussed. Different auxiliary techniques employing extra circuitry were introduced to counter some of the covered non-idealities.

CHAPTER 4

OVERSAMPLED A/D CONVERSION

The A/D conversion of a signal implies sampling of the analog input signal, where the Nyquist frequency is the minimum sampling frequency to avoid aliasing [13]. Oversampling is defined by sampling the analog input signal at a higher rate than the Nyquist frequency. For a signal bandwidth f_{b}, the oversampling ratio (OSR) is defined by: $\mathrm{OSR} = \frac{f_{\mathrm{s}}}{2f_{\mathrm{b}}}$.

Over the last couple of decades, a class of A/D converters employing oversampling, known as $\Sigma\Delta$-modulators, have become increasingly popular due to their relaxed requirements for analog error sources [27]. The theory of these ADCs is the subject of this chapter. Besides from oversampling, these ADCs also employ noise filtering of the quantization noise, effectively shifting the majority of the quantization noise out of the signal bandwidth. By digital post-filtering of the out-of-band frequency range, a high SQNR can be obtained, even for a very low number of bits in the quantizer.

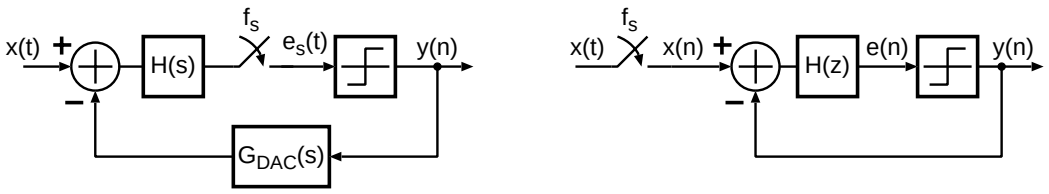
A drawback of oversampling is that the signal band is limited to some fraction of the sampling frequency. This drawback can be countered by increasing the number bits in the quantizer, hence lessening the quantization noise, and enabling a lower sampling frequency for the same SQNR performance. In this thesis however, only single-bit $\Sigma\Delta$-modulators will be considered.

The stability of especially higher order $\Sigma\Delta$-modulators has been under scrutiny since the emergence of this type of ADC [28, 29]. To counter the potential instability of modulators of order higher than two, cascaded architectures such as MASH [27] have been proposed. The MASH architecture employs low order ($\mathrm{N} \leq 2$) $\Sigma\Delta$-modulators in cascade to achieve higher-order performance, and thus avoids the stability issues of higher order modulators. We will in this chapter present a method for evaluating the stability of higher order modulators, and use this for estimating the ADC performance. The modeling framework used, was developed and presented in [30].

The metrics described in chapter 3, for specifying the performance of D/A converters, are also valid in the case of ADCs. The description of $\Sigma\Delta$-modulators however, rely heavily on signal processing concepts, as their performance is inherently linked to their frequency domain behavior. For this reason, much of the presented material in this chapter will rely on signal processing theory.

4.1 The $\Sigma\Delta$-Modulator

$\Sigma\Delta$-modulators are besides from oversampling, characterized by a loopfilter which can be either continuous-time or discrete-time as illustrated in fig. 4.1. In the continuous-time case, the sampling in the A/D conversion process takes places deep in the modulator loop, just prior to the quantizer. Whereas



(a) Continuous-time loopfilter

(b) Discrete-time loopfilter

Figure 4.1: Continuous- and discrete-time loopfilters in $\Sigma\Delta$-modulators.

the input signal for the discrete-time loopfilter, has already been sampled prior to entering the $\Sigma\Delta$-modulator loop. In fig. 4.1, a feedback D/A converter (DAC) with transfer function $G_{\mathrm{DAC}}(s)$ is shown. This feedback DAC is strictly speaking also present in the discrete-time loop, it is however as we will see in section 4.6, an important part of the overall loopfilter design in the continuous-time case whereas it in the discrete modulator can be set equal to unity as it merely converts the quantized output to the appropriate analog level which is however still discrete-time.

4.1.1 Basic Functionality

A linear model of the $\Sigma\Delta$-modulator employing a discrete-time loopfilter is shown in fig. 4.2.

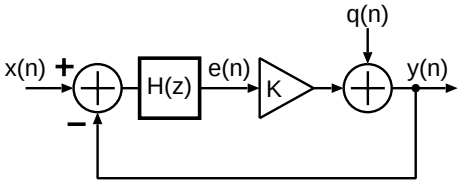


Figure 4.2: $\Sigma\Delta$-modulator linear model.

The quantizer is modeled by a gain block with an arbitrary gain K, and an additive noise source $q(n)$. The gain factor stems from the fact that the quantizer will truncate the signal to one of the quantization levels and will hence have an input determined gain delimited: $K \in \pm[1;\infty]$, where the sign is the input signal the sign. In the case of a 1-bit quantizer, the quantizer becomes equal to the signum function. The quantization noise power in a Nyquist rate data converter for a step size Δ, can be shown to be [12]:

$$P_{\mathrm{q}} = \frac{\Delta^2}{12} \tag{4.1}$$

An expression for the quantization noise power in oversampling converters is derived in section 4.3.2. For demonstrating the basic functionality of $\Sigma\Delta$-modulator ADCs, the expression given in eq. (4.1) suffices. A common assumption is that the quantization noise is white and hence has a constant frequency independent magnitude [31]. For a sampling frequency f_{s}, this noise power will consequently have a power spectral density (PSD):

$$S_{\mathrm{q}}(f) = \frac{\Delta^2}{12 f_{\mathrm{s}}} \tag{4.2}$$

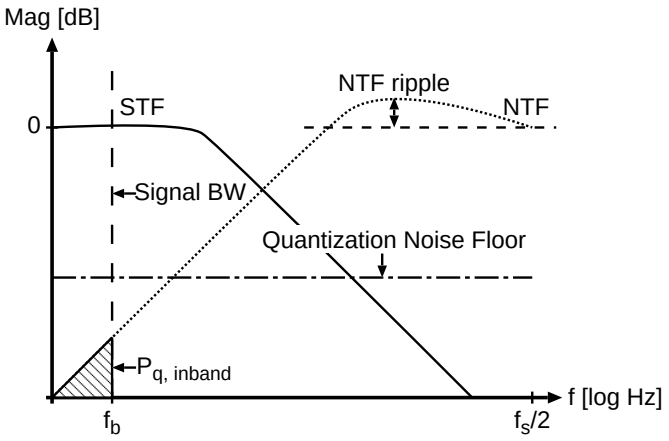


Figure 4.3: STF and NTF for a low-pass loopfilter.

From the schematic in fig. 4.2, we see that the output signal of the modulator can be written as a sum of two transfer functions. One for the signal input $x(n)$, henceforth denoted the signal transfer function (STF), and one for the quantization noise input, the noise transfer function (NTF):

$$Y(z) = STF(z)X(z) + NTF(z)Q(z) \tag{4.3}$$

Using the fully linear model allows us define the two transfer functions independently as the rules of superposition apply. From fig. 4.2, we have:

$$\begin{array}{rcccl} STF(z) & = & \frac{Y(z)}{X(z)} & = & \frac{KH(z)}{1+KH(z)} \\ NTF(z) & = & \frac{Y(z)}{Q(z)} & = & \frac{1}{1+KH(z)} \end{array} \tag{4.4}$$

The equivalent transfer functions for the continuous-time modulator employing the linear quantizer model are given by:

$$\begin{array}{rcccl} STF(s) & = & \frac{Y_{\mathrm{s}}(s)}{X(s)} & = & \frac{KH(s)}{1+KG_{\mathrm{DAC}}(s)H(s)} \\ NTF(s) & = & \frac{Y_{\mathrm{s}}(s)}{Q(s)} & = & \frac{1}{1+KG_{\mathrm{DAC}}(s)H(s)} \end{array} \tag{4.5}$$

If we assume a high DC-gain, low-pass function of the loopfilter $H(z)$ and $H(s)$ respectively, we see that for frequencies below the filter cut-off frequency f_{c}, we can approximate the signal transfer function with: $STF \approx 1$, and the noise transfer function: $NTF \approx 0$. If we assume a zero-order hold transfer function for the DAC present in the continuous-time loopfilter, we see that the frequency sinc-response does not alter the noise suppression at low frequencies, compared to the discrete-time case.

The two transfer functions are shown qualitatively in fig. 4.3 for a low-pass loopfilter function. As the signal is heavily oversampled, the signal bandwidth f_{b}, is much lower than the sampling frequency. The resulting low-pass STF is matched by a corresponding high-pass NTF. The quantization noise results in a noise floor which is shaped by the NTF, hence moving most of the noise out of the signal band. The resulting inband quantization noise, can be found by integrating the remaining noise over the

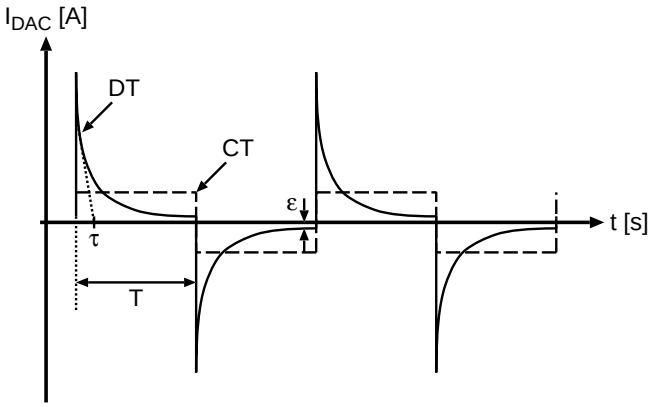


Figure 4.4: Switched-capacitor and switched-current DAC feedback current.

signal bandwidth:

$$P_{\mathrm{q,\,inband}} = \int_0^{f_\mathrm{b}} S_\mathrm{q}(f)\,|NTF(f)|^2\,df = \frac{\Delta^2}{12 f_\mathrm{s}} \int_0^{f_\mathrm{b}} |NTF(f)|^2\,df \tag{4.6}$$

Assuming a single-bit quantizer where only the levels $\pm\Delta$ exist, a maximum amplitude sinusoidal signal without overloading the quantizer has it's power given by: $P_\mathrm{s} = \Delta^2/2$. The resulting maximum SQNR is then:

$$\mathrm{SQNR}_\mathrm{max} = 10\log\left(\frac{P_\mathrm{s}}{P_{\mathrm{q,\,inband}}}\right) = 10\log\left(\frac{6 f_\mathrm{s}}{\int_0^{f_\mathrm{b}} |NTF(f)|^2\,df}\right) \tag{4.7}$$

The result given in eq. (4.7) tells us that a very high SQNR can be achieved even for a single-bit quantizer as long as we choose a high sampling rate coupled with a steep quantization noise transfer function. The steepness of the NTF can be chosen by increasing the loopfilter order.

4.1.2 Continuous-time vs. Discrete-time loopfilters

Several advantages can be gained from employing a continuous-time loopfilter over it's discrete-time equivalent. As the quantizer is located deep inside the continuous-time modulator loop, the preceding loopfilter will act as an anti-alias (AA) filter. In the discrete-time case, no such implicit filtering takes place prior to the sampling process, necessitating an explicit anti-aliasing filter to be implemented. However, for the high OSRs which are common for $\Sigma\Delta$-modulators, the requirements for slope of the AA-filter can be relaxed [32]. Nonetheless, if the signal being preprocessed has to retain a very high quality (e.g. CD quality $\sim$ 16 bits) the design of an analog filter with sufficient dynamic range may present a non-trivial task.

Whereas the AA-filtering issue presents a system-block level issue, the power consumption issue is related to the implementation of the individual integrators present in the loopfilter. A widely used approach to implement discrete-time $\Sigma\Delta$-modulators is the switched capacitor (SC) technique [33, 34], where an OTA is used for driving the integrating capacitor. This type of integrator relies on transferring

the signal charge in a short period as shown qualitatively in fig. 4.4 (curve DT). At the end of the clock period, the settling error ε, will remain and should be below 1/2 LSB for the specified accuracy. If we assume a first order model of the OTA, the settling time constant τ, is given by [35]:

$$\tau = \frac{1}{2\pi \mathrm{GBW}} \tag{4.8}$$

For a resolution of B bits, the number of time-constants n, necessary for sufficient settling of the integrator output, is found from:

$$n > (B+1)\ln(2) \tag{4.9}$$

Hence for a given resolution of the discrete-time $\Sigma\Delta$-modulator, a required GBW can be found dictating the necessary speed requirements, which determines the integrator power consumption. In continuous-time $\Sigma\Delta$-modulators, the feedback waveform can be arbitrarily chosen limited only by circuit complexity [36]. In fig. 4.4, the curve marked CT shows the simplest imaginable DAC feedback waveform; a NRZ curve. The amount of feedback charge is unchanged, but as the current is delivered at a constant amplitude, the speed requirements of the OTA delivering the current can be relaxed compared to those of the corresponding SC integrator [37]. In [38], it is reported that the OTA GBW requirement for a CT loopfilter can be lessened to 1/3 the GBW of the equivalent DT SC loopfilter. Here, OTA-C integrators are considered with the OTA operating in class A in both cases.

Another issue for SC implementations is switch resistance. As technology scales downward, so does the maximum supply voltage available. Lower supply voltages implies higher switch resistances as the maximum achievable gate overdrive diminishes, cf. A.2.2. For the common NMOS-PMOS transmission gate, this might result in an input voltage region where the transmission gate does not conduct at all. A common approach to resolve this issue, is to employ clock voltage doublers for the critical switches [34, 39]. This method to alleviate the effects of reduced supply voltage, will however not be viable for deep-submicron technologies due to the limit posed by gate dielectric reliability [40].

In the recent years, a new variant of the switched capacitor technique has emerged, namely the switched-opamp technique. Rather than relying on switch conductance, the output-stage of the OTA is itself switched on-off, thereby avoiding the use of transmission gates at voltage critical nodes. Using class AB OTAs, very low power consumption has been reported ($\leq 40\,\mu W$) [41], while retaining high dynamic range.

4.2 Continuous- and Discrete-time Loopfilter Equivalence

From fig. 4.1, it is seen that the $\Sigma\Delta$-modulator based on a continuous-time loopfilter, actually constitutes a mixed discrete- and continuous-time system. The CT $\Sigma\Delta$-modulator loopfilter can be said to be equivalent to the DT counterpart, if the loopfilters at the sampling instant produce the same input to the quantizer, for the same modulator input $x(t)$ [42]:

$$e_{\mathrm{s}}(t)|_{\mathrm{t=nT}} = e(n) \tag{4.10}$$

where T is the sampling clock period. This condition is fulfilled if the impulse response of the loopfilters is the same at the sample instants:

$$\mathcal{Z}^{-1}\left\{H(z)\right\} = \mathcal{L}^{-1}\left\{H(s)G_{\mathrm{DAC}}(s)\right\}\big|_{\mathrm{t=nT}} \tag{4.11}$$

where $\mathcal{Z}^{-1}\left\{\cdot\right\}$ is the inverse Z transform and $\mathcal{L}^{-1}\left\{\cdot\right\}$ is the inverse Laplace transform. With the above conditions satisfied, the signal sampled by the quantizer will be identical in the CT and DT loopfilters. Hence DT simulations will capture the CT loopfilter time response at the sample-time instants.

Mathematical theory allows us to obtain equivalent DT and CT models of the same system. With regards to simulation time, this is a great improvement as DT simulations have a significantly lower computational cost.

The necessary mathematical tools for performing the mapping between the two domains is derived in appendix B. In the following, this equivalence will be utilized and thus results from discrete-time filters will be presented since these can be directly related to the corresponding continuous-time filter.

4.3 Quantizer Modeling

In the following we will estimate the parameters which characterize the 1-bit quantizer. The linearized gain factor will be estimated, the quantizer output variance and noise variance are deducted and a noise amplification factor is introduced. The analysis will be performed for a digital $\Sigma\Delta$-modulator configuration, since the theoretical framework presented is most easily explained in the discrete-time domain as opposed to mixed discrete- continuous-time analysis. The results obtained however, are also applicable to a CT $\Sigma\Delta$-modulator configuration, since we have from section 4.2 that equivalence between the DT- and the CT $\Sigma\Delta$-modulator can be obtained.

4.3.1 The Linear Gain Factor

The quantizer constitutes a highly nonlinear element in the modulator. In fig. 4.2 this non-linearity is modeled by a gain factor and an additive noise source. In [43] two linearized gain factors are proposed under the assumption of a DC-input to the modulator. The two gains correspond to the gain of the DC-component K_{DC}, and the gain for the AC-component K_{AC}, of the quantizer input signal respectively. The gain factors are assigned in such a manner as to minimize the mean square error. Consider the quantizer input signal of fig. 4.2 split into the two components:

$$e(n) = e_{\mathrm{DC}}(n) + e_{\mathrm{AC}}(n) \tag{4.12}$$

where $e_{\mathrm{AC}}(n)$ is assumed to be a zero mean stochastic signal. Let $H(n)$ denote the nonlinear transfer function of the quantizer. The output of the quantizer can thus be written:

$$y(n) = H\left[e(n)\right] = H\left[e_{\mathrm{DC}}(n) + e_{\mathrm{AC}}(n)\right] \tag{4.13}$$

For the modelling error $\varepsilon(n)$, to be minimized in the mean square sense, we thus have that the error variance σ_ε^2, should be minimized:

$$\sigma_\varepsilon^2 = E\left\{\varepsilon^2(n)\right\} = E\left\{\left[y(n) - K_{\mathrm{DC}}e_{\mathrm{DC}}(n) - K_{\mathrm{AC}}e_{\mathrm{AC}}(n)\right]^2\right\} \tag{4.14}$$

Taking the partial derivatives with respect to each of the gain factors of eq. (4.14) and setting each of those equal to zero we obtain:

$$\frac{\partial\sigma_\varepsilon^2}{\partial K_{\mathrm{AC}}} = 2K_{\mathrm{AC}}E\left\{e_{\mathrm{AC}}^2(n)\right\} - 2E\left\{y(n)e_{\mathrm{AC}}(n)\right\} = 0 \tag{4.15}$$

$$\frac{\partial\sigma_\varepsilon^2}{\partial K_{\mathrm{DC}}} = 2K_{\mathrm{DC}}E\left\{e_{\mathrm{DC}}^2(n)\right\} - 2E\left\{y(n)e_{\mathrm{DC}}(n)\right\} = 0 \tag{4.16}$$

Solving eqs. (4.15,4.16) for the resulting gain factors we obtain the following definitions:

$$K_{\mathrm{AC}} \triangleq \frac{E\left\{y(n)e_{\mathrm{AC}}(n)\right\}}{\sigma_{\mathrm{e}}^2} \tag{4.17}$$

$$K_{\mathrm{DC}} \triangleq \frac{m_{\mathrm{y}}}{m_{\mathrm{e}}} \tag{4.18}$$

where $m_{\mathrm{e}} = E\left\{e_{\mathrm{DC}}(n)\right\}$, $m_{\mathrm{y}} = E\left\{y(n)\right\}$ and $\sigma_{\mathrm{e}}^2 = E\left\{e_{\mathrm{AC}}^2(n)\right\}$. The last identity holds as we are considering an AC-signal only system for the derivation of K_{AC}. Consider fig. 4.2 for mean values only, expressing the mean quantizer input m_{e}, in terms of the mean modulator input m_{x}, gives us:

$$m_{\mathrm{e}} = \frac{H(z)}{1 + H(z)K_{\mathrm{DC}}}m_{\mathrm{x}} \tag{4.19}$$

As we are considering mean values only, there is no contribution from the quantization noise which is modeled as zero mean. As the loopfilter has a very high DC-gain, ideally $\rightarrow \infty$, we can simplify the expression given in (4.19):

$$m_{\mathrm{e}} \approx \frac{m_{\mathrm{x}}}{K_{\mathrm{DC}}} \tag{4.20}$$

Inserting eq. (4.20) in the derived definition of K_{DC} given in eq. (4.18), we find that $m_{\mathrm{y}} \approx m_{\mathrm{x}}$. This restates the result given in the introduction to the basic operation of $\Sigma\Delta$-modulators, that the high DC-gain of the loopfilter will ensure that the mean output value of the modulator will equal the mean value of the input.

As stated earlier, a prerequisite of the models is that only DC-input is applied at the modulator input. For the high OSRs typically employed for single-bit quantizer modulators, the input frequencies will typically be low enough to honor this constraint. This constraint furthermore allows us to conclude that all activity is due to quantization noise. Hence the AC gain factor applies only to quantization noise in the loop.

4.3.2 Quantizer Output Variance

In the following, only the AC-gain factor of the quantizer will be considered and will be denoted 'K'.

The negative feedback employed in the loopfilter will seek to minimize the amount of quantization

noise in the loop. As the linear model employed for the quantizer is derived under the constraint to minimize mean square error, as much as possible of the output variance is accounted for by the model. The AC output of the quantizer can thus be found as the amplified AC-component, and the added quantization noise:

$$\begin{array}{rcl} y(n) - m_{\mathrm{y}} & = & K\left[e(n) - m_{\mathrm{e}}\right] + q(n) \\ \Leftrightarrow \quad y(n) & = & m_{\mathrm{y}} + K\left[e(n) - m_{\mathrm{e}}\right] + q(n) \end{array} \tag{4.21}$$

Calculating the quantizer output variance from eq. (4.21) we have:

$$\sigma_y^2 = V\left\{y(n)\right\} = E\left\{y^2(n)\right\} - E^2\left\{y(n)\right\} \tag{4.22}$$

$$= K^2\sigma_{\mathrm{e}}^2 + \sigma_{\mathrm{q}}^2 \tag{4.23}$$

where σ_{e}^2 and σ_{q}^2 are the quantizer input variance and the quantization noise variance respectively. For the derivation of eq. 4.23, it was assumed that all signal components are orthogonal and that the noise is zero mean. Alternatively, the output variance can be expressed:

$$\sigma_y^2 = 1 - m_{\mathrm{y}} \tag{4.24}$$

where the output power is equal to one as the only values the quantizer output can assume are ± 1. As we have seen earlier, the output mean value is given by the input mean: $m_{\mathrm{y}} \approx m_{\mathrm{x}}$. So the expression given in eq. (4.24) shows us that indeed the output variance must be caused by the quantization noise. Combining eq. (4.23), (4.24) and inserting the definition of the AC-gain factor derived in the previous section, we can obtain an expression for the resulting quantization noise:

$$\sigma_{\mathrm{q}}^2 = 1 - m_{\mathrm{y}} - \frac{E\left\{y(n)e(n)\right\}}{\sigma_{\mathrm{e}}^2} \tag{4.25}$$

Hence, the resulting quantization noise is dependent on the quantizer AC-gain and the mean modulator output. From the previous considerations, we have that $m_{\mathrm{y}} \approx m_{\mathrm{x}}$. In [43], the AC-gain factor has been evaluated assuming that the quantizer input stream has a Gaussian PDF, resulting in the following expression for the quantization noise variance:

$$\sigma_{\mathrm{q}}^2 = 1 - m_{\mathrm{x}} - \frac{2}{\pi}\exp\left[-2\left(\mathrm{erf}^{-1}\left(m_{\mathrm{x}}\right)\right)^2\right] \tag{4.26}$$

where erf^{-1} is the inverse error function. For a uniform PDF the noise variance can be shown to be [44]:

$$\sigma_{\mathrm{q}}^2 = 1 - m_{\mathrm{x}} - \frac{3}{4}\left(1 - m_{\mathrm{x}}^2\right)^2 \tag{4.27}$$

The derived expressions are seen to be dependent only on the mean modulator input. The resulting quantization noise variance is shown in fig. 4.5 for modulator inputs from zero to full scale. From the figure, it is seen that the quantization noise goes to zero as the modulator input approaches unity. This is due to the fact, that the only input levels a 1-bit quantizer can describe arbitrarily well are ± 1. Hence

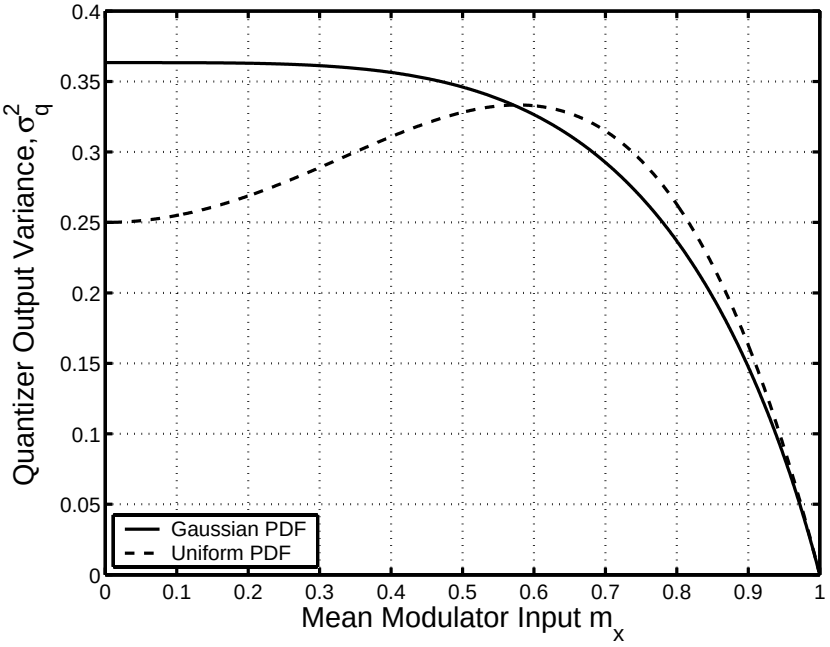


Figure 4.5: Quantization noise power versus mean modulator input.

for the shown quantizer input PDFs, the quantization noise power is restricted to values in the interval:

$$\sigma_{\mathrm{q}}^2 \in \left[0\,;\,1-\frac{2}{\pi}\right] \tag{4.28}$$

It is seen that this quantization noise description varies significantly from the Nyquist rate expression. Henceforth the above derived theory will be utilized.

4.3.3 Noise Amplification Factor

In the previous section, we saw that the quantizer output variance was due to quantization noise. An important parameter when estimating the stability of $\Sigma\Delta$-modulators is the so-called noise amplification factor A, proposed in [45]. The noise amplification factor is defined as the ratio between the quantizer's output variance, over the quantization noise variance, as a function of the linearized AC-gain factor:

$$A(K) \triangleq \frac{\sigma_{\mathrm{y}}^2}{\sigma_{\mathrm{q}}^2} \tag{4.29}$$

Since the transfer function between $q(n)$ and $y(n)$ is known, i.e. $NTF(f)$, the output variance could also be expressed as the filtered quantization noise:

$$\sigma_{\mathrm{y}}^2 = \sigma_{\mathrm{q}}^2 \int_{-f_s/2}^{f_s/2} |NTF\,(f)|^2\,df \tag{4.30}$$

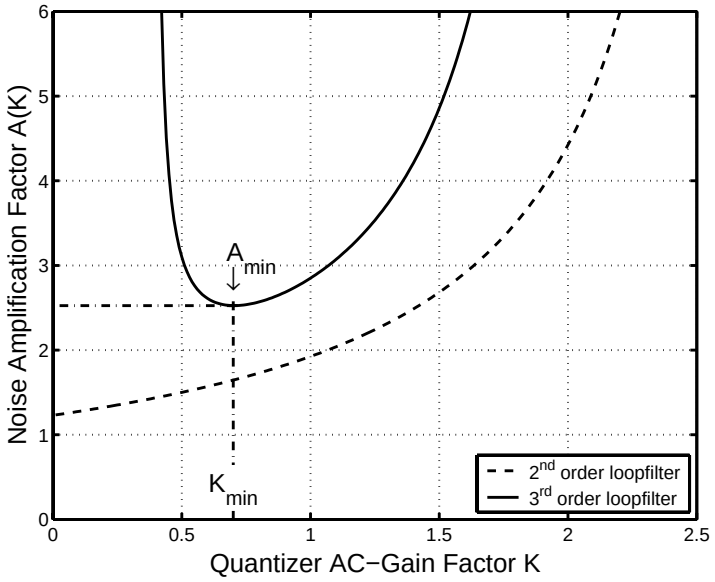


Figure 4.6: Example $A(K)$-curves obtained from a second and a third order loopfilter respectively.

Using the definition in eq. (4.29), the noise amplification factor can be expressed in terms of the noise transfer function:

$$A(K) = \int_{-f_s/2}^{f_s/2} |NTF(f)|^2 \, df \tag{4.31}$$

$$= \sum_{n=0}^{n=\infty} ntf^2(n) = \|ntf(n)\|_2^2 \tag{4.32}$$

To obtain 4.32, Parseval's therorem for discrete-time signals has been used. The $\|\cdot\|_2$ notation is the 2-norm.

For causality, $H(z)$ contains one sample delay. This is also the case for CT-loopfilter as the sampler is present in the loop. This ensures that the discrete-time impulse response of the noise transfer function has the property: $ntf(0) = 1$. This implies that:

$$A(K) \geq 1 \tag{4.33}$$

In fig. 4.6, two example $A(K)$ curves are depicted for a second order and a third order Butterworth high-pass NTF, where the filter cutoff frequency is set to 1/10 of the sampling frequency.

The main difference between the two curves, is that the second order curve is monotonically increasing, whereas the third order curve is $\cup$-convex and thus has a global minimum at $K \neq 0$. In the following section, we will see that this convexity is crucial for stability assessment in higher order $\Sigma\Delta$-modulators.

In [30] it is shown that three qualitatively different $A(K)$ curves exist. All higher order modulators ($N \geq 3$) are shown to have $\cup$-convex $A(K)$-curves. The two other curve types are both monotonically increasing and are distinguished by whether their global minimum is given by: $A(0) = 1$.

4.4 ΣΔ-Modulator Stability

It is a well known fact that ΣΔ-modulators of orders higher than one, become unstable for inputs exceeding some threshold value [46, 45, 47]. This threshold value will henceforth be denoted the maximum stable amplitude (MSA).

Due to the strongly nonlinear quantizer being present in the modulator loop, the normal methods relying on completely linear systems for predicting loopfilter stability are insufficient. In [30] an alternative method for predicting stability which employs the NTF impulse response has been developed.

Consider the ∪-convex curve shown for the third order loopfilter shown in fig. 4.6. As the quantizer gain K, is based on long term statistical averaging, it can be assumed that the modulator will be working in an equilibrium operating point. Due to the ∪-convexity of the $A(K)$ curve, any equilibrium operating point A_{eq}, can have two different K values assigned to it. Only the minimum noise amplification point $A_{\min}$, is seen to have a single gain $K_{\min}$, assigned to it. However, as we will see in the following, only one of the two K values is stable.

$K > K_{\min}$: The slope around equilibrium is positive. If K is increased slightly, the noise amplification is increased and more noise will result at the quantizer output, and hence in the modulator loop. As the noise amplification factor is defined by quantizer output noise over input noise, the increased loop noise will result in a reduction of K and therefore the system is forced back into equilibrium. If K is decreased slightly, the noise amplification is decreased and less noise will be present in the loop and hence result in an increase of K, which again forces the system into equilibrium.

$K < K_{\min}$: The slope around equilibrium is negative. If K is decreased slightly, the noise amplification is increased and more noise in the loop results. This further reduces K and hence equilibrium escapes. An increased K results in an increased noise amplification, which again raises K and equilibrium escapes.

So due to the small perturbations of the equilibrium, only the K-value relating to the positive slope of $A(K)$ yields a stable equilibrium. In the case of positive $A(K)$ slope, a negative feedback effect forced the equilibrium to remain, whereas a positive feedback effect caused the equilibrium to escape for a negative $A(K)$ slope. The other curve shown in fig. 4.6 is seen to be monotonically increasing, and any equilibrium point on this curve will thus always remain stable.

The above discussed $A(K)$ curves were produced by employing eq. (4.32), i.e. utilizing the NTF impulse response. An analytical expression, dependent only on the modulator mean input, can be obtained from the noise amplification factor definition coupled with the expressions given for the output variance in eq. (4.24), and the quantization noise variance given in eq. (4.26):

$$A(m_{\mathrm{x}}) = \frac{\sigma_{\mathrm{y}}^2}{\sigma_{\mathrm{q}}^2} = \frac{1 - m_{\mathrm{x}}^2}{1 - m_{\mathrm{x}} - \frac{2}{\pi} \exp\left[-2\left(\mathrm{erf}^{-1}\left(m_{\mathrm{x}}\right)\right)^2\right]} \tag{4.34}$$

where it has been used that $m_{\mathrm{x}} \approx m_{\mathrm{y}}$ and a Gaussian PDF for the quantizer input has been assumed. The resulting noise amplification curve is shown as the solid line in fig. 4.7. As was derived in section

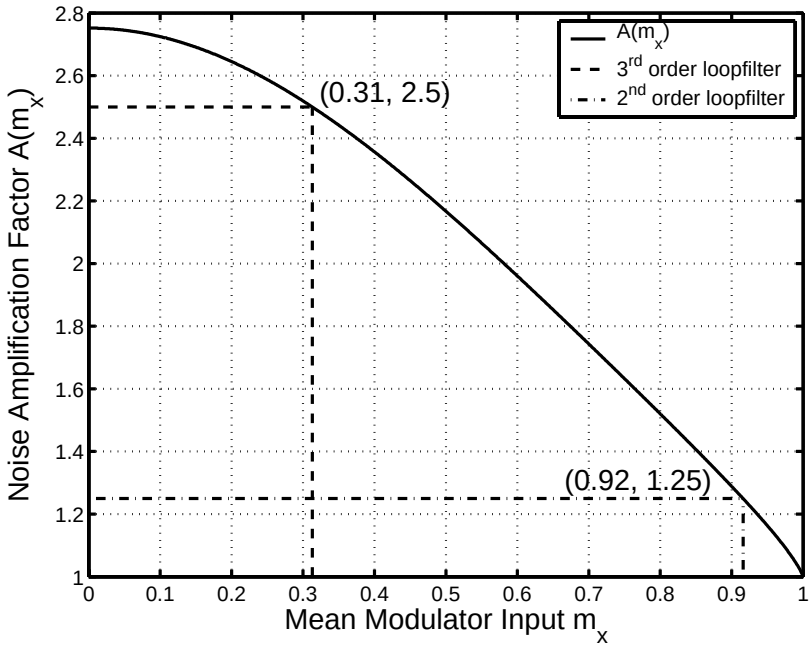


Figure 4.7: Noise amplification factor versus mean modulator input m_{x}.

4.3.3, the minimum value of $A(K)$ is 1. Solving (4.34) for the maximum value yields:

$$\max\{A(m_{\mathrm{x}})\} = A(0) = \frac{\pi}{\pi - 2} \approx 2.752 \tag{4.35}$$

So under the Gaussian assumption, the noise amplification factor lies in the interval:

$$A(m_{\mathrm{x}}) \in \left[1; \frac{\pi}{\pi - 2}\right] \tag{4.36}$$

The $A(m_{\mathrm{x}})$ curve shows that the noise amplification factor is inversely proportional to the modulator input, i.e. the higher modulator input the lower noise amplification factor. Consider once more the $\cup$-convex curve given in fig. 4.6. If we have a stable equilibrium operating point located on the positive slope of $A(K)$, increasing the modulator input m_{x}, will cause a decrease in $A(m_{\mathrm{x}})$, i.e. $A(K)$. So the operating point is pushed closer to the $A_{\min}$ point. If we keep increasing m_{x}, the $A_{\min}$ point will be passed and we will enter the part of the $A(K)$ slope which has a negative slope and instability will ensue. From these considerations, it follows that the MSA can be determined by solving for the m_{x}, for which the following equation is fulfilled:

$$A(K_{\min}) = A(m_{\mathrm{x}}) \tag{4.37}$$

For the two curves given in fig. 4.6 we have $A_{\min} \simeq 2.5$ and 1.25 for the third order- and the second order loopfilter respectively. For each of these $A_{\min}$ values, the corresponding MSA can be found as shown in fig. 4.7, giving us an approximate MSA of 0.31 and 0.92 for the example curves.

4.5 Performance Prediction

Assuming that analog noise sources are negligible, the resolution of a $\Sigma\Delta$- modulator is given by the SQNR. A simple formula for estimating the maximum SQNR for a digital interpolative $\Sigma\Delta$- modulator is derived in appendix C, and replicated below:

$$\mathrm{SQNR}_{\max} = 10\log\left(\frac{3}{2}\left(\mathrm{OSR}\right)^{2N+1}\cdot\frac{2N+1}{\pi^{2N}}\right) \tag{4.38}$$

This simple expression however, only gives a quick estimate of the performance which can be expected. The resulting SQNR is typically too optimistic as a number of presumptions have been made:

- The formula presumes that the quantization noise power is given by: $\sigma_{\mathrm{q}}^2 = \Delta/12$, as for Nyquist rate ADCs.
- Furthermore, it presumes that the modulator is stable for full-scale inputs, i.e. MSA=1.
- The quantizer AC-gain K, is presumed equal to 1.

Variants of this formula is often encountered in the literature for quick performance estimation. In the analysis done so far, we have seen that the above mentioned presumptions are *not* constants, but depends on the chosen modulator order, feedback filter cut-off frequency and sampling frequency. These parameters can be resolved, and hence the modulator performance predicted, if the working equilibrium can be resolved. These issues are the subject of this section.

4.5.1 Equilibrium Estimation

The overall performance of the modulator in terms of SQNR, is determined by the amount of inband noise and the MSA, i.e. the maximum input tone energy. In the previous sections, we have seen that both these parameters are dependent on the working equilibrium of the quantizer. Hence obtaining an estimate for the working equilibrium is important for prediction of overall modulator performance.

As was stated in section 4.4, the gain factor model assumes that the modulator will attempt to operate with the least amount of quantization noise in the loopfilter, as long as the operating point is located on the positive slope of the $A(K)$ curve. This is also intuitively clear, since the overall modulator constitutes a negative feedback system. At this point, it is convenient to introduce another transfer function: The error transfer function (ETF), between the quantization noise source and the quantizer input. From fig. 4.2 we have:

$$ETF(z) = \frac{E(z)}{Q(z)} = -\frac{H(z)}{1+KH(z)} \tag{4.39}$$

The error transfer function can also be expressed in terms of the NTF:

$$ETF(z) = \frac{NTF(z)-1}{K} \tag{4.40}$$

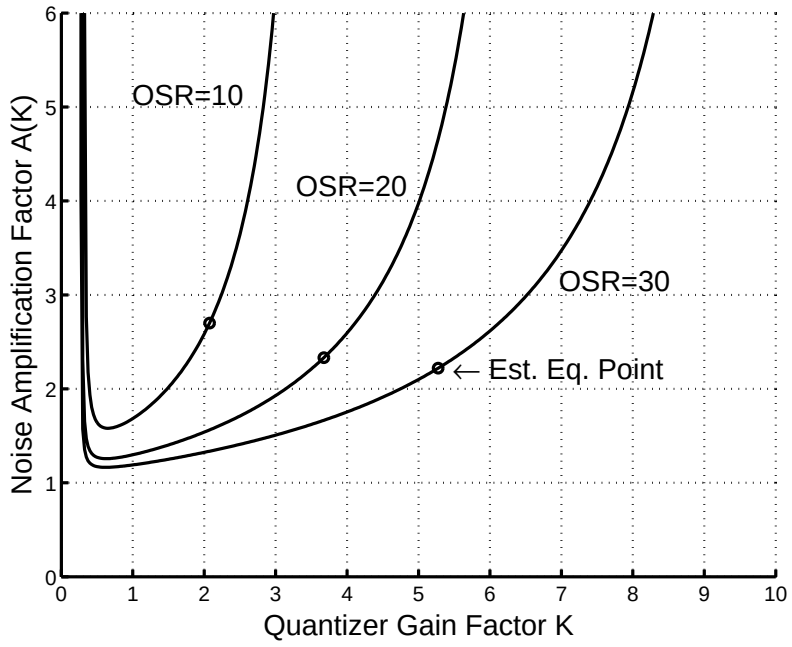


Figure 4.8: Noise amplification factor curves, from a third order loopfilter, for three different OSRs. Estimated equilibrium working points are indicated.

which in the time domain gives us:

$$etf(n) = \frac{ntf(n) - \delta(n)}{K} \tag{4.41}$$

The intrinsic loop noise minimization thus prescribes that a quantizer gain K, should be found such that $etf(n)$ is minimized in a least squares sense, i.e. that a minimum of quantization noise is present at the quantizer input. The squared two 2-norm of $etf(n)$ can be found to be:

$$\|etf(n)\|_2^2 = \frac{\|ntf(n)\|_2^2 - 1}{K^2} = \frac{A(K) - 1}{K^2} \tag{4.42}$$

which is seen to yield a result based on the quantizer gain dependent noise amplification factor $A(K)$. The resulting equilibrium gain K_{eq}, is thus found from the K that minimizes eq. (4.42). Fig. 4.8 shows three $A(K)$ curves for a third order modulator with the estimated equilibrium operating points shown. From the curves, it can be seen that for increasing OSR the modulator will become increasingly stable as the minimum point of the $A(K)$ curves move downward.

4.5.2 Equilibrium Performance

Once the equilibrium point has been established, the quantization noise power can be determined. From the definition of the noise amplification factor we have $A(K) = (1 - m_{\mathrm{x}}) / \sigma_{\mathrm{q}}^2$. Hence for zero input at

an equilibrium quantizer gain K_{eq}, the total quantization noise power is:

$$\sigma_{\mathrm{q}}^2 = \frac{1}{A(K_{\mathrm{eq}})} \tag{4.43}$$

Once the equilibrium point has been established, the total inband noise power can be evaluated:

$$\sigma_{\mathrm{b}}^2 = \frac{1}{A(K_{\mathrm{eq}})f_{\mathrm{s}}} \int_0^{f_{\mathrm{b}}} \left| \frac{1}{1 + K_{\mathrm{eq}} H(e^{j2\pi f})} \right|^2 df \tag{4.44}$$

Once the MSA is determined, the resulting maximum SQNR for a sinusoidal input signal at the MSA level, is given by:

$$SQNR_{\max} = 10 \log \left(\frac{\mathrm{MSA}^2 A(K_{\mathrm{eq}}) f_{\mathrm{s}}}{2 \int_0^{f_{\mathrm{b}}} \left| 1/\left[1 + K_{\mathrm{eq}} H(e^{j2\pi f})\right] \right|^2 df} \right) \tag{4.45}$$

In fig. 4.9, the predicted MSA and SQNR for different OSRs, is shown for a third order $\Sigma\Delta$-modulator, versus the relative NTF cutoff frequency $f_{\mathrm{c}}/f_{\mathrm{b}}$, where f_{b} is the signal bandwidth. The NTF filter type is a Butterworth high-pass.

From the curves we see that the SQNR increases with both the OSR and the NTF cutoff frequency. As the OSR is increased, the MSA increases while the quantization noise floor is lowered hence giving the higher SQNR. By increasing the NTF cutoff frequency, we choose to filter the quantization noise in a more aggressive manner, i.e. we move more noise out of the baseband. The price paid is the adverse effect on the MSA; as the cutoff frequency increases, the MSA decreases as seen from the figure.

The abrupt drop-off of the SQNR curves occur as the MSA approaches zero. From the figure, it is seen that the maximum SQNR is obtained at a MSA of approximately 0.3. This rather low value of the MSA is due to the fact that the noise suppression grows "faster" for increasing f_{c}, than the corresponding MSA decreases.

4.6 Loopfilter Design

The loopfilter $H(z)$, of the modulator can be designed from the desired NTF. If the modulator is of the type which has a low-pass STF and a high-pass NTF, it was found in section 4.1.1, that the feedback filter should be a low-pass filter with a high DC-gain.

Once the DT loopfilter has been designed, the CT counterpart $H(s)$, can be established using the mathematical tools derived in appendix B.

4.6.1 Noise Transfer Function Prototype

In [45] a method for designing the $\Sigma\Delta$-modulator loopfilter from a NTF prototype is given. For a NTF prototype given by $NTF(z) = A(z)/B(z)$ and using eq. (4.4) we have:

$$NTF(z) = \frac{1}{1 + KH(z)} = \frac{A(z)}{B(z)} \tag{4.46}$$

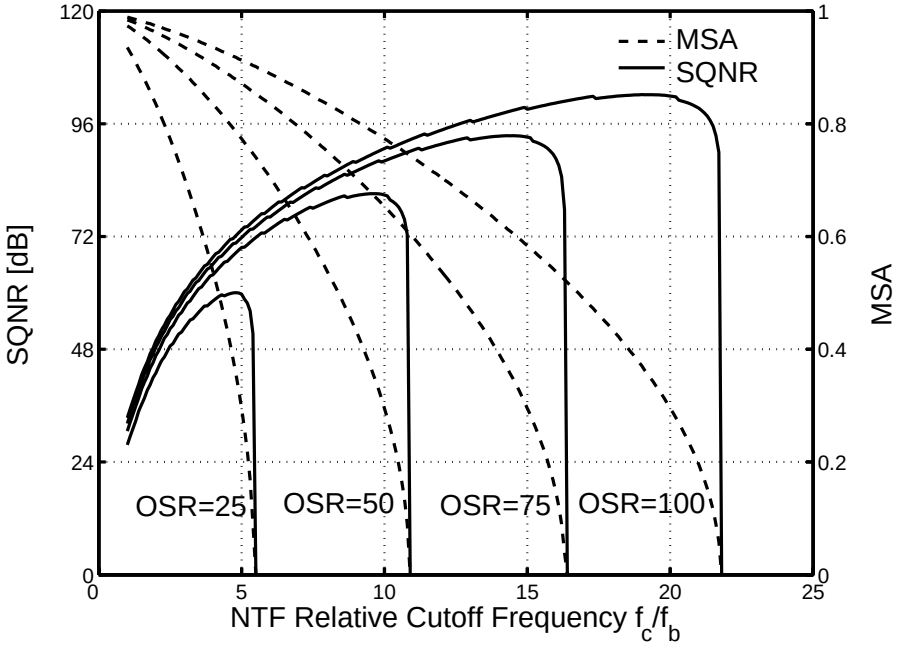


Figure 4.9: SQNR and MSA in a a 3^{rd} order $\Sigma\Delta$ ADC vs. f_c for increasing OSR.

Rearranging (4.46) to express $H(z)$ gives the following expression:

$$H(z) = \frac{1}{K} \cdot \frac{B(z) - A(z)}{A(z)} \tag{4.47}$$

Due to the single bit quantizer, which effectively implements a signum function, the loopfilter is invariant to scaling [30, 31] and the noise amplification factor K, can thus be set to unity:

$$H(z) = \frac{B(z) - A(z)}{A(z)} \tag{4.48}$$

What this simplification states, is that the modulator is defined only by it's poles and zeros and knowledge of the AC-gain factor is not necessary for design.

So, by designing the prototype NTF of the modulator, the $A(z)$ and $B(z)$ polynomials can be determined, and thus the feedback filter is defined. However, a scaling of these polynomials is necessary in order to ensure that the feedback filter $H(z)$ is not delay free. This is done by ensuring that $A(z)$ and $B(z)$ have the same highest order z-term, so that these will cancel each other in the numerator of $H(z)$, hence $ntf(0) = 1$ as we prescribed in section 4.3.3.

4.6.2 Coefficient Mapping

Consider the CT $\Sigma\Delta$-modulator architecture of order *N*, shown in fig. 4.10, this architecture is known as an interpolative modulator [48, 13] and is a typical choice for low-power $\Sigma\Delta$-modulator implementations, due to it's relaxed analog requirements [49]. This architecture is also the choice for the implemented $\Sigma\Delta$-modulator ADC presented in chapter 6.

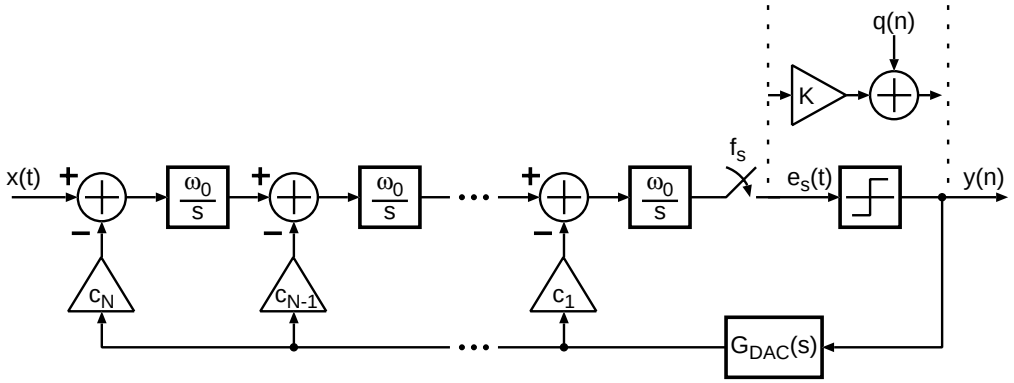


Figure 4.10: CT interpolative ΣΔ-modulator of order N.

By inspection of the block diagram we can establish the transfer functions:

$$STF(s) = \frac{K\left(\omega_0/s\right)^N}{1+K\left[c_{\mathrm{N}}\left(\omega_0/s\right)^N + c_{\mathrm{N}-1}\left(\omega_0/s\right)^{N-1} + \cdots + c_1\left(\omega_0/s\right)\right]G_{\mathrm{DAC}}(s)} \tag{4.49}$$

$$NTF(s) = \frac{1}{1+K\left[c_{\mathrm{N}}\left(\omega_0/s\right)^N + c_{\mathrm{N}-1}\left(\omega_0/s\right)^{N-1} + \cdots + c_1\left(\omega_0/s\right)\right]G_{\mathrm{DAC}}(s)} \tag{4.50}$$

From the NTF definition given in eq. (4.5) we can establish that for an interpolative ΣΔ-modulator of order N, the loopfilter is given by:

$$H(s) = \sum_{n=1}^{N} c_{\mathrm{n}}\left(\frac{\omega_0}{s}\right)^n \tag{4.51}$$

For a zero order hold feedback DAC, the z-transform of the combined feedback filter: $\widehat{H(s)} = H(s)\cdot G_{\mathrm{DAC}}(s)$, can be established by employing the mathematical tools derived in appendix B. Taking the z-transform of each order of the combined feedback filter $\widehat{H(s)}$, thus results in an expression of the given form:

$$\widehat{H(z)} = \mathcal{Z}\left\{\widehat{H(s)}\right\} = \sum_{n=1}^{N} \frac{c_{\mathrm{n}}\left(\omega_0 T\right)^n}{n!}\frac{B_{\mathrm{n}}(z)}{\left(1-z^{-1}\right)^n} \tag{4.52}$$

where $B_{\mathrm{n}}(z)$ is a polynomial in z of order *n*. By comparing the resulting coefficients for each order of z in the z-transformed interpolative feedback filter $\widehat{H(z)}$, with the coefficients resulting from the digital feedback filter derived from the NTF prototype $H(z)$, *N* linear equations result. These can be resolved using standard methods. Let d_{n} denote the coefficient for z of order *n*, in $H(z)$. The resulting mapping functions for obtaining the CT ΣΔ-modulator feedback coefficients from the digital NTF prototype for modulators of orders two through four is given in the table 4.1.

4.7 CT ΣΔ-Modulator Non-Idealities

Analog mismatches in the implementation of ΣΔ-modulators will cause the performance to deviate. Some of these mismatches can be viewed as errors from a system viewpoint. The subject of this section

N	Coefficient mapping functions
2	$c_1 = \frac{1}{2}\left(\frac{f_s}{\omega_0}\right)(d_2 - d_1)$, $c_2 = \left(\frac{f_s}{\omega_0}\right)^2 \sum_{n=1}^{2} d_n$
3	$c_1 = \frac{1}{6}\left(\frac{f_s}{\omega_0}\right)(2d_1 - d_2 + 2d_3)$, $c_2 = \left(\frac{f_s}{\omega_0}\right)^2(-d_1 + d_3)$, $c_3 = \left(\frac{f_s}{\omega_0}\right)^2 \sum_{n=1}^{3} d_n$
4	$c_1 = \frac{1}{12}\left(\frac{f_s}{\omega_0}\right)(-4d_1 + d_2 - d_3 + 4d_4)$, $c_2 = \frac{1}{12}\left(\frac{f_s}{\omega_0}\right)^2(11d_1 - d_2 - d_3 + 11d_4)$ $c_3 = \frac{1}{2}\left(\frac{f_s}{\omega_0}\right)^3(-3d_1 - d_2 + d_3 + 3d_4)$, $c_4 = \left(\frac{f_s}{\omega_0}\right)^4 \sum_{n=1}^{4} d_n$

Table 4.1: Coefficient mapping functions from discrete-time $\Sigma\Delta$-modulators to continuous-time.

is to discuss some of these errors and their performance impact.

4.7.1 Clock Jitter

In a sampling process, the exact sampling instant will be subject to stochastic perturbations due to clock generator imperfections. Let $\Delta t(n)$ denote the sampling instant time offset for sample n. The *n*th sampling instant $s(n)$, is then given by:

$$s(n) = nT + \Delta t(n) \tag{4.53}$$

where T is the sample clock period. This sampling instant uncertainty is termed clock jitter. For a NRZ, DAC feedback waveform, the jitter will affect the amount of charge which is fed back, as shown in fig. 4.11, for a current feedback waveform.

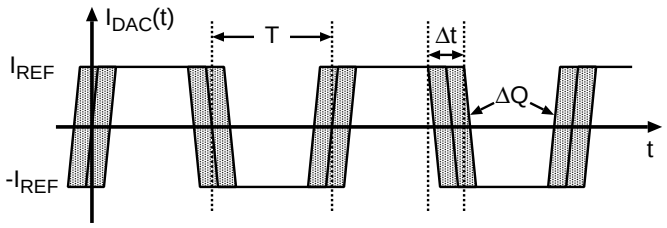


Figure 4.11: Clock jitter effect for a NRZ current feedback DAC.

For a modulator output sequence $y(n)$, the jitter induced error can be modeled as an error sequence given by [30]:

$$e_{\mathrm{NRZ}}(n) = [y(n) - y(n-1)]\,\frac{\Delta t}{T} \tag{4.54}$$

If we assume the jitter to have a white noise characteristic, i.e. wideband and uncorrelated with the output bit-stream difference sequence, the error sequence variance can be found:

$$\sigma_{\mathrm{e}}^2 = \sigma_{\Delta \mathrm{y}}^2 \frac{\sigma_j^2}{T^2} \tag{4.55}$$

where $\sigma^2_{\Delta \mathrm{y}}$ is the variance of $y(n) - y(n-1)$ and σ_j^2 is the variance Δt. Consider a sinusoid input signal with magnitude A, applied to a modulator running at some oversampling ratio OSR. If we ignore all other noise sources, the jitter limited SNR can be found [30]:

$$\mathrm{SNR}_{\mathrm{NRZ}} = 10 \log \left(\frac{\mathrm{OSR} \cdot A^2}{2\sigma^2_{\Delta \mathrm{y}} \left(\sigma_j / T\right)^2} \right) \tag{4.56}$$

The modulator output difference variance $\sigma^2_{\Delta \mathrm{y}}$, can be estimated by simulation, where-after the acceptable upper limit of the clock jitter can be calculated. Very good correspondence between simulation and eq. (4.56) is demonstrated in [50], verifying the validity of the formula.

4.7.2 Intersymbol Interference

If the DAC output waveform has unequal rise- and falltimes, the ADC will experience intersymbol interference. I.e. as explained in section 3.4.3: The amount of charge present in two pulsestreams will differ, though the average value of the two pulsestreams is the same. When a pulsestream which suffers from memory effects is used as feedback from the DAC, harmonic distortion can arise as the feedback charge is signal dependent.

To determine the effect of intersymbol interference analytically, prediction of the ΣΔ ADC output symbol density would be necessary, which may prove an intractable task for higher order loopfilters. In [51], the output symbol densities are estimated for first- and second order loops.

A possible solution, as previously discussed, is to employ a RZ scheme for the feedback pulse. This has the adverse effect though, that the jitter sensitivity increases as the number of transitions in the feedback stream effectively doubles.

4.7.3 Integrator Leakage

As infinite gain in the integrators is not attainable, perfect noise suppression at DC can not be obtained. The finite gain in effect pushes the zeros of the NTF transfer function from DC toward higher frequency. The integrator transfer function (ITF) can be augmented by a parasitic zero to model this effect:

$$\mathrm{ITF}(s) = \frac{\omega_0}{s + \omega_{\mathrm{p}}} \tag{4.57}$$

An example of the effect on the NTF, is in fig. 4.12. Here a third order Butterworth NTF with a cutoff frequency of 100 kHz, is exposed to ITF zero frequencies from 1 Hz to 1 kHz, clearly showing the deteriorated noise suppression capability. As the effect of this non-ideality is that quantization noise will 'leak' into the signal band, this is termed integrator leakage. The DC-gain of the integrator is related to the zero frequency by:

$$A_{\mathrm{DC}} = \frac{f_0}{f_{\mathrm{p}}} \tag{4.58}$$

If we limit our discussion to Butterworth type NTFs, then for signal band frequencies, an *N*th order NTF can be approximated by a differentiator of order *N*:

$$\mathrm{NTF}_{\mathrm{ideal}}(s) = \left(\frac{s}{\omega_0}\right)^N,\ \omega \ll \omega_0 \tag{4.59}$$

In the following, the integrator unity-gain frequency f_0, will be related to the sampling frequency by a ratio α, such that $f_0 = \alpha f_\mathrm{s}$. The double-sided inband noise can then be found by integrating the NTF over the signal band, which can be shown to be:

$$\begin{aligned}\sigma^2_{\mathrm{b,ideal}} &= \frac{\sigma^2_\mathrm{q}}{f_\mathrm{s}}\int_{-\omega_\mathrm{b}}^{\omega_\mathrm{b}} \left|\mathrm{NTF}_{\mathrm{ideal}}(\omega)\right|^{2N} d\omega = \frac{\sigma^2_\mathrm{q}}{f_\mathrm{s}}\int_{-f_\mathrm{b}}^{f_\mathrm{b}} \left|\frac{f}{f_0}\right|^{2N} df \\ &= \frac{\sigma^2_\mathrm{q}}{(2N+1)\,\alpha^{2N}\mathrm{OSR}^{2N+1}}\end{aligned} \tag{4.60}$$

Similarly, an NTF approximation for the NTF employing the leaky integrator is given by:

$$\mathrm{NTF}_{\mathrm{leak}}(s) = \left(\frac{s+\omega_\mathrm{p}}{\omega_0}\right)^N,\ \omega \ll \omega_0 \tag{4.61}$$

Resolving the double-sided inband noise for the leaky NTF, we find:

$$\begin{aligned}\sigma^2_{\mathrm{b,leak}} &= \frac{\sigma^2_\mathrm{q}}{f_\mathrm{s}}\int_{-\omega_\mathrm{b}}^{\omega_\mathrm{b}} \left|\mathrm{NTF}_{\mathrm{leak}}(\omega)\right|^{2N} d\omega = \frac{\sigma^2_\mathrm{q}}{f_\mathrm{s}}\int_{-f_\mathrm{b}}^{f_\mathrm{b}} \left|\frac{f^2+f_\mathrm{p}^2}{f_0^2}\right|^{N} df \\ &= \frac{\sigma^2_\mathrm{q}}{\alpha^{2N}\mathrm{OSR}^{2N+1}}\left[\left(\frac{f_\mathrm{p}}{2f_\mathrm{b}}\right)^{2N} + \sum_{n=1}^{N}\frac{N\,(N-1)\ldots(N-n+1)}{(2n+1)\,n!}\left(\frac{f_\mathrm{p}}{2f_\mathrm{b}}\right)^{2(N-n)}\right]\end{aligned} \tag{4.62}$$

By dividing eq. (4.62) by eq. (4.60), an expression for the resulting excess quantization noise can be found:

$$\frac{\sigma^2_{\mathrm{b,leak}}}{\sigma^2_{\mathrm{b,ideal}}} = 1 + (2N+1)\left[\left(\frac{f_\mathrm{p}}{2f_\mathrm{b}}\right)^{2N} + \sum_{n=1}^{N-1}\frac{N\,(N-1)\ldots(N-n+1)}{(2n+1)\,n!}\left(\frac{f_\mathrm{p}}{2f_\mathrm{b}}\right)^{2(N-n)}\right] \tag{4.63}$$

For a given loopfilter order, the required DC gain of the integrator can be found for a desired maximum inband excess noise. As an example case, let's consider a third order loopfilter. Using eq. (4.63), we find the noise ratio:

$$\frac{\sigma^2_{\mathrm{b,leak}}}{\sigma^2_{\mathrm{b,ideal}}} = 1 + 7\left[\left(\frac{f_\mathrm{p}}{2f_\mathrm{b}}\right)^6 + \left(\frac{f_\mathrm{p}}{2f_\mathrm{b}}\right)^4 + \frac{3}{5}\left(\frac{f_\mathrm{p}}{2f_\mathrm{b}}\right)^2\right] \tag{4.64}$$

Eq. (4.64) is most easily solved by graphical or numerical methods. If we require that the excess noise should be less than 3 dB, the following requirement results:

$$\frac{f_\mathrm{p}}{2f_\mathrm{b}} \leq 0.36 \tag{4.65}$$

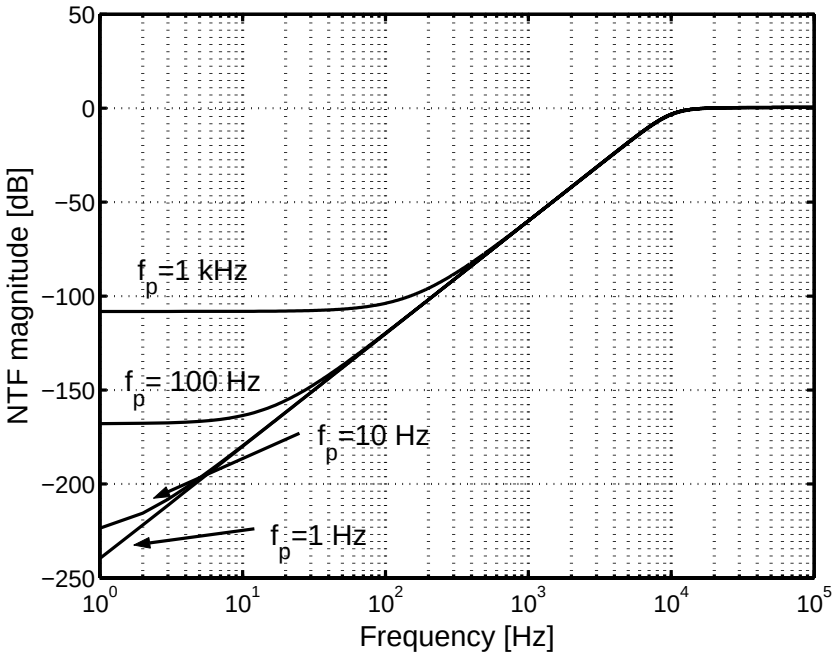


Figure 4.12: Third order Butterworth NTF with displaced integrator zeros.

This result can be related to the integrator DC gain, by isolating f_{p} in eq. (4.58) and inserting in the above expression. Thus we get:

$$A_{\mathrm{DC}} \geq \frac{\alpha}{0.36} \cdot \mathrm{OSR} \tag{4.66}$$

where we have used: $f_0 = \alpha f_{\mathrm{s}}$. Expressions similar to eq. (4.66) can be derived for any given order of Butterworth NTF and gives a simple rule-of-thumb for the integrator design. A similar expression is derived in [52] for discrete-time ΣΔ-modulators.

4.7.4 Integrator Gain Error

Due to analog mismatches in an implementation of a ΣΔ-modulator, the integrator unity-gain frequency ω_0, will be subject to perturbations. This is termed integrator gain error. If we assume that the errors will affect all integrators in a similar manner, all unity-gain frequencies will shift by the same amount $\Delta\omega$, such that the shifted integrator transfer function is given by:

$$\mathrm{ITF}(s) = \frac{\omega_0 + \Delta\omega}{s} \tag{4.67}$$

From eq. (4.50), we see that the integrator gain error will manifest itself as a shift in the ΣΔ-modulator NTF cutoff frequency. The effect of such a shift is most easily shown by an example. Consider the MSA and SQNR curves for a third order loopfilter at two different OSRs, given in fig. 4.13. For both curves, an initial cutoff frequency is chosen such that the resulting SQNR is close to maximum, but somewhat below the frequency where the SQNR slopes downward. On the figure, a cutoff frequency variation $\Delta\omega_1$ and $\Delta\omega_2$, of +/-20 % is shown from the initial cutoff frequency. It is seen that the impact

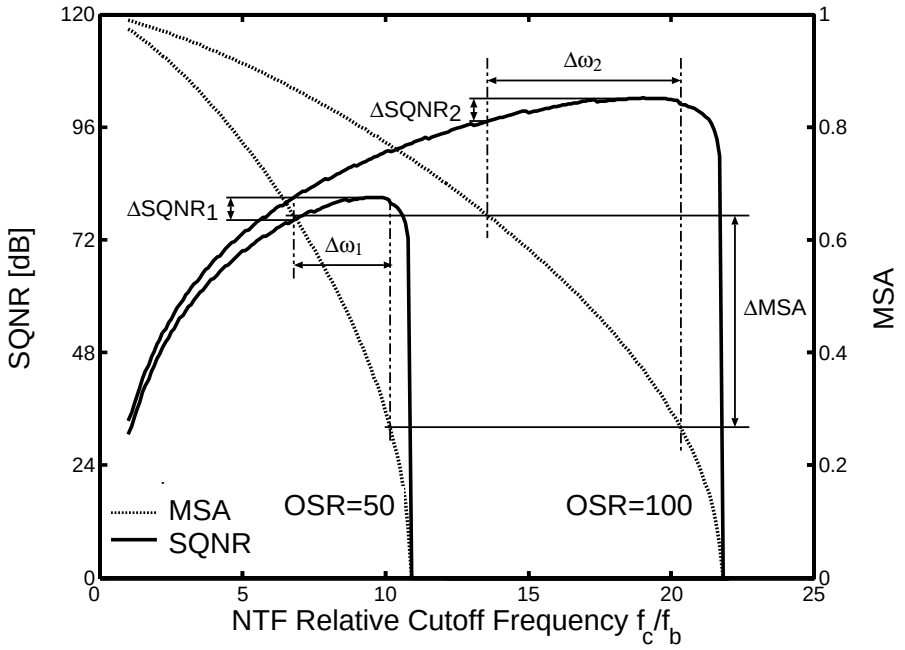


Figure 4.13: Gain error effect on SQNR and MSA for a third order loopfilter.

on the resulting SQNR for both curves is minimal. The SQNR variation is in both cases found to be less than 4 dB. However for the MSA, a significant variation $\Delta\mathrm{MSA} = 0.38$, is observed over this range.

Though the SQNR variation is seen to be minimal, the large MSA variation implies that the applicable signal energy may be much lower than the initial design was specified for. This further implies a higher power consumption in the analog implementation, in order retain sufficient dynamic range for the degraded input range. These considerations suggest that trimming capability of the NTF cutoff frequency should be considered for a given implementation, if large integrator gain errors are expected.

4.8 Summary

The subject of this chapter was the theoretical background necessary for the design of $\Sigma\Delta$-modulators. The basic functionality of the $\Sigma\Delta$-modulator was discussed and two types of loopfilters were presented: CT and DT loopfilters. Some drawbacks and advantages were discussed for the two filter types, and one advantage pointed out for CT loopfilters is the potentially lower power consumption compared to equivalent DT loopfilters.

Equivalence between the two loopfilter types for a given order was demonstrated. By mathematical theory, a set of mapping functions can be obtained which allows the transformation of a CT loopfilter to it's DT counterpart. This proves to be very useful as the wealth of knowledge on the design and modeling of DT systems can be used for CT systems as well. Moreover, the mapping allows fast simulation by employing the DT equivalent of a CT loopfilter.

The modeling framework for $\Sigma\Delta$-modulators is presented. The presented modeling demonstrates that oversampling converters vary significantly from Nyquist rate converters, and properties such as

quantization noise power and maximum input amplitude are closely linked to the choice of design parameters.

Some of the non-idealities which affect the performance of CT $\Sigma\Delta$-modulators are presented. High sensitivity toward clock jitter and intersymbol interference is to be expected for a NRZ feedback waveform. The intersymbol interference error can be canceled if an RZ waveform is used, however at the price of increased jitter sensitivity. By analysis, a rule of thumb was found for the minimum integrator gain required for a maximum performance degradation. Finally, integrator gain error was seen to heavily modulate the MSA, however without significantly degrading overall performance.

Part II

Applications

CHAPTER 5
INSTRUMENTATION AMPLIFIERS

Amplifiers used for specific measurement purposes are commonly termed instrumentation amplifiers. Instrumentation amplifiers are characterized by having a controlled gain whereas OTAs are usually designed to have gain which is above some minimum value. Standard OTAs can be used for constructing instrumentation amplifiers, where an example is the classical three OTA instrumentation amplifier [53]. This topology uses two OTAs for voltage buffers enabling high input impedance of the overall amplifier. The third OTA comprises a second stage for signal amplification with the gain controlled by resistor ratios.

Such a solution is potentially quite power hungry if the input signals have a small magnitude, which is typically the case for many sensors. The power consumption of an input amplifier is typically determined by the necessary noise suppression, yielding poor power efficiency for a three OTA design. In this chapter, two high gain CMOS instrumentation amplifiers suitable nerve cuff recorded signals are presented. To achieve the high gain, two amplifying stages are used in both designs. Each stage has a differential input enabling very high input impedance and the gain is controlled by loading.

As CMOS is the employed technology, several challenges arise. A CMOS differential input pair will typically have input-referred offset voltages in the range of several millivolts, which necessitates some scheme for offset compensation as saturation of the amplifier would otherwise result. CMOS circuits are often plagued by low frequency 1/f-noise. For low frequency sensor output, this noise type may well dominate overall noise performance of the amplifier.

The first amplifier presented solves the 1/f noise problem simply by MOST sizing as 1/f noise is inversely proportional to size. The second amplifier uses the chopper modulation technique for 1/f noise reduction.

The thermal noise level is regulated by the bias current. The intrinsic noise level of the cuff electrode is determined by the cuff resistance, which is reported to be in the range of $1\,\mathrm{k}\Omega$ to $20\,\mathrm{k}\Omega$ [54]. A compromise value of $5\,\mathrm{k}\Omega$ is chosen as a design parameter, indicating a thermal noise level of $9\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$. The input referred noise of the amplifiers should thus remain below this level.

As the nerve signal bandwidth is approximately 400 Hz - 4 kHz, the amplifier should have a bandwidth of at least 4 kHz. Adding some margin, the amplifiers were designed for a 10 kHz bandwidth.

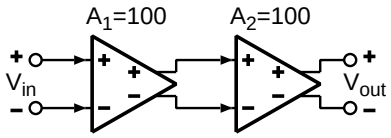


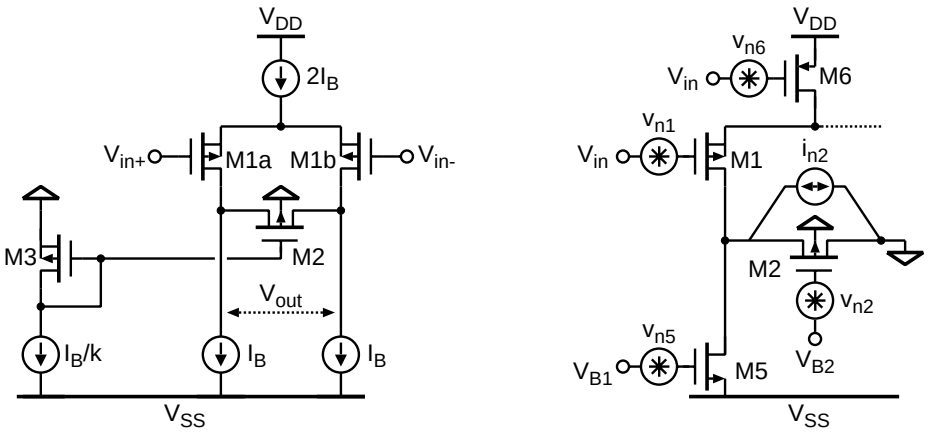
Figure 5.1: Amplifier block diagram.

5.1 Nerve Signal Amplifier V1

The desirable high gain needed for the instrumentation amplifier can be difficult to obtain in a controlled manner for a single stage. Instead a two-stage configuration for the overall amplifier is employed as shown in fig. 5.1, where each gain stage is configured to have a gain of 100, thus yielding an overall gain of 80 dB. The same gain circuit is employed for each gain stage in fig. 5.1.

5.1.1 Gain Stage Principle of Operation

In fig. 5.2(a), a schematic is shown of the gain stage employed in this instrumentation amplifier. The figure only shows the principle MOSTs for clarifying the gain stage operation. All current sources are implemented as high output impedance cascoded current sources. The input pair MOSTs M1a and M1b are biased in saturation whereas M2 is biased in the linear triode region controlling the gain as load. M3 is used for generating a correct gate bias for the M2 MOST.



(a) Principle of operation.

(b) Noise analysis half-circuit.

Figure 5.2: Gain stage principle circuit schematic and half circuit for noise analysis.

The gain of stage can be shown to be given by:

$$A_{\mathrm{V}} = -\frac{g_{\mathrm{m1}}}{2g_{\mathrm{ds2}}} \tag{5.1}$$

Depending on the inversion level of the channel, the transconductance and the channel conductance will vary, however if the inversion level of all MOSTs is the same, the gain will remain unaffected by the

level of inversion. Consider the resulting gain expressions for the weak- and strong inversion regions respectively:

Weak inversion: From section A.2.2, we have the respective transconductance and conductance for a MOST in weak inversion:

$$g_{\mathrm{m}} = 2K'(W/L)V_{\mathrm{T}}e^{\frac{V_{\mathrm{G}}-V_{\mathrm{th0}}-nV_{\mathrm{S}}}{V_{\mathrm{T}}}} \tag{5.2}$$

$$g_{\mathrm{ds}} = 2nK'(W/L)V_{\mathrm{T}}e^{\frac{V_{\mathrm{G}}-V_{\mathrm{th}}-nV_{\mathrm{S}}}{nV_{\mathrm{T}}}} \tag{5.3}$$

Inserting these small-signal expressions into eq. (5.1), we have for the resulting gain expression for weak inversion operation:

$$A_{\mathrm{V,WI}} = -\frac{2K'(W/L)_1 e^{\frac{V_{\mathrm{G1}}-V_{\mathrm{th1}}-nV_{\mathrm{S1}}}{V_{\mathrm{T}}}}}{4nK'(W/L)_2 e^{\frac{V_{\mathrm{G2}}-V_{\mathrm{th2}}-nV_{\mathrm{S2}}}{nV_{\mathrm{T}}}}} \tag{5.4}$$

Strong inversion: Equivalently, the small-signal parameters for strong inversion operation are:

$$g_{\mathrm{m}} = \frac{K'}{n}\frac{W}{L}(V_{\mathrm{G}}-V_{\mathrm{th0}}-nV_{\mathrm{S}}) \tag{5.5}$$

$$g_{\mathrm{ds}} = K'(W/L)(V_{\mathrm{G}}-V_{\mathrm{th0}}-nV_{\mathrm{S}}) \tag{5.6}$$

Which results in the following gain expression for the gain stage operating in strong inversion:

$$A_{\mathrm{V,SI}} = -\frac{\frac{K'_{\mathrm{P}}}{n}(W/L)_1(V_{\mathrm{G}}-V_{\mathrm{th0}}-nV_{\mathrm{S}})}{K'_{\mathrm{P}}(W/L)_2(V_{\mathrm{G}}-V_{\mathrm{th0}}-nV_{\mathrm{S}})} \tag{5.7}$$

If both the differential input pair and the load have the same gate to source voltage and no bulk effect has modified the threshold voltage, the above gain expressions can both be simplified to a single unified expression:

$$A_{\mathrm{V}} = A_{\mathrm{V,WI}} = A_{\mathrm{V,SI}} = -\frac{(W/L)_1}{2n(W/L)_2} \tag{5.8}$$

Except for the slope factor n, the above expression is solely dependent on transistor dimensions. The slope factor can be regarded as a technology constant [55], but can be shown to ultimately depend on the level of doping in the semiconductor. Consequently once n is determined, the gain is well defined for the proposed gain stage. In the proposed two-stage instrumentation amplifier, the gain is chosen to 100 in each stage.

In order to ensure that indeed both the load MOST and the input pair have the same inversion level, these are laid out in a common-centroid fashion where a unit-sized finger is used throughout in multiples for determining the gain scaling. The same unit finger is used for realizing M3 such that for a scaling of M3 dimensions: $(W/L)_3 = 1/k\,(W/L)_1$, accompanied by an equivalent bias current scaling: $I_{\mathrm{D3}} = 1/k \cdot I_{\mathrm{D1}}$ the same inversion level is assured for M3 and M1 and thus also for M1 and M2 as the gate of M3 is used for biasing M2. Thereby fulfilling the bias requirements necessary for eq. (5.8) being valid.

5.1.2 Noise Analysis

A half-circuit depicting the gain stage noise sources is shown in fig. 5.2(b). Here, the principal current source MOSTs and their corresponding noise sources are shown. Cascodes are not shown as their noise contribution is negligible [56]. For all MOSTs, except M2, both their thermal and 1/f-noise components are referred to the gate.

The noise source associated with the current source MOST M6, can be omitted from our calculations as the noise stemming from this source will couple to the outputs as a common-mode signal and will thus not be seen at the differential output.

The 1/f-noise component of the load MOST M2, is referred to the gate. As there is no drain bias current flowing in M2, the M2 transconductance g_{m2}, is equal to zero and the 1/f-noise will do little more than modulate the M2 channel conductance g_{ds2}, slightly. Hence the 1/f-noise for M2 can be ignored. The remaining noise sources give rise to the following expression for the total output noise power:

$$v_{\mathrm{no,tot}}^2(f) = v_{\mathrm{n1}}^2(f)\left(\frac{g_{\mathrm{m1}}}{g_{\mathrm{ds2}}}\right)^2 + v_{\mathrm{n5}}^2(f)\left(\frac{g_{\mathrm{m5}}}{g_{\mathrm{ds2}}}\right)^2 + \left(\frac{i_{\mathrm{n2}}(f)}{g_{\mathrm{ds2}}}\right)^2 \tag{5.9}$$

where it has been used that all noise sources are considered uncorrelated. For comparison with the amplifier input noise source, the total output noise is referred to the amplifier input, i.e. eq. (5.9) is divided by the M1 power gain $(g_{\mathrm{m1}}/g_{\mathrm{ds2}})^2$:

$$v_{\mathrm{ni,tot}}^2(f) = v_{\mathrm{n1}}^2(f) + v_{\mathrm{n5}}^2(f)\left(\frac{g_{\mathrm{m5}}}{g_{\mathrm{m1}}}\right)^2 + \left(\frac{i_{\mathrm{n2}}(f)}{g_{\mathrm{m1}}}\right)^2 \tag{5.10}$$

If we require that the main noise contributor should be M1, all other noise sources should be negligible compared the M1 inherent noise. Thus this requirement implies:

$$v_{\mathrm{n1}}^2(f) \gg v_{\mathrm{n5}}^2(f)\left(\frac{g_{\mathrm{m5}}}{g_{\mathrm{m1}}}\right)^2 + \left(\frac{i_{\mathrm{n2}}(f)}{g_{\mathrm{m1}}}\right)^2 \tag{5.11}$$

From section A.3, we have that the noise of a MOST in saturation can be written:

$$v_{\mathrm{n}}^2(f) \quad = \quad \underbrace{\frac{c4kT}{g_{\mathrm{m}}}}_{\mathrm{Thermal}} \quad + \quad \underbrace{\frac{K_{\mathrm{F}}'}{WLC_{\mathrm{ox}}f}}_{1/\mathrm{f}} \tag{5.12}$$

Consider first the thermal noise. A MOST biased in triode has it's thermal noise given by the resistive channel:

$$i_{\mathrm{n}}^2(f) = 4kTg_{\mathrm{ds}} \tag{5.13}$$

Combining eq. (5.11-5.13), we find that:

$$g_{\mathrm{m1}} \gg g_{\mathrm{m5}} + g_{\mathrm{ds2}}/c \tag{5.14}$$

The second term of 5.14 is easily fulfilled as the ratio between g_{m1} and g_{ds2} is recognized as the amplifier gain and since the inversion level dependent factor c, is approximately equal to one. The

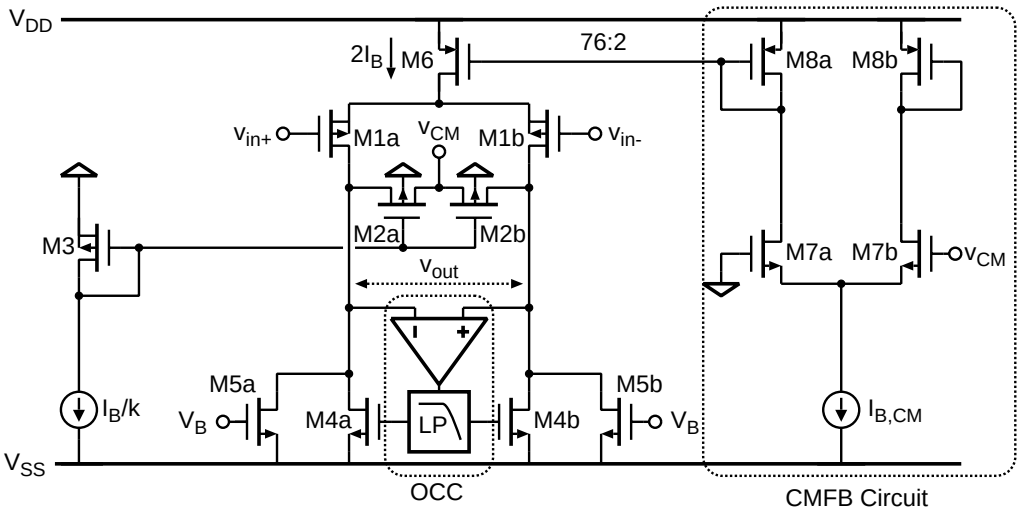


Figure 5.3: Amplifier first stage schematic.

first term suggests the relative dimensioning between M1 and M5 since they both carry the same drain current I_{D}: The $g_{\mathrm{m}}/I_{\mathrm{D}}$-ratio, i.e. the layout aspect ratio, of M1 should be maximized and minimized for M5.

For the 1/f-noise, inserting the 1/f-noise expression in the inequality eq. (5.11), gives us the following relationship between M1 and M5 areas:

$$\frac{(WL)_5}{(WL)_1} \gg \frac{K'_{\mathrm{F,N}}}{K'_{\mathrm{F,P}}} \left(\frac{g_{\mathrm{m5}}}{g_{\mathrm{m1}}}\right)^2 \tag{5.15}$$

This inequality is more difficult to fulfil though the M1 transconductance g_{m1}, is maximized relative to g_{m5}. The technology dependent 1/f-noise constant K'_{F}, it typically a factor ten lower in PMOSTs than in their NMOST counterparts. A maximized g_{m1}, also implies a large M1 area which then requires an equivalently large M5 area if eq. (5.15) is to be honored.

5.1.3 First Stage

A schematic of the implemented amplifier first stage is shown in fig. 5.3, though omitting cascode MOSTs for simplicity. Some additional circuitry is shown here which will be explained later. As the first stage receives the weak nerve signal directly from the cuff electrode, minimal noise is crucial for this stage. As the transconductance of a MOST is maximized in weak inversion, this operating mode is employed for the M1 input pair. For a given cuff electrode resistance R_{cuff}, we thus get the necessary drain current for the input pair to achieve a thermal noise floor below that of the cuff resistance:

$$I_{\mathrm{D1}} > \frac{2n^2 V_{\mathrm{T}}}{R_{\mathrm{cuff}}} \tag{5.16}$$

In our case, a total bias current for the stage input pair of $76\,\mu\mathrm{A}$ was chosen for a cuff resistance of approx. $5\,\mathrm{k}\Omega$.

A prerequisite of eq. (5.16), is that the M1 pair operates in weak inversion. This is ensured by setting their aspect ratio such that the inversion coefficient (IC), fulfills: $IC < 0.1$ (see appendix 1). For the rather large bias current, an aspect ratio of $7200\,\mu m/1\,\mu m$ was found to be necessary in order to ensure that the M1 pair is operating in deep subthreshold.

M4 and M5 are multiples of the same unit MOST. To minimize the M4 and M5 MOSTs noise contribution, they are biased in strong inversion to lower their transconductance compared to g_{m1}. To bring the 1/f-noise contribution of M4 and M5 approximately on par with the M1 contribution, the total aspect ratio of M4 and M5 was set to $152\,\mu m/6\,\mu m$.

5.1.3.1 Common-Mode Feedback

As the gain stage is fully differential, a common-mode feedback (CMFB) circuit is necessary to stabilize the operating point. Furthermore, as the M3 bias MOST has it's source clamped to ground, so does the amplifier output need to have a common-mode voltage equal to gnd, in order to bias the M2 load MOST correctly.

Since the output magnitude of the signal is very small $(\pm 1\,\mathrm{mV})$, the common-mode signal (CMS) can be readily extracted by splitting the load M2, into two MOSTs hence giving the CMS in the "middle" of the load. The CMFB circuit shown in fig. 5.3 is simply a differential pair which compares the CMS to gnd and controls the current source M6 in order to regulate the amplifier output common voltage to gnd.

5.1.3.2 Offset Compensation

Though the input pair MOSTs are quite large, the inherent offset in the pair still has a worst case statistical value of approx. $\Delta v_{\mathrm{th1}} = 1\,\mathrm{mV}$. For a gain of 100, this would result in an output referred offset of 100 mV bringing the load MOST out of weak inversion and thus disrupting the amplifier operation.

This points to the necessity of implementing an offset compensation scheme. Typically, the well-known auto-zero technique [9] is used for offset-compensation. This technique suffers from downfolding of wideband thermal noise though and is thus deemed unsuitable for our needs as minimization of thermal noise is critical to our application.

A block diagram of the employed offset compensation circuit (OCC) is shown in fig. 5.3. Here, the differential output voltage is amplified and passed through a low-pass filter, the output of which is used for controlling the M4 current source MOSTs thus canceling the offset in a current-steering manner. For an input pair offset Δv_{th1}, the error current in the input pair is:

$$\Delta i_{\mathrm{err}} = g_{\mathrm{m1}} \Delta v_{\mathrm{th1}} = \frac{I_{\mathrm{D}}}{n V_{\mathrm{T}}} \Delta v_{\mathrm{th1}} \tag{5.17}$$

The current of each branch in the input pair is partitioned between M4 and M5. Let α denote the partitioning ratio and Δv_{os} the small signal output voltage of the OCC. The resulting current steered by

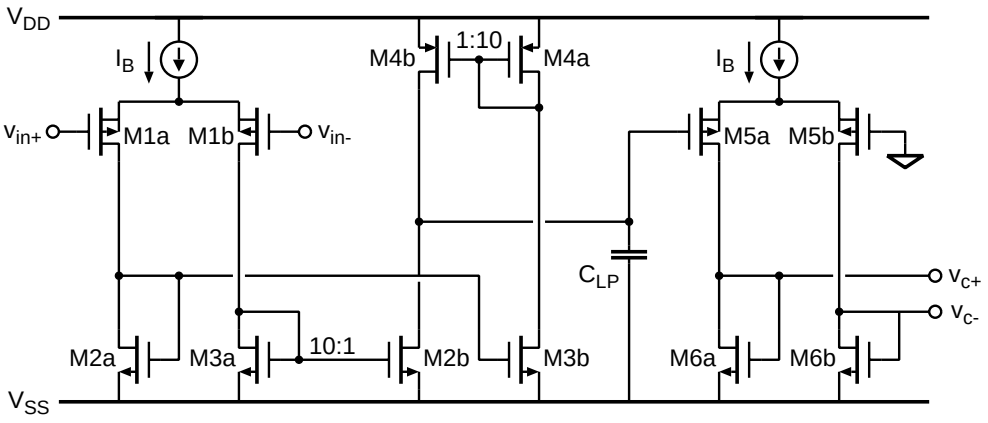


Figure 5.4: Stage 1 offset compensation circuit (OCC).

the OCC is then given by:

$$\Delta i_{\mathrm{corr}} = g_{\mathrm{m4}} \Delta v_{\mathrm{os}} = \frac{2\alpha I_{\mathrm{D}}}{V_{\mathrm{eff4}}} \Delta v_{\mathrm{os}} \tag{5.18}$$

Requiring that the OCC should be well capable of nulling any offset induced error current we find that:

$$\alpha > \frac{nV_{\mathrm{eff4}}}{2V_{\mathrm{T}}} \frac{\Delta v_{\mathrm{th1}}}{\Delta v_{\mathrm{os}}} \tag{5.19}$$

The partitioning ratio is determined by the worst-case estimate of the input pair offset, i.e. $3\sigma\{\Delta v_{\mathrm{th1}}\}$ and the allowable OCC output swing Δv_{os}. The maximum OCC output swing is in practice some fraction of V_{eff4}, as the internal OCC devices generating Δv_{os} are scaled replicas of the M4 pair to ensure good matching. Allowing a good margin for the offset error, the partitioning ratio α, was in our case set to 13%.

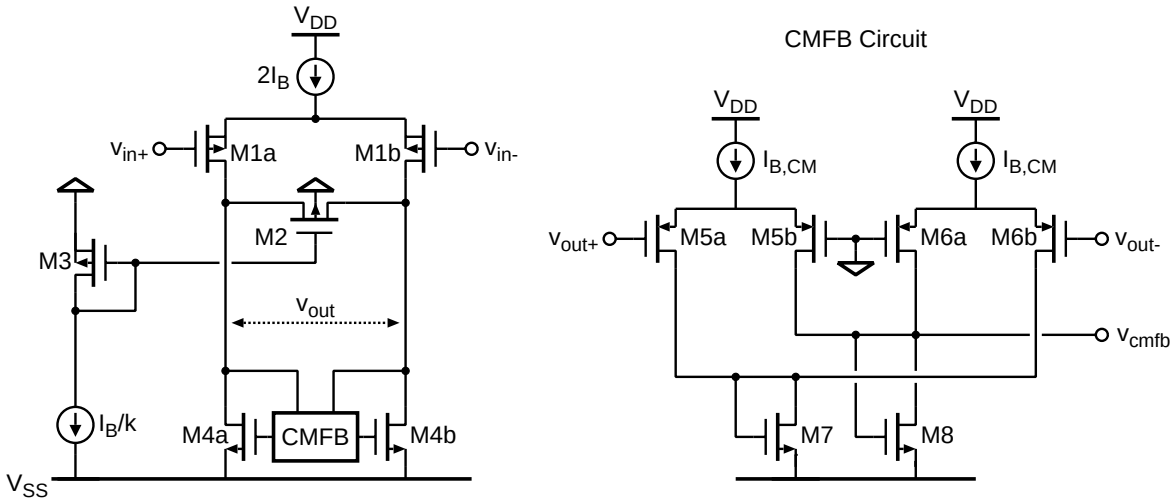
The speed of the OCC should be sufficiently low not to filter out part of the signal band, i.e. the OCC bandwidth should remain below 400 Hz.

The realized OCC is shown in fig. 5.4. The differential output of the main amplifier is sensed through the OCC M1 input pair. The resulting small signal current is divided by a factor of ten before being integrated on the C_{LP} capacitor. The resulting control voltage on the integrating capacitor is converted to a fully differential control signal through the M5 pair which is diode-loaded by the M6 pair. The M6 MOSTs are scaled replicas of the steered M4 current sources of fig. 5.3 for good matching.

The bandwidth of the OCC is given by the input pair downscaled transconductance to the integrating capacitor:

$$f_{\mathrm{3dB}} = \frac{1}{10} \cdot \frac{g_{\mathrm{m1}}}{2\pi C_{\mathrm{LP}}} \tag{5.20}$$

A bias current of $1\,\mu$A, was used in each branch of the input pair and an integrating capacitor of $1\,$nF to give a 3dB frequency of 100 Hz.



(a) Amplifier second stage.

(b) CMFB circuit for amplifier second stage.

Figure 5.5: Amplifier second stage schematic.

5.1.4 Second Stage

The second amplifier stage is shown in fig. 5.5. As the input signal to the second stage has been preamplified by a factor 100 by the first stage, the noise performance requirements of this stage can be significantly relaxed indicating reduced biasing currents. The gain of the second stage is chosen to 100, thus the peak-peak output voltage of the second stage will attain values of $\pm 100\,\mathrm{mV}$, indicating that weak inversion operation is not applicable here. Instead, all devices are biased in the strong inversion region with an effective voltage sufficient for handling the output swing.

The biasing current is rather determined by the bandwidth requirements of the stage. However, a PMOST input pair is still appropriate due to their superior 1/f-noise performance. For a 3dB bandwidth of 10 kHz, the input pair bias current was set to $2.5\,\mu\mathrm{A}$ in each branch.

5.1.5 Measurements

The amplifier was implemented in a standard single-poly, 3-metal, N-well $0.5\,\mu\mathrm{m}$ CMOS process. A microphotograph of the implemented chip is shown in fig. 5.6. Unfortunately, the chip was covered with metal 3 dummy filler blocks deteriorating the chip photo considerably. Two versions of each of the amplifying stages were implemented and are contained in the top half of the chip.

The bottom half of the chip contains an RF transceiver unit for a distributed transducer system applicable for implantable systems and is described in [5]. To be noticed is the square block in the lower right corner of the chip containing the designers initials. This block is actually two 1 nF capacitors implemented as poly-substrate capacitors.

Frequency response: An example frequency response curve of the two-stage amplifier is shown in

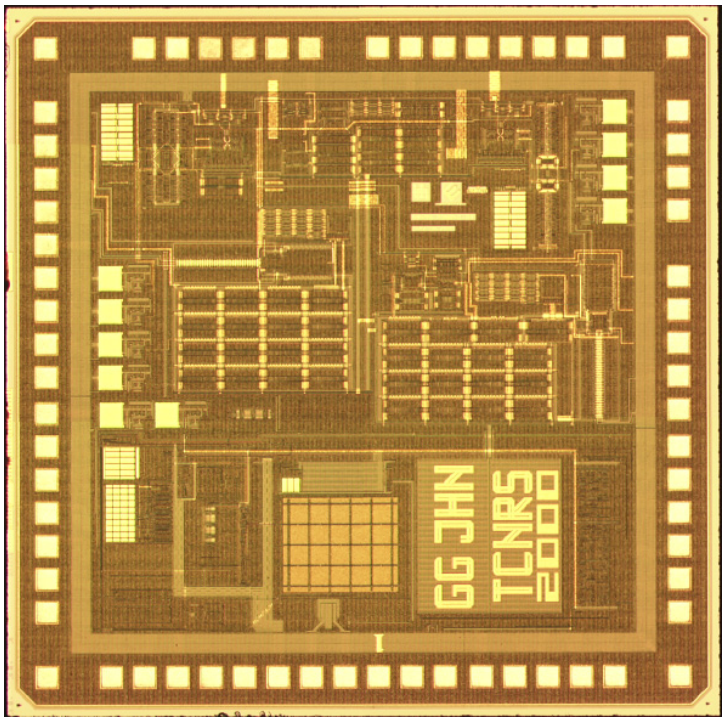


Figure 5.6: Test chip microphotograph.

fig. 5.7(a). The lower and upper 3-dB frequencies of the amplifier are seen to be approximately:

$$f_{\mathrm{3dB,lower}} = 100\,\mathrm{Hz} \tag{5.21}$$

$$f_{\mathrm{3dB,upper}} = 10\,\mathrm{kHz} \tag{5.22}$$

Some gain variation was seen over the set of 10 test chips, but remained within a +/- 10 % margin of the nominal gain of 80 dB. The main cause of this variation is thought to be due to the fact that the amplifier gain is sensitive to any shift in the common-mode voltage. I.e. shifting the common-mode voltage will cause a deviation in the M2 drain conductance g_{ds}, and hence modulate the gain. So if the CMFB circuit has some offset associated to it, this will alter the gain accordingly.

Harmonic distortion: Fig. 5.7(b) illustrates the measured output spectrum for a full amplitude input sinusoid at 1 kHz. The spectrum has been normalized to the maximum output signal level, which for a maximum nominal input of $20\,\mu\mathrm{V_{pp}}$, is $200\,\mathrm{mV_{pp}}$. The odd order harmonics are seen to dominate the spectrum as the amplifier is fully differential. The utilized linearized load MOST operating in triode combined with a traditional differential input pair does not provide a high THD, which on average was measured to be 2.2 % for the set of test chips.

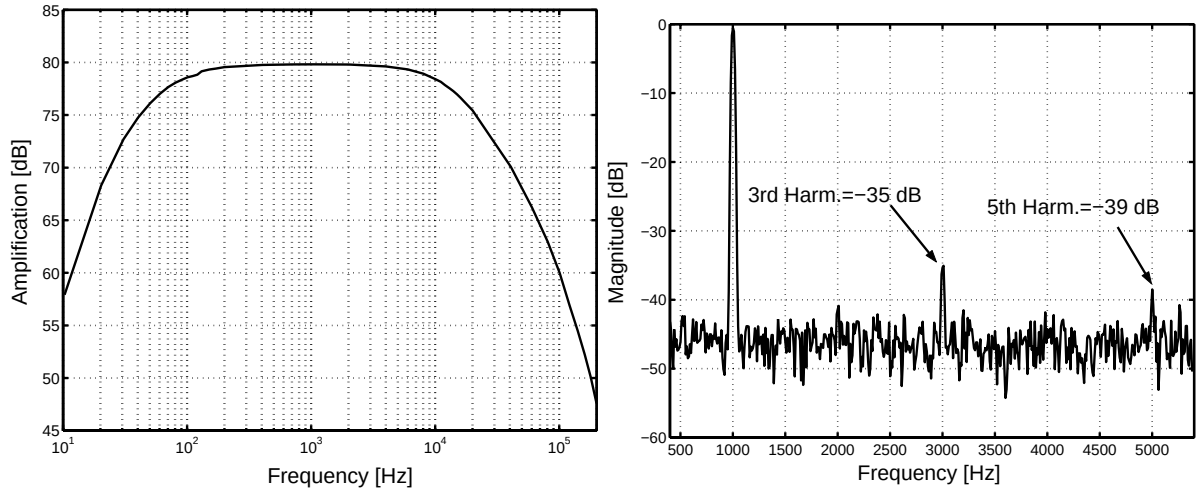
PSRR and CMRR: The CMRR and the PSRR were measured respectively by applying a 100 mV input tone as a common-mode signal and by AC-coupling the signal to the positive supply. The resulting measured rejection ratios are depicted in fig. 5.8(a). We see that for input frequencies below 100 kHz, the PSRR is above 84 dB, whereas the CMRR remains above 87 dB.

Noise: The measured output noise spectral density is shown in fig. 5.8(b). For lower frequencies, the noise spectrum is clearly dominated by 1/f-noise. If we define the noise corner to be the 3dB point, an equivalent noise corner of 800 Hz is observed from the spectrum. So even though very large input MOSTs were utilized in the layout ($7200\,\mu m/1\,\mu m$), 1/f-noise still dominate the noise performance at frequencies below 800 Hz. The ripple and steep declining slope observed at the end of the spectrum is due to an inserted Chebychev filter at the amplifier output. The average output noise floor over the signal bandwidth is $50\,\mu\mathrm{V}/\sqrt{\mathrm{Hz}}$, equivalent to an amplifier input referred noise of $5\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$.

Offset performance: The offset compensation scheme by current steering was tested by applying a differential offset voltage to the amplifier input. The resulting amplifier stage 1 output voltages $v_{\mathrm{O}+}\,\mathrm{and}\,v_{\mathrm{O}-}$, are shown versus the applied offset voltage in fig. 5.9. It is seen that offsets with a magnitude smaller than 3.5 mV are canceled by the scheme. Alongside the measured output voltages, the simulated steered biasing currents for each branch of the input differential pair is shown illustrating the current partitioning in each branch. Some residual offset will remain due to offsets in the offset regulating loop. For the set of test chips, this residual offset was observed to be less than 4 mV at the output of stage 1, indicating a stage 1 input offset smaller than $40\,\mu\mathrm{V}$.

Power consumption: An average amplifier bias current of $91.5\,\mu\mathrm{A}$ was measured at a nominal supply voltage of 3 V, thus the circuit consumes $\simeq 275\,\mu\mathrm{W}$ of power. The employed output buffer amplifiers were not included in this power measurements. Bias circuitry was also exempt from this measurement as large biasing currents were employed in order not to face measurement problems due to weak current levels in the bias circuitry. This implies that a slightly higher power consumption should be expected when a dedicated bias circuit is included in the design.

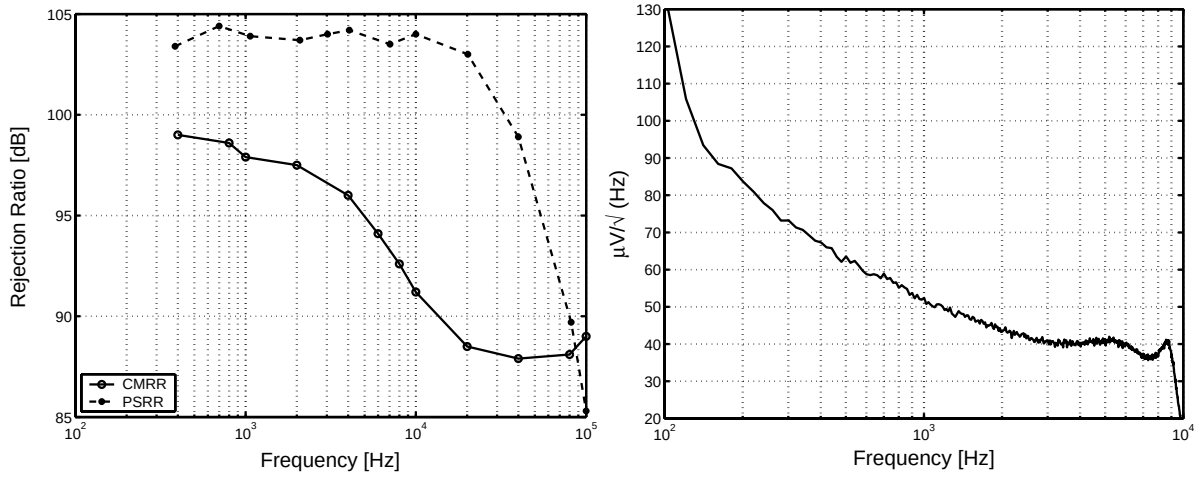
Slew rate (SR): In case of overloading the amplifier input, the output signal will be dominated by slewing. The amplifier slew rate was measured by applying a square wave signal with an amplitude of $400\,\mathrm{mV_{pp}}$. The observed slew rate was $500\,\mathrm{V/ms}$, ensuring fast recovery from an overload condition of the amplifier input.



(a) Typical amplifier frequency response.

(b) Typical measured amplifier output spectrum.

Figure 5.7: Frequency response and harmonic distortion of the two-stage neural signal amplifier.



(a) Typical measured CMRR and PSRR.

(b) Measured output noise power spectral density.

Figure 5.8: PSRR, CMRR and output noise PSD for the two-stage nerve signal amplifier.

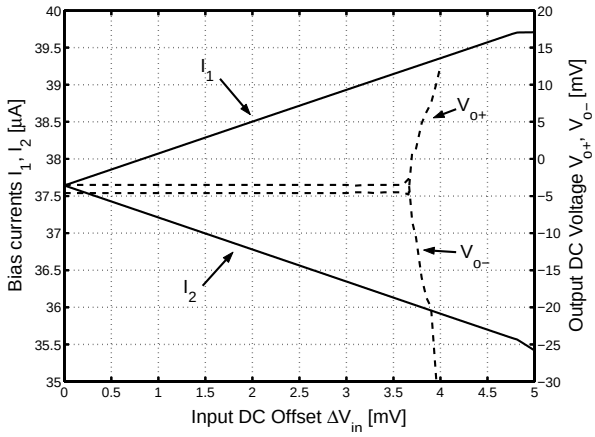


Figure 5.9: Offset cancellation by current steering.

Supply Voltage	$3\,\mathrm{V}$
Power Consumption	$275\,\mu\mathrm{W}$
THD(@ f=1 kHz)	$2.2\,\%$
Typical Input Offset (Stage 1)	$< 40\,\mu\mathrm{V}$
CMRR	$> 87\,\mathrm{dB},\ f < 100\,\mathrm{kHz}$
PSRR	$> 84\,\mathrm{dB},\ f < 100\,\mathrm{kHz}$
Gain	$10000 \pm 10\%$
Equivalent input referred noise	$5\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$
Slew rate	$500\,\mathrm{V}/\mathrm{ms}$

Table 5.1: Measured amplifier V1 performance.

5.1.6 Nerve Amplifier V1 Summary

A summary of the measured amplifier performance is given in table 5.1.

The implemented amplifier is seen to have a performance which is within the boundaries set forth for an implantable amplifier. However several pros and cons are identifiable in this particular amplifier implementation:

Advantages

- Each of the amplification stages employ an extremely simple architecture employing very few transistors indicating a low noise contribution from the circuit.
- Very high power efficiency due to the simple architecture.
- The circuit is realizable using only MOSTs, i.e. no extra masks are needed for special layers e.g. resistors etc.

Drawbacks

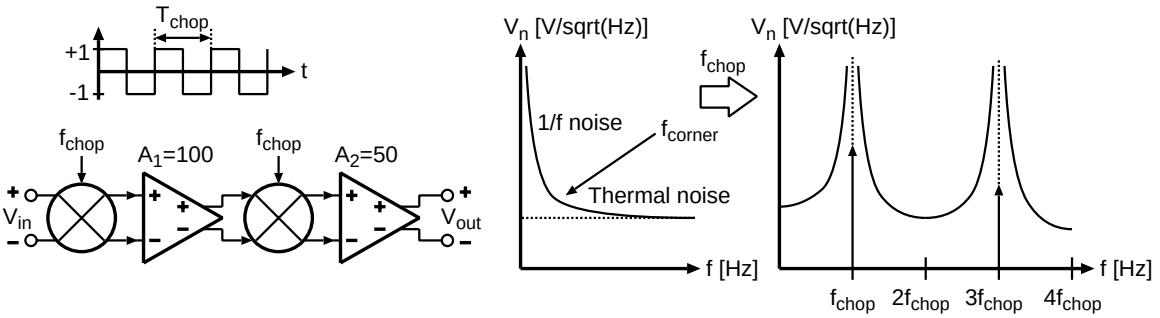
- A dual supply is needed as the circuit employs ground connection for biasing. As current needs to be drawn from the ground pin, some active circuitry would be needed to realize this supply as the inductive power supply intrinsically only provides a single supply.
- Very large capacitors are needed for stabilizing the OCC.
- Very large input MOSTs are employed for the weak inversion operation of the first stage input pair however the 1/f-noise still dominates at the lower inband frequencies.
- The output offset of the first stage of the amplifier is large enough to warrant either an AC-coupling scheme between the stages or some OCC for the second stage as well. This implies either realization of on-chip floating capacitors coupled with very large resistors which could very well constitute a major design challenge, or increased circuit complexity and power consumption for a stage two OCC.

- The amplifier gain is dependent on the accuracy of the CMFB circuit and thus sensitive to common-mode voltage shifts.

5.2 Nerve Signal Amplifier V2

In the previous section, an implantable instrumentation amplifier for nerve signals was presented. However, the presented architecture suffered from several deficiencies. To counter the drawbacks of the first implemented amplifier, another amplifier topology coupled with the well-known chopper technique is employed in the amplifier which is the subject of this section.

A block diagram of the proposed improved amplifier is shown in fig. 5.10(a). Here, the first stage of the amplifier is chopped, resulting in the noise spectrum depicted in fig. 5.10(b). Using the chopper technique, we can effectively shift the 1/f-noise out of the signal band as derived in section 2.3, hence resolving the issue of low-frequency inband 1/f-noise as was seen to be present in the instrumentation amplifier presented in the previous sections.



(a) Amplifier block diagram.

(b) Chopped CMOS amplifier noise.

Figure 5.10: Chopper modulated amplifier block diagram and chopped noise spectrum.

5.2.1 Gain Stage Principle of Operation

A schematic including only the principal MOSTs of the employed gain stage is shown in fig. 5.11. Instead of employing a linearized MOST as load, as was the idea in the first version of the nerve signal amplifier, we have loaded the M1 PMOST input pair by a secondary PMOST pair M2 in a folded cascode manner.

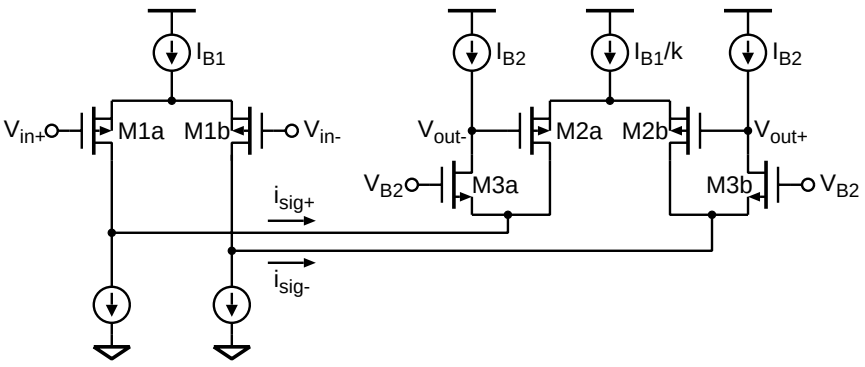


Figure 5.11: Gain stage principle schematic.

Using this scheme, the gain is defined by the ratio of transconductances:

$$A_{\mathrm{V}} = -\frac{g_{\mathrm{m1}}}{g_{\mathrm{m2}}} \tag{5.23}$$

If both the input pair and load pair are assumed to operate in the same region, the gain function can be written for the weak- and strong inversion regions respectively:

Weak inversion:

$$A_{\mathrm{V,WI}} = -\frac{2K'_{\mathrm{P}}\left(W/L\right)_1 e^{\frac{V_{\mathrm{G1}}-V_{\mathrm{th1}}-nV_{\mathrm{S1}}}{V_{\mathrm{T}}}}}{2K'_{\mathrm{P}}\left(W/L\right)_2 e^{\frac{V_{\mathrm{G2}}-V_{\mathrm{th2}}-nV_{\mathrm{S2}}}{nV_{\mathrm{T}}}}} \tag{5.24}$$

Strong inversion:

$$A_{\mathrm{V,SI}} = -\frac{\frac{K'_{\mathrm{P}}}{n}\left(\frac{W}{L}\right)_1\left(V_{\mathrm{G1}}-V_{\mathrm{th0}}-nV_{\mathrm{S1}}\right)}{\frac{K'_{\mathrm{P}}}{n}\left(\frac{W}{L}\right)_2\left(V_{\mathrm{G2}}-V_{\mathrm{th0}}-nV_{\mathrm{S2}}\right)} \tag{5.25}$$

By doing the layout in a common centroid fashion and utilizing the same unit MOST for both the input and load pair, good matching can be achieved. If we furthermore scale the bias current of the load, according to the implemented ratio between the input- and output pair aspect ratios, i.e. set the scaling factor *k* equal to the desired gain, the same channel inversion level will exists in both the input- and output pair, giving equal gate-source voltages for both pairs.

If these requirements are met, both the expressions for weak- and strong inversion converge to a simple gain expression only dependent on geometrics:

$$A_{\mathrm{V}} = A_{\mathrm{V,WI}} = A_{\mathrm{V,SI}} = -\frac{\left(W/L\right)_1}{\left(W/L\right)_2} \tag{5.26}$$

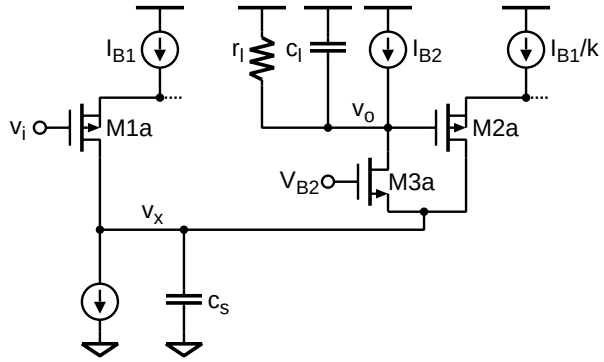
Eq. (5.26) is valid if both pairs have the same threshold voltage, which is easily achieved by using PMOSTs in an N-well allowing the bulk terminal to be connected to the source.

Using this gain stage, no dual supply is needed as was the case for the first version of the amplifier. The gain is dependent solely on geometric ratios and insensitive to the common-mode level and there is no dependence on the threshold slope *n*. Also, no special biasing circuit is needed for the load pair.

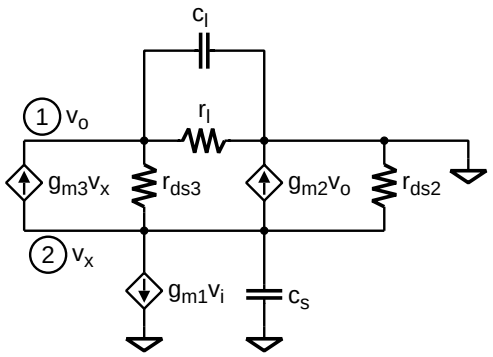
The M3 pair ensures that a low impedance path exists from the M1 input pair to the M2 load pair. As the M3 pair should have it's transconductance maximized, these MOSTs operate close to the weak inversion region having an inversion coefficient of approximately 0.3.

The bias current of the input pair is determined by the noise requirements as will be discussed later, hence also setting the bias current of the output pair. The M3 pair bias current is set by the required settling behavior of the output.

Consider the half-circuit schematic shown in fig. 5.12(a). Here the sources of the M1- and M2 pair are considered to be AC-ground. A load resistance r_{l}, and a load capacitance c_{l}, has been associated to the amplifier output. The node v_{x}, has a parasitic capacitance c_{s} connected to it. The equivalent small-signal diagram is shown in fig. 5.12(b), where the MOST parasitic capacitors are lumped into the two shown capacitors. The M1 drain-source resistance has been ignored in this diagram. For the two



(a) Half circuit for small-signal analysis.



(b) Small-signal diagram.

Figure 5.12: Half-circuit schematic and small signal diagram for settling analysis of amplifier gain stage.

internal nodes in the small circuit diagram, we have the following nodal equations:

$$1 \quad : \quad v_{\mathrm{x}}\left(g_{\mathrm{ds3}}+g_{\mathrm{ds2}}+sc_{\mathrm{s}}+g_{\mathrm{m3}}\right)+v_{\mathrm{o}}\left(g_{\mathrm{m2}}-g_{\mathrm{ds3}}\right)+g_{\mathrm{m1}}v_{\mathrm{i}}=0 \tag{5.27}$$

$$2 \quad : \quad v_{\mathrm{o}}\left(g_{\mathrm{ds3}}+g_{\mathrm{l}}+sc_{\mathrm{l}}\right)-v_{\mathrm{x}}\left(g_{\mathrm{ds3}}+g_{\mathrm{m3}}\right)=0 \tag{5.28}$$

Solving these nodal equations, one arrives at a transfer function given by:

$$A_{\mathrm{V}}(s)=\frac{v_{\mathrm{o}}}{v_{\mathrm{i}}}\approx-\frac{\frac{g_{\mathrm{m1}}}{g_{\mathrm{m2}}+G}}{1+s\left(\frac{Gc_{\mathrm{s}}+g_{\mathrm{m3}}c_{\mathrm{l}}}{g_{\mathrm{m3}}(g_{\mathrm{m2}}+G)}\right)+s^2\left(\frac{c_{\mathrm{l}}c_{\mathrm{s}}}{g_{\mathrm{m3}}(g_{\mathrm{m2}}+G)}\right)},\ G=g_{\mathrm{ds3}}+g_{\mathrm{l}} \tag{5.29}$$

where the following assumptions have been made in the derivation:

$$g_{\mathrm{m2}} \quad \gg \quad g_{\mathrm{ds3}} \tag{5.30}$$

$$g_{\mathrm{m3}} \quad \gg \quad g_{\mathrm{ds3}}+g_{\mathrm{ds2}} \tag{5.31}$$

Eq. (5.29) is recognized as a second order all-pole transfer function. In [57], the Q-factor and the transfer function resonant frequency are used alongside the DC-gain A_0, to characterize these types of transfer functions. From [57] we have:

$$A_{\mathrm{V}}(s)=\frac{A_0}{1+\frac{s}{\omega_0 Q}+\frac{s^2}{\omega_0^2}} \tag{5.32}$$

By comparing eq. (5.32) to eq. (5.29) we can identify the DC-gain:

$$A_0\approx-\frac{g_{\mathrm{m1}}}{g_{\mathrm{m2}}},\ g_{\mathrm{m2}}\gg g_{\mathrm{ds3}}+g_{\mathrm{l}} \tag{5.33}$$

which is seen to match the desired expression given in eq. (5.26).

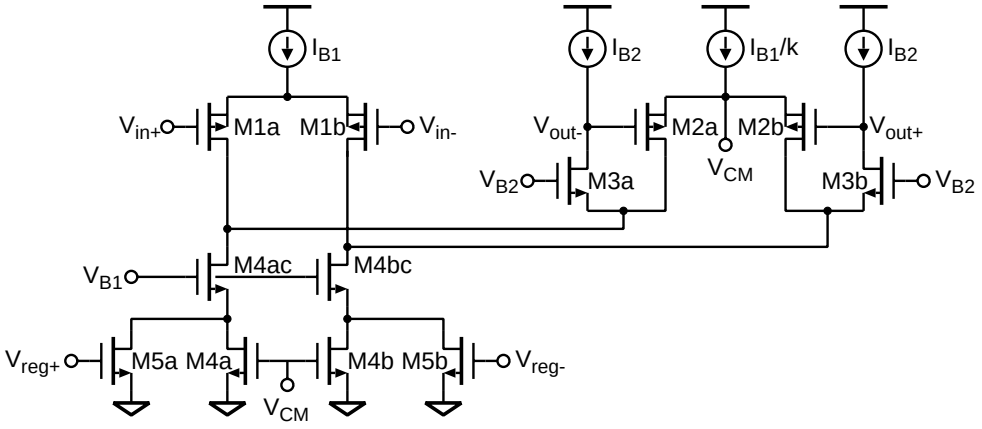


Figure 5.13: Chopper amplifier first stage schematic.

The squared resonant frequency ω_0, can be found to be:

$$\omega_0^2 \approx \frac{g_{\mathrm{m3}} g_{\mathrm{m2}}}{c_{\mathrm{l}} c_{\mathrm{s}}},\ g_{\mathrm{m2}} \gg g_{\mathrm{ds3}} + g_{\mathrm{l}} \tag{5.34}$$

Finally the Q factor can be found:

$$\begin{aligned} Q &= \frac{\sqrt{g_{\mathrm{m3}} \left(g_{\mathrm{m2}} + g_{\mathrm{ds3}} + g_{\mathrm{l}}\right) c_{\mathrm{s}} c_{\mathrm{l}}}}{\left(g_{\mathrm{ds3}} + g_{\mathrm{l}}\right) c_{\mathrm{s}} + g_{\mathrm{m3}} c_{\mathrm{l}}} \\ &\approx \frac{\sqrt{g_{\mathrm{m3}} g_{\mathrm{m2}} c_{\mathrm{s}} c_{\mathrm{l}}}}{g_{\mathrm{ds3}} c_{\mathrm{s}} + g_{\mathrm{m3}} c_{\mathrm{l}}},\ g_{\mathrm{m2}} \gg g_{\mathrm{ds3}} + g_{\mathrm{l}},\ g_{\mathrm{ds3}} \gg g_{\mathrm{l}} \end{aligned} \tag{5.35}$$

In order not to have any peaking in the amplifier transfer function, the Q factor should remain below $\sqrt{1/2}$ [57]. The Q factor is linked to overshoot in the step response by [13]:

$$\mathrm{overshoot} = e^{\frac{-\pi}{\sqrt{4Q^2-1}}} \tag{5.36}$$

By choosing a Q factor of $\sqrt{1/2}$, the step overshoot is below 5 %, while retaining the fast settling desirable in a chopped amplifier stage. Studying eq. (5.35) we see that most of the parameters are already fixed. The source capacitance c_{s}, is set by the dimensions of the employed bias current sources which also fix g_{m2} as this parameter is set by the desired gain. Increasing the load capacitance c_{l}, would be a poor choice as reduced bandwidth would result. Finally, increasing g_{m3} causes the Q factor to drop approximately by the square root. I.e. by increasing the I_{B2} bias current, the desired Q factor can be achieved.

5.2.2 First Stage

A schematic illustrating the amplifier first stage is shown in fig. 5.13. All the shown current sources are implemented as cascoded MOST sources.

Besides the gain stage, the functionality of which was the subject of the previous section, the stage

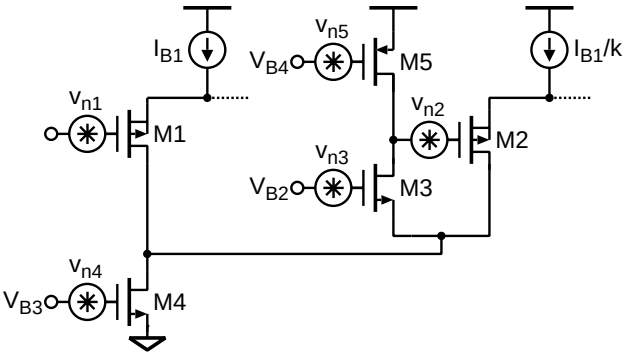


Figure 5.14: Chopper amplifier first stage half circuit for noise analysis.

also needs a CMFB scheme as the circuit is fully differential. Common-mode stabilization is achieved by realizing that the source potential of the M2 pair is in fact the level shifted output common-mode voltage. Using this potential for biasing the M4 current source pair thus realizes an extremely compact CMFB circuit as all the utilized MOSTs are already part of the amplifier circuit. The M4 pair is biased in the triode region and thus act as voltage controlled resistors. M4ac and M4bc are cascodes included to ensure a sufficiently high output impedance for the M4 current source pair.

The M5 pair realizes an offset nulling port in the same manner as explained in section 5.1.3.2, similar analysis can be performed for this stage giving the necessary partitioning ratio between M4 and M5. Running a large number of Monte-Carlo simulations resulted in a current ratio of 25 %. The scheme for controlling the offset nulling port is the subject of section 5.2.5.

As the first stage directly amplifies the weak sensor signal, a noise analysis for the stage is needed, which is the subject of the following section.

5.2.2.1 Noise Analysis

A half circuit for noise analysis is shown in fig. 5.14, excluding the M4c cascode MOSTs. The noise sources stemming from the input and load pair current sources are ignored as they will only contribute common-mode noise. As the 1/f-noise is dealt with using the chopper modulation technique, only the thermal noise is included in this analysis.

Assuming all noise sources to be uncorrelated the total input referred noise power can be found:

$$v_{\mathrm{ni,tot}}^2(f) \approx v_{\mathrm{n1}}^2(f) + v_{\mathrm{n2}}^2(f)\left(\frac{g_{\mathrm{m2}}}{g_{\mathrm{m1}}}\right)^2 + v_{\mathrm{n4}}^2(f)\left(\frac{g_{\mathrm{m4}}}{g_{\mathrm{m1}}}\right)^2 + v_{\mathrm{n5}}^2(f)\left(\frac{g_{\mathrm{m5}}}{g_{\mathrm{m1}}}\right)^2 \tag{5.37}$$

Here, the M3 contribution has been ignored as this transistor operates as a folded cascode.

Rewriting eq. (5.37), by inserting the noise expressions we get:

$$v_{\mathrm{ni,tot}}^2(f) = 4kT\left[\frac{c_1}{g_{\mathrm{m1}}} + \frac{c_2 g_{\mathrm{m2}}}{g_{\mathrm{m1}}^2} + \frac{c_4 g_{\mathrm{ds4}}}{g_{\mathrm{m1}}^2} + \frac{c_5 g_{\mathrm{m5}}}{g_{\mathrm{m1}}^2}\right] \tag{5.38}$$

Here it should be noted that M4 is operated in the triode region. By requiring that the noise contribution

from the input transistor M1 should dominate and by inserting the *c* factors we have:

$$1 \gg \underbrace{\frac{g_{\mathrm{m2}}}{g_{\mathrm{m1}}}}_{\mathrm{I}} + \underbrace{2\frac{g_{\mathrm{ds4}}}{g_{\mathrm{m1}}}}_{\mathrm{II}} + \underbrace{\frac{4g_{\mathrm{m5}}}{3g_{\mathrm{m1}}}}_{\mathrm{III}} \tag{5.39}$$

where we have for each contribution:

`I:` This ratio is the reciprocal of the stage gain which is set to 100, so it is 1/100.

`II:` M4 operates in the triode region and M1 is near weak inversion. Inserting the conductance expression we have:

$$2\frac{g_{\mathrm{ds4}}}{g_{\mathrm{m1}}} = 2\frac{V_{\mathrm{T}} n^{3/2}\sqrt{2K'_N\,(W/L)_4}}{\sqrt{I_{\mathrm{D}}}} \tag{5.40}$$

This expression points to minimizing the M4 aspect ratio and increasing the bias current. I_{D} is the M1 bias current, and will thus have a relatively high value. Thus minimizing eq. (5.40) does not incur severe design constraints.

`III:` M5 operates in the strong inversion region thus giving:

$$\frac{4g_{\mathrm{m5}}}{3g_{\mathrm{m1}}} = \frac{8nV_{\mathrm{T}}I_{\mathrm{D5}}}{3V_{\mathrm{eff5}}I_{\mathrm{D1}}} \tag{5.41}$$

As M4 is designed for good matching, V_{eff} is set to a few hundred millivolts, i.e. much larger than the thermal voltage V_{T}. This noise contribution can be further minimized, by keeping the M5 bias current sufficiently low. The M5 bias current is ultimately set by the necessary M3 transconductance g_{m3}, as this parameter sets the regulated cascode loop Q factor as was seen in section 5.2.1. Hence a trade-off exists between the Q factor and noise contribution of M3.

5.2.2.2 Dimensioning

As a maximized transconductance is desired for the input pair for optimum thermal noise suppression, the weak inversion operating region is attractive. To obtain deep subthreshold operation (IC<0.1), the first version of the amplifier employed an input pair with a very large aspect ratio of 7200 for a bias current of $76\,\mu\mathrm{A}$.

In fig. 5.15, the inversion coefficient and the transconductance of a single NMOST biased at $10\,\mu\mathrm{A}$ is shown versus the aspect ratio. This figure reveals that most of the transconductance gain is obtained going from strong- to moderate inversion (10<IC<1). Hereafter, the g_{m} slope is limited as the g_{m} grows toward it's asymptotic end value obtained in the weak inversion region. A similar curve done for the input pair M1, biased at $100\,\mu\mathrm{A}$, shows that 90 % of the asymptote value is obtained for an aspect ratio of $300\,\mu\mathrm{m}/0.6\,\mu m$, which are the chosen dimensions and bias current. Hence similar noise performance can be achieved for a much smaller aspect ratio than was used in the first version of the amplifier.

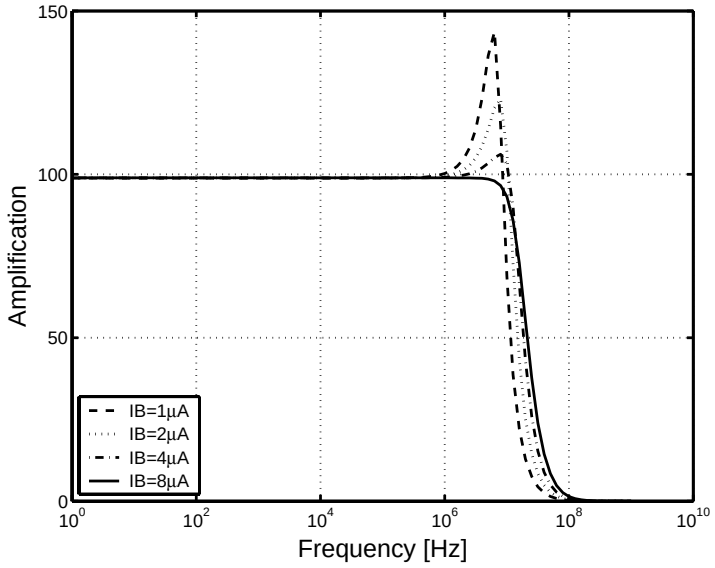


Figure 5.16: Stage 1 frequency response for increasing M3 bias current.

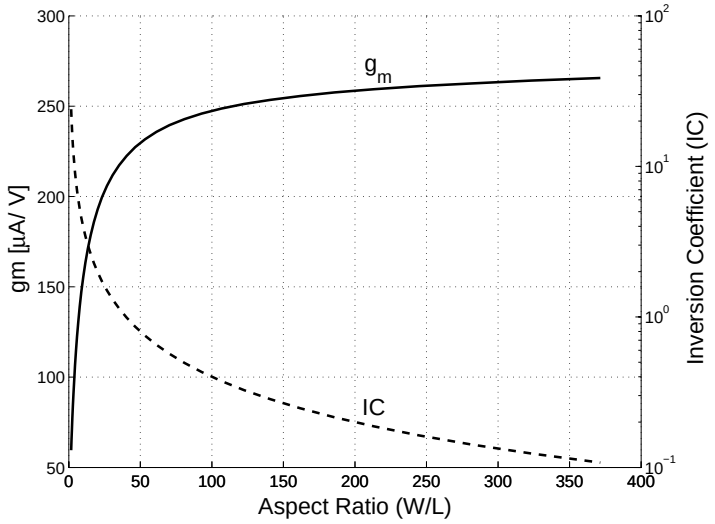
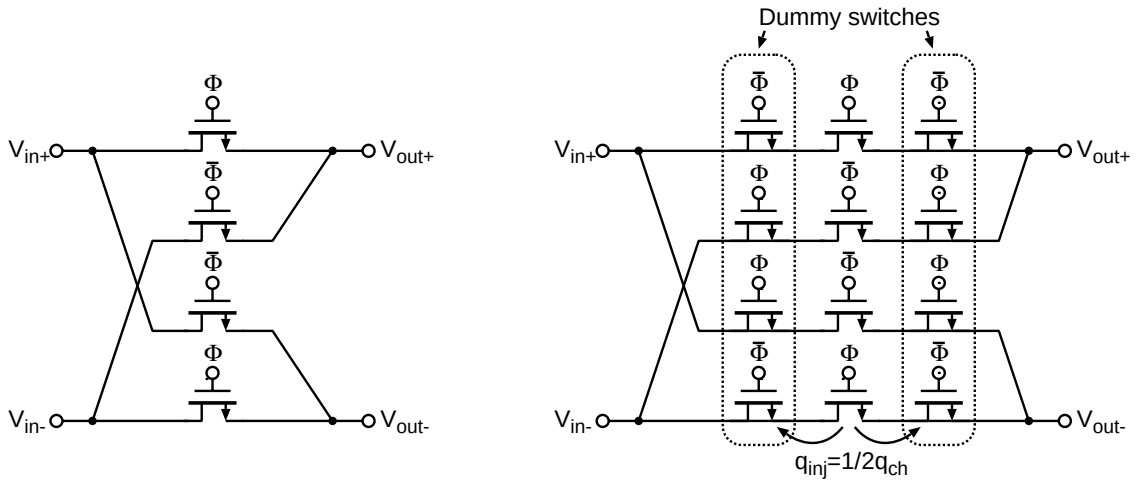


Figure 5.15: Inversion coefficient and transconductance vs. aspect ratio.

In [58], an aspect ratio in excess of 30000 at a tail current of $400\,\mu$A is stated to be necessary for implementing a low-noise CMOS preamplifier for nerve cuff signals. The resulting large MOSTs are used as a justification for using bipolar inputs due to their superior transconductance and 1/f-noise performance. However the above discussion shows that using the MOST in moderate inversion will suffice for thermal noise suppression and we have from section 2.3 that the 1/f-noise can be dealt with by using chopper modulation.

Choosing the first stage gain to 100, fixes the load pair bias current at $1\,\mu$A.

As previously stated, setting the bias current for the M3 transistors affects both the Q factor and the noise performance. The simulated stage 1 frequency response for different bias currents is shown



(a) Differential NMOST chopper

(b) Chopper with dummy switches

Figure 5.17: Fully differential NMOST chopper realization.

in fig. 5.16. A clear peak is observable for low M3 bias current levels, i.e. the Q factor is higher than $\sqrt{1/2}$. A flat frequency response was obtained at an M3 bias current of $5\,\mu$A, which still does not provide a high enough M3 transconductance for it to significantly contribute to the total output noise and hence is a good trade off.

As M3 is the main noise contributor in the circuit, apart from the M1 input pair, the remaining MOSTs do not contribute significantly to the output noise. The total MOST noise contribution excluding M1, was simulated to be less than 15 %.

5.2.3 Chopper Realization

Chopper circuits are easily implemented in MOS technology as only four switches are needed for a fully differential chopper implementation as shown in fig. 5.17(a). Here, the inversion of signal sign is done simply by alternating the signal input using a clock Φ, and it's counter-phase $\overline{\Phi}$.

5.2.3.1 Input Chopper

The chopper circuit which is placed at the front of the amplifier stage, is noise critical. The thermal noise due to the resistive channel in the MOST is given by:

$$i_{\mathrm{n}}^2(f) = \quad 4kT \quad \underbrace{K'(W/L)V_{\mathrm{eff}}}_{g_{\mathrm{ds}}} \tag{5.42}$$

In eq. (5.42), the maximum V_{eff}, is given by the available supply voltage. As NMOSTs have a higher K', these are an appropriate choice for the switches. The aspect ratio is then set by requiring the noise level be below that of the cuff resistance R_{cuff}, and is in our case set to $12\,\mu\mathrm{m}/0.35\,\mu\mathrm{m}$.

Using non-minimal switch sizes also implies increased charge injection and clock feedthrough. Charge injection occurs as the inversion charge present in channel is released when the gate voltage

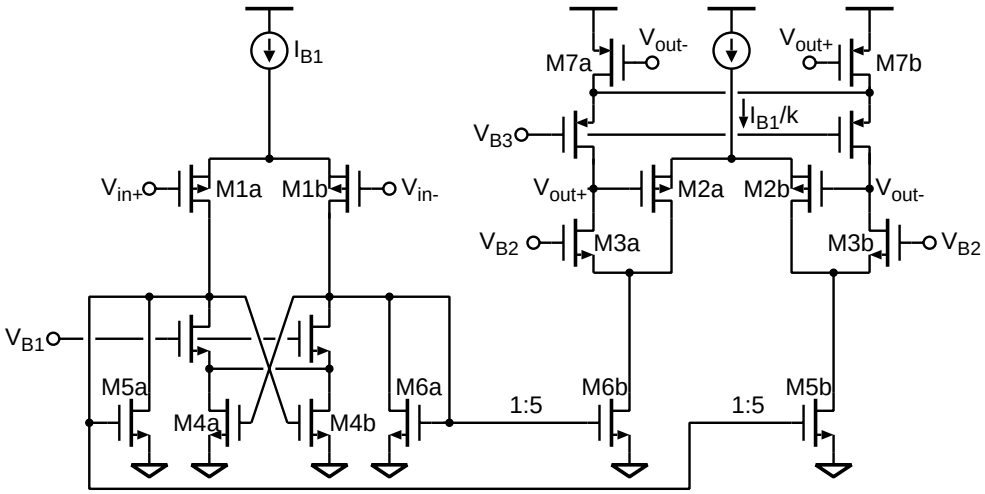


Figure 5.18: Second gain stage schematic.

drops to zero. To counter the charge injection, dummy switches with half the aspect ratio of the actual switch, are included in the design as shown in fig. 5.17(b). As they do not serve as active switches, their drain and source terminals are shorted together. Assuming a very fast clock transition, the channel charge of the active switch will split equally to the drain and source terminals [9]. Using the inverted clock phase for the dummy switches will then cause these to take up the injected charge, hence canceling the charge injection error.

The clock feedthrough which stems from the parasitic gate-drain and gate-source capacitances, is also dealt with by employing the dummy switches. As the dummy switches have their drain and source terminals shorted, the total parasitic gate to drain and source capacitance equals that of the active switch. Hence the feedthrough from the transition waveforms should cancel each other.

5.2.3.2 Output Chopper

The chopper placed at the output is not noise critical as the output signal has gained 40 dB in magnitude. Hence, minimum sized NMOSTs can be used for the output chopper. In order to be able to employ dummy switches, larger than minimum sized switches were employed. Instead NMOST with an aspect ratio of $1.4\,\mu\mathrm{m}/0.35\,\mu\mathrm{m}$ were used.

5.2.4 Second Stage

The second amplifier gain stage is shown in fig. 5.18. For this stage, the noise performance is no longer critical, and the dimensioning is rather determined by the requirements for achieving the desired gain. The gain for the second stage was set to 50.

As the signal swings in this stage are significantly higher than in the first stage, all MOSTs are biased in the strong inversion region.

The M1 input pair tail current was set to $10\,\mu\mathrm{A}$. In order not have very low output bias currents for a gain of 50, i.e. $200\,\mathrm{nA}$, some of the gain of the stage is done in current mode. Specifically, the M5

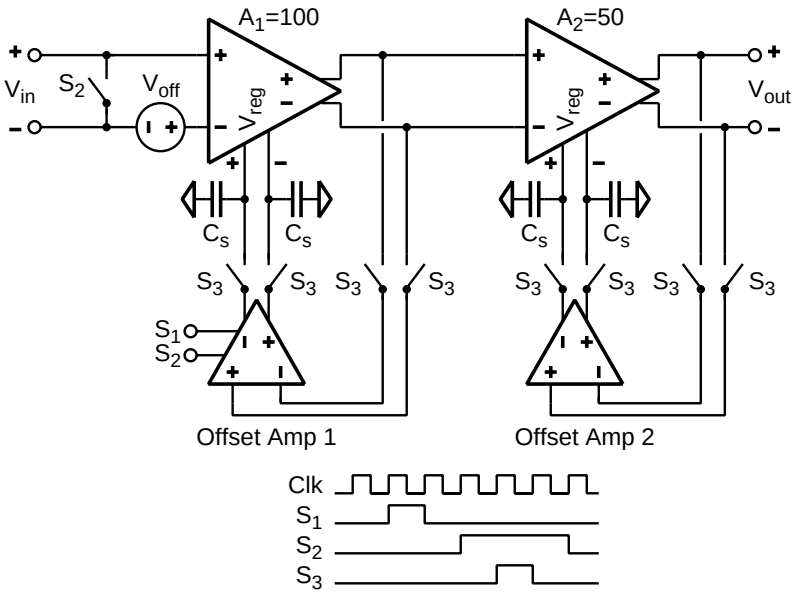


Figure 5.19: Offset compensation scheme.

and M6 current mirror pairs each have a gain of 5. Thereby the M2 output pair tail current could be set to $1\,\mu$A.

Choosing such a scheme, renders a slightly more complicated circuit. As primarily the signal current should be mirrored to the output load pair, the cross-coupled M4 pair takes up most of the input pair bias current. Shorting the M4 pair drains ensures that there will be no differential signal current uptake.

Common-mode stabilization is done in the output stage by cross-coupling the M7 current source pair to the output nodes and shorting their drains. This of course necessitates the use of cascodes for establishing the current source differential output resistance. The M7 pair is then biased in the triode region and hence act as voltage-controlled resistors.

Each of the M3 MOSTs was biased at $2\,\mu$A taking up the bulk of the current mirrored and amplified from the input stage.

An offset nulling port is also included in this stage, but is not shown in the schematic in order not to clutter the schematic unnecessarily. The current partitioning was for this stage set to 10 %.

5.2.5 Offset Compensation Scheme

Each of the amplifier stages included an offset nulling port. These could simply be controlled by negative feedback with a large time constant as was the case for the first amplifier of section 5.1. Such a scheme will typically require large capacitors, which will either take up a large part of the chip area, or be included off-chip. Instead we propose a digitally controlled offset compensation scheme. For this purpose a small state-machine was included for controlling the scheme. A block diagram schematic and timing diagram of the proposed scheme is shown in fig. 5.19. For each of the stages, an auxiliary opamp is included. This scheme requires a special calibration phase where the instrumentation amplifier

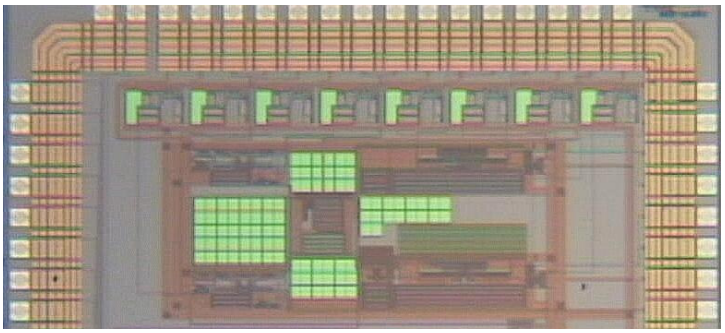


Figure 5.20: Nerve amplifier V2 test chip microphotograph.

is taken out of operation.

During calibration, the input referred offset v_{off}, is measured, amplified and fed back to offset nulling port for each of the stages, in a negative feedback manner. Hereby the branch currents of each of the input pairs are regulated. When the calibration phase ends, the calibration voltage is stored on the capacitors C_{s}, and the offset regulated amplifier is ready for operation. Due to parasitic leakage from the storage capacitors, the calibration has to run at regular intervals, during which the amplifier will not be operational.

The first auxiliary amplifier, offset amp 1, also includes internal offset compensation, utilizing the well-known auto-zero technique [9]. For this purpose an extra signal S1, is included for this opamp. The second stage offset amp does not include this feature.

5.2.6 Measurements

A chip microphotograph of the implemented amplifier is shown in fig. 5.20. The chip contains two different versions of the proposed amplifier.

The employed $0.35\,\mu\mathrm{m}$ CMOS process includes two poly layers, enabling all capacitors to be implemented as poly-poly capacitors. Furthermore, four metal layers are available. The technology employed is an N-well process.

In the following measurements, the chopper frequency was set to 200 kHz if nothing else is mentioned.

Offset performance: The auto-zero scheme employed for the first stage offset opamp appeared to have some problems, leaving us with an offset of approximately 60 mV at the stage 1 output. Instead, an external capacitor of 100 nF was placed at the offset nulling port input, and offset regulation was thus done in a continuous manner, circumventing the switched scheme. No auto-zero scheme was employed in the second stage, and the resulting output offset is shown vs. time is shown in fig. 5.21. Here, the calibration is run every 160 s, and it is seen that the offset is canceled at these intervals. The maximum observed residual output offset was 5 mV, equivalent to an input referred offset below $2\,\mu\mathrm{V}$.

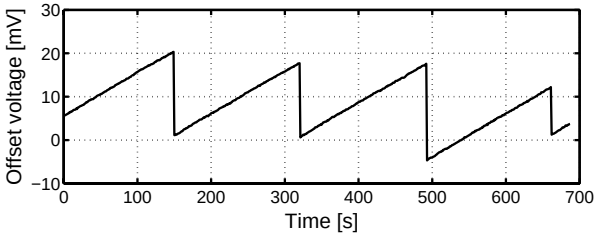
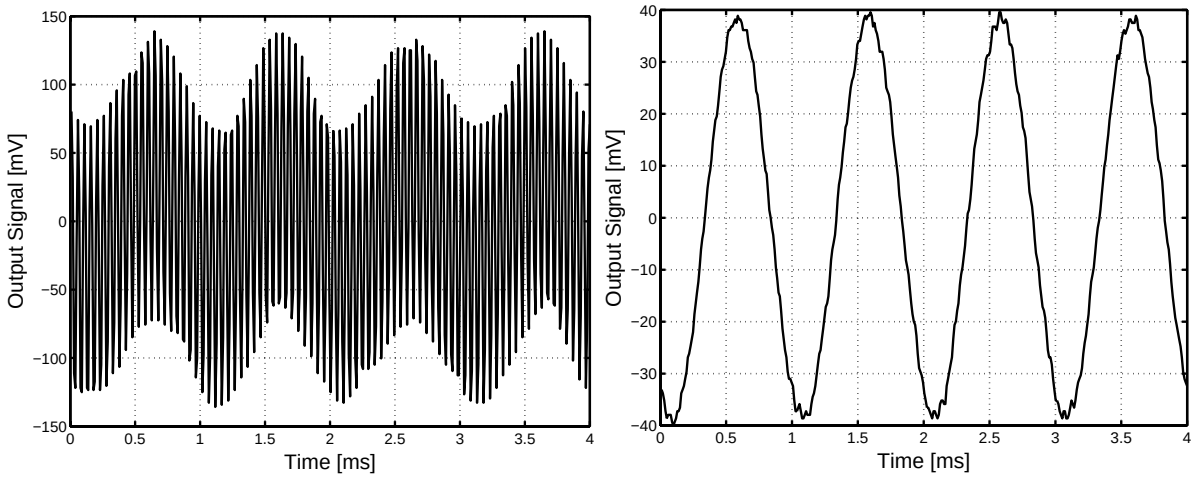


Figure 5.21: Stage 2 output offset vs. time when employing the DT offset compensation scheme.



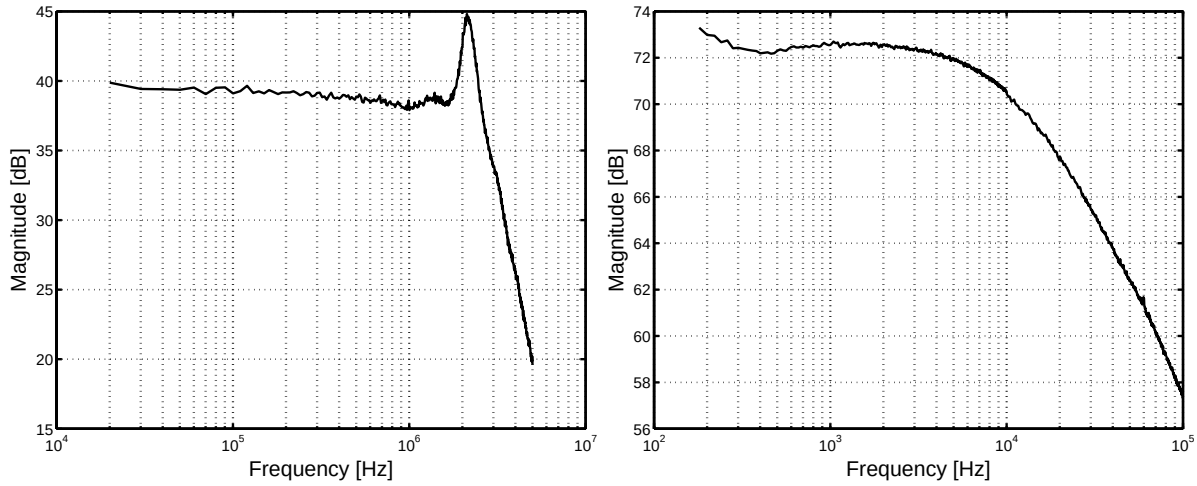
(a) No intermediate filter.

(b) With intermediate filter.

Figure 5.22: Output waveforms for the chopped amplifier, with and without filtering between stages.

In fig. 5.22(a), a typical output waveform is shown. Dynamic offsets induced by the choppers, result in a large signal present at the chopping frequency and it's harmonics. In several of the test chips, this magnitude was sufficient to saturate the second amplifier stage. To eliminate these chopped offsets, a Stanford Research Systems Model SR640 low-pass filter was inserted between the two stages in the test setup. If nothing else is mentioned, this intermediate filter had it's corner frequency set to 99 kHz in the succeeding measurements. A typical output waveform with the intermediate filter is shown in fig. 5.22(b), where it is seen that the chopped offset has been eliminated.

Frequency response: The magnitude frequency response curve for the first amplifier stage is shown in fig. 5.23(a). The curve is the measured average of the ten test chips. A clear peak is observed at 200 MHz, indicating a Q factor slightly higher than expected. The high bandwidth of the stage is due to the high biasing current, which makes the stage well-suited for chopping. The frequency response of the overall amplifier is shown in fig. 5.23(b). A slightly lower gain than the prescribed 74 dB is observed as the average gain is approximately 72.5 dB. The use of matched PMOS pairs provided good matching however, and a gain variation of less than +/- 5 % was observed over the set of test chips. No chopping was employed for these measurements.

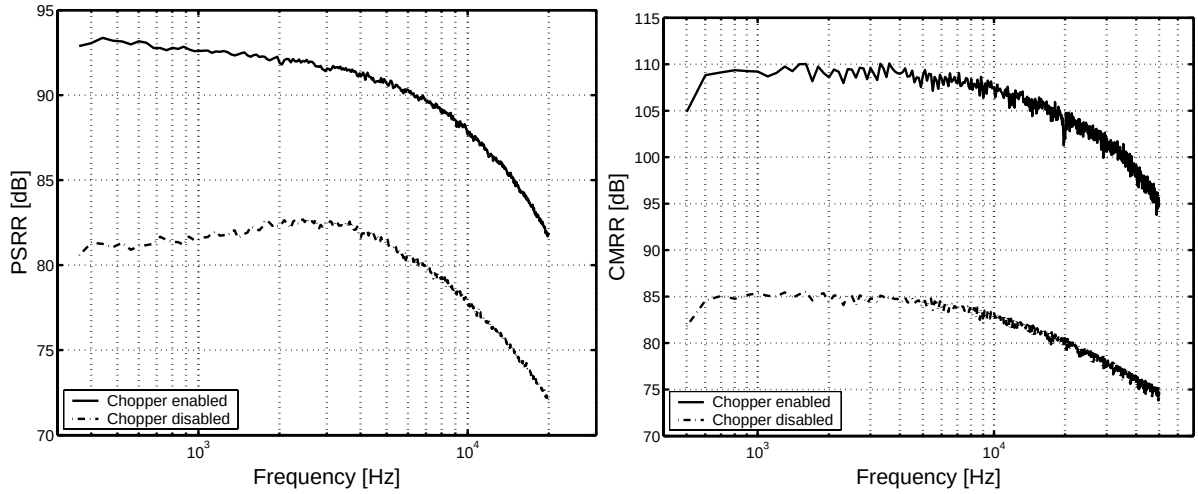


(a) Stage 1 frequency response.

(b) Overall amplifier frequency response.

Figure 5.23: Average frequency magnitude response curves for stage 1 and the overall amplifier.

PSRR and CMRR: The average PSRR and CMRR are shown in fig. 5.24(a) and (b) respectively. The measurements were done both with- and without enabling the choppers. A significant improvement is observable when the choppers are enabled. This is due to the fact, that any signal that will couple through the power supply, or the common-mode input to the differential outputs, will be modulated by the output chopper. Hence these signals will be shifted upwards in frequency to the chopper frequency and it's harmonics, where they can be removed by post filtering.

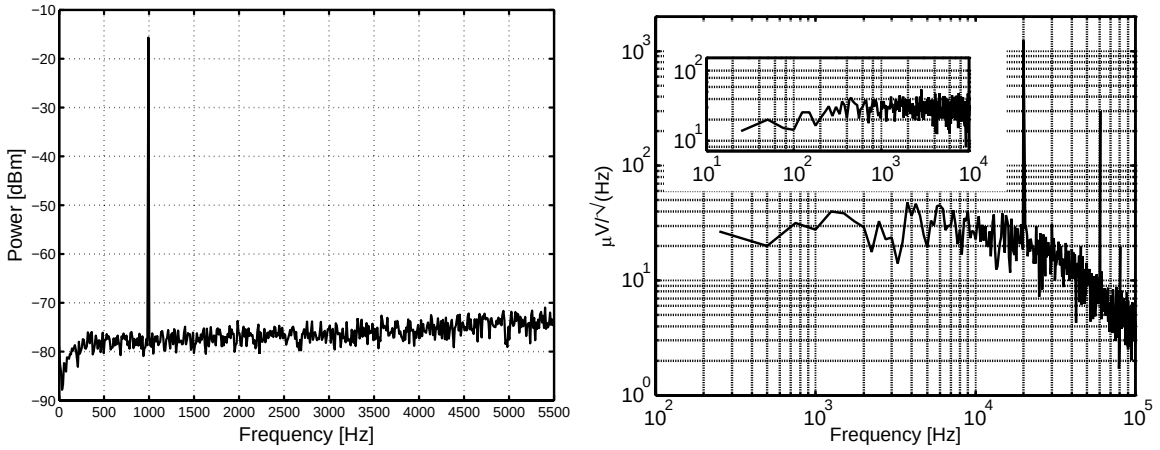


(a) Average PSRR with and without chopping.

(b) Average CMRR with and without chopping.

Figure 5.24: Average power-supply- and common-mode rejection ratios.

Harmonic distortion: Fig. 5.25(a) shows an example output spectrum for the amplifier. The harmonics are not visible as these are buried beneath the noise floor in this spectrum. On average, the total harmonic distortion (THD) in the test chips was dominated solely by the third harmonic, and was observed to be below 0.5 % for all test chips at an applied input amplitude of $40\,\mu\mathrm{V_{pp}}$.



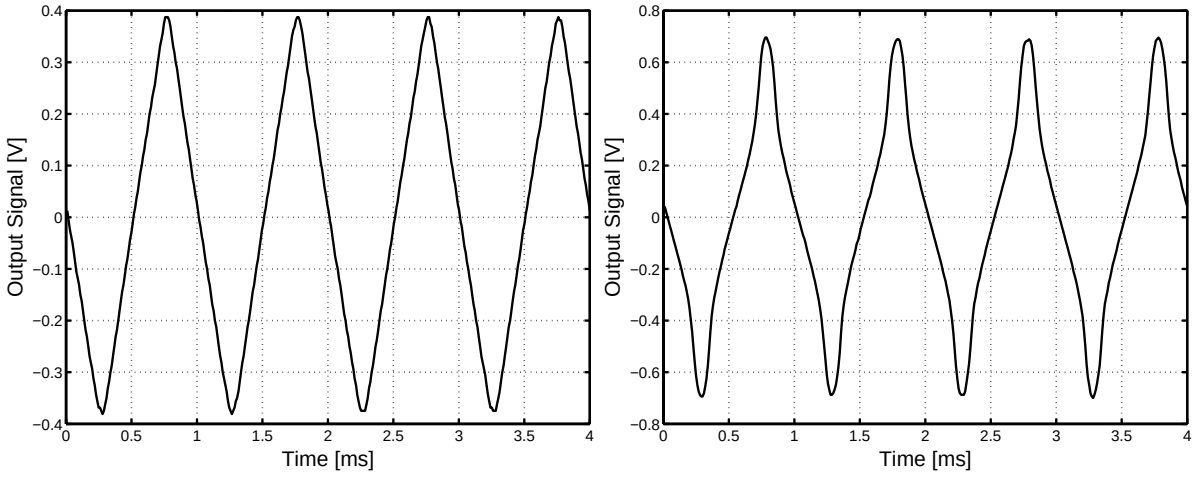
(a) Example output spectrum.

(b) Example output noise spectrum.

Figure 5.25: Example output spectrum and noise spectrum.

Noise: An example output noise spectrum for the chopper amplifier is depicted in fig. 5.25(b). It is noticed that there is no sign of the 1/f-noise at low frequencies due to the chopping. The equivalent input referred noise was found to approximately $7\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$. The chopping frequency was set to 20 kHz in this case.

Slew rate (SR): In fig. 5.26(a) an example waveform showing the amplifier output during slewing is shown. The measured slew rate is, SR=1.8 V/ms. A typical overload output curve is shown in fig. 5.26(b), where the output goes toward the supply rails as the input signal exceeds the maximum input range. These curves indicate that the amplifier will recover promptly from an overload condition.



(a) Amplifier output during slewing.

(b) Amplifier output during an overload.

Figure 5.26: Example output slewing waveform and output overload waveform.

Supply Voltage	$1.8\,\mathrm{V}$
Power Consumption	$242\,\mu\mathrm{W}$
THD(@ f=1 kHz)	$< 0.5\,\%$
CMRR	$> 105\,\mathrm{dB},\ f < 10\,\mathrm{kHz}$
PSRR	$> 85\,\mathrm{dB},\ f < 10\,\mathrm{kHz}$
Gain	72.5 dB
Equivalent input referred noise	$7\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$
Slew rate	$1.8\,\mathrm{V}/\mathrm{ms}$

Table 5.2: Measured amplifier V2 performance.

Power consumption: The current drawn by the circuit was measured to be approximately $134\,\mu\mathrm{A}$, excluding bias circuitry and buffers. For a supply voltage of 1.8 V, this yields a power consumption of approximately $242\,\mu\mathrm{W}$.

5.2.7 Nerve Amplifier V2 Summary

A summary of the measured amplifier performance is given in table 5.2.

In the following a summary of the observed advantages and drawbacks for the nerve signal amplifier V2 is given:

Advantages

- 1/f-noise is effectively removed from the signal band using the chopper modulation technique.
- Single supply operation.
- No need for large on-chip capacitors.
- The circuit is realizable using only MOSTs, i.e. no extra masks are needed for special layers e.g. resistors etc.
- Using moderate inversion input MOSTs in the first stage, significantly reduces the necessary chip area for these.

Drawbacks

- A slightly more complex topology reduces the power efficiency.
- An extra low-pass filter is necessary in between stages.
- Two poly layers are needed for poly-poly capacitors.
- A calibration phase were the amplifier is brought off-line is needed.

5.3 Summary

In this chapter, two different versions of a nerve signal amplifier has been presented. The first amplifier features a gain of 80 dB, whereas the second was designed to have a gain of 74 dB.

The first version utilized a very simple topology, using very few transistors. However it suffered from residual 1/f-noise, though extremely large input transistors were used in the design. Furthermore, using a continuous-time offset compensation scheme necessitated the use of very large capacitors, taking up significant chip area.

The second version uses a slightly more complex topology based on using transimpedance loads for amplification. The use of matching PMOST pairs for input and load resulted in an improved THD. No specific CMFB circuits were needed in this design as common-mode stabilization could be achieved using MOSTs already present in the design.

Both designs operate at a few hundred microwatts, thus having a very low power consumption. The second design features a lower power consumption than the first design due to lower supply voltage.

The use of the chopper modulation technique in the second design resulted in elimination of the low-frequency 1/f-noise, leaving only thermal noise to dominate the signal band noise performance. Dynamic offset introduced by the choppers, showed that intermediate low-pass filtering is necessary for acceptable performance, increasing the power consumption. The complexity of such a filter however could probably remain low, as the signal bandwidth is very limited to the applicable chopper frequency.

To eliminate the need for very large time constants in the offset compensation scheme, as was seen to be necessary in the case of the first amplifier version, a digitally controlled scheme based on periodic calibration was utilized for the second amplifier version. For the first stage, an internal auto-zero scheme for the auxiliary amplifier did not operate correctly giving a high output offset. A run of 200 Monte-Carlo simulations across process variations and mismatch did not reveal any such tendency toward faulty operation of the auto-zero scheme. The discrete-time OCC of the second stage did operate correctly and verified the feasibility of this scheme.

CHAPTER 6

A CONTINUOUS-TIME ΣΔ-MODULATOR

Initially, the ability of ΣΔ-modulators to achieve very high resolution by heavy oversampling, was primarily utilized to implement high quality, audio-range ADCs. Examples of such modulators can be found in [59, 60, 61], where modulators with resolutions of 16 bits and above are reported.

The fast-growing telecommunications market, coupled with the availability of modern sub-micron processes, has since extended the applicability of ΣΔ ADCs to increasingly wide-band signals [62, 63, 64]. Most of these architectures employ multibit quantizers, in order to reduce the necessary OSR, and thereby extend the modulator signal bandwidth. The price paid for using multibit quantizers is that some calibration scheme is often necessary for the feedback DAC, as the DAC linearity needs to match the overall modulator linearity, and since these architectures do *not* enjoy the intrinsic linearity of the 1-bit quantizer.

In recent years, extreme low-power and medium resolution ΣΔ ADCs for biomedical applications have begun to emerge. The low- to medium resolution range has otherwise typically been the domain of other A/D schemes e.g. successive approximation. Two examples of ΣΔ ADCs for pacemaker applications, are reported in [65] and [66]. Cardiac signals have a bandwidth typically less than 150 Hz, and since only a limited ADC resolution is needed (<8 bits), the sampling frequency can be set quite low. This enables ultra-low power consumption, which is below 2 μW in the mentioned examples. The circuit reported in [66] only has simulated results reported though.

In this chapter, the design and measurements of a continuous-time ΣΔ ADC suitable for quantizing preamplified nerve signals is presented. As discussed in section 4.2, employing a CT loopfilter in ΣΔ-modulators holds several advantages over their DT counterparts. Some of the mentioned drawbacks is increased jitter sensitivity and intersymbol interference, both linked to the feedback waveform. Continuous-time loopfilters have primarily been applied in bandpass ΣΔ-modulators for communication purposes, an example of which is given in [67].

First, the general architecture of the ΣΔ ADC is presented, where after the building blocks of the modulator are described. A design methodology for choosing the optimum filter order, feedback filter cutoff frequency and oversampling ratio from a power-efficiency viewpoint, is then presented. Finally, the measurement results of a prototype ΣΔ ADC are presented.

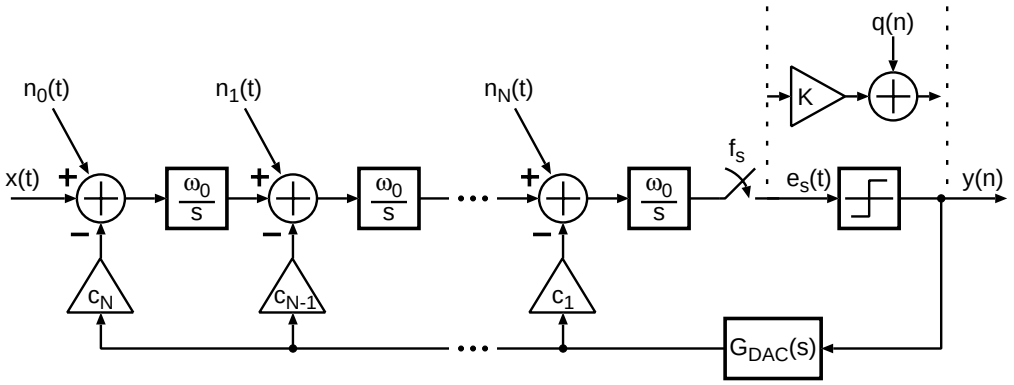


Figure 6.1: CT ΣΔ-modulator of order N with noise injected into the integrator inputs.

6.1 Architecture

The general architecture which is employed for the ΣΔ ADC, is shown in fig. 6.1 for order *N*. Also shown in fig. 6.1, are noise sources $n_0..n_{\mathrm{N}}$, injected into the various nodes of the modulator loop.

For a given noise source $n_{\mathrm{x}}(t)$, the transfer function $NTF_{\mathrm{x}}(s)$, from the noise source to the modulator output can be obtained from the block diagram:

$$NTF_{\mathrm{x}}(s) = \frac{Y(s)}{N_{\mathrm{x}}(s)} = \frac{K\left(\omega_0/s\right)^{N-x}}{1 + KH(s)G_{\mathrm{DAC}}(s)} \tag{6.1}$$

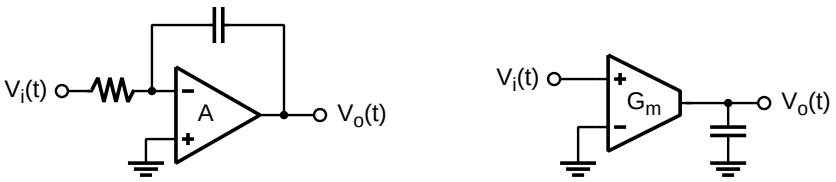
The transfer function for noise source $n_0(t)$ is clearly just the STF, indicating that the first integrator in the loopfilter needs to fulfill all the requirements for the overall ΣΔ-modulator, as any noise components of this block will go unattenuated to the output. All the internal nodes are however noise shaped with a slope proportional to their location in the loopfilter, where ultimately the quantization noise is filtered by the NTF.

For a white noise source with a constant PSD: $S_{\mathrm{x}}(f)$, injected into node *x*, a scale factor k_{x}, can be deduced as the ratio of integrated inband noise seen at the modulator output to the inband noise stemming from the first integrator:

$$k_{\mathrm{x}} = \frac{\int_0^{f_{\mathrm{b}}} STF^2(f)df}{\int_0^{f_{\mathrm{b}}} NTF_{\mathrm{x}}^2(f)df} = \frac{f_{\mathrm{b}}}{\int_0^{f_{\mathrm{b}}} NTF_{\mathrm{x}}^2(f)df} \tag{6.2}$$

The scale factor k_{x}, is thus the ratio which will bring the noise from the internal node noise source n_{x}, on a par with the noise from the first integrator. We see that a higher noise level can be accepted for the inner integrators. As the power consumption in analog circuitry is inversely proportional to the noise level, this indicates that the power consumption of the inner loop integrators can be lowered.

In the following section, the topology of the individual building blocks, i.e. the integrator and the quantizer, will be presented.



(a) Active-RC Integrator (b) G_m-C Integrator

Figure 6.2: Two common CT integrator types.

6.2 Integrator

In fig. 6.2, two traditional integrator types are presented. Each of integrator relies on converting an input signal voltage to an equivalent current signal which is then integrated on a capacitor.

The active-RC integrator relies on an opamp for maintaining a virtual ground node at the negative input, and a resistor for the $V - I$ conversion. The use of resistors can provide a very low THD [68], as these can be manufactured with a very high linearity. However, for integrated circuits production, the use of on-chip resistors requires extra processing, as masks for the manufacture of these are needed.

Another option is the $G_{\mathrm{m}} - C$ integrator. Here, a transconductor is used for $V - I$ conversion, where some linearized internal element is employed. Many transconductors are reported in the literature which employ MOSTs only [69, 70], and thus the $G_{\mathrm{m}} - C$ integrator is more suitable for integrated solutions. Achieving high linearity is more difficult for MOST only solutions, as the MOST is inherently nonlinear. However, depending on the application the higher level of harmonics may be acceptable.

In our case, we have chosen the $G_{\mathrm{m}} - C$ integrator type. The solution proposed is however not MOST only, as poly-poly capacitors are employed. These could potentially be replaced by metal-metal capacitors, or even by gate-oxide to substrate capacitors for a full MOST only implementation. In order to avoid any unnecessary problems due MOST capacitors and their inversion level, the poly-poly capacitor solution was chosen for the prototype.

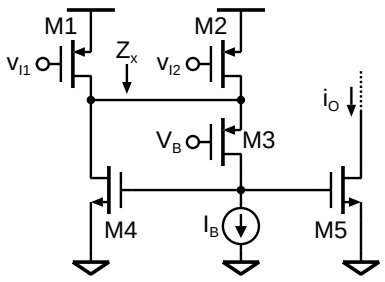


Figure 6.3: Single-ended transconductor.

6.2.1 Transconductor

A principle schematic of the employed two-input transconductor is shown in fig. 6.3. This transconductor is an extended version of the transconductor introduced in [70]. A two-input transconductor is necessary to facilitate both the input signal and the feedback DAC signal.

The M1 and M2 input MOSTs are biased in the triode region, and thus has a linear $V - I$ characteristic given by:

$$I_{\mathrm{D}} = K_{\mathrm{P}}^{'} \left(W/L\right) \left[V_{\mathrm{G}} - V_{\mathrm{th,p}} + \frac{n}{2} \left(V_{\mathrm{D}} - V_{\mathrm{S}}\right)\right] \left(V_{\mathrm{D}} - V_{\mathrm{S}}\right), \, V_{\mathrm{G}} - V_{\mathrm{th,p}} \leq n \left(V_{\mathrm{D}} - V_{\mathrm{S}}\right) \tag{6.3}$$

From eq. (6.3), it is seen that the output current has a linear dependence on the gate voltage, if the drain-source voltage is held constant. The resulting transconductance is:

$$G_{\mathrm{m,single}} = g_{\mathrm{m}} = \frac{\partial I_{\mathrm{D}}}{\partial V_{\mathrm{G}}} = K_{\mathrm{P}}^{'} \left(W/L\right) \left(V_{\mathrm{D}} - V_{\mathrm{S}}\right) \tag{6.4}$$

for each of the input MOSTs, M1 and M2. The ratio between the M1 and M2 transconductances can be set by scaling their aspect ratios accordingly.

The drain-source voltage of M1 and M2 is fixed by M3, which forms a negative feedback loop in conjunction with M4. This negative feedback loop creates a virtual ground node at the M3 source, which is well suited for current summing. By varying the M3 gate bias voltage V_{B}, the drain-source voltage of M1 and M2 can be varied and hence the G_{m}, of the transconductor can be tuned.

As the signal current produced by M1 and M2 has to flow through M4, it is accessible from the M5 drain terminal, as it is mirrored through the M4-M5 pair. The M4-M5 pair also include cascodes for better matching performance, these are however not shown in the schematic for simplicity.

M3 is biased in weak- to moderate inversion to maximize it's transconductance, whereas the M4-M5 current mirror is biased in strong inversion for good matching.

The impedance Z_{x}, seen from each of the transconductor MOSTs M1 and M2, is approximately $1/g_{\mathrm{m4}}g_{\mathrm{m3}}r_{\mathrm{ds3}}$, where the feedback action is seen as the $1/g_{\mathrm{m3}}$ is multiplied by the inverse loop gain. The regulating loop is dominated by the pole at the M4 gate node. An approximate expression for the loop GBW is thus given by:

$$GBW_{\mathrm{loop}} \approx \frac{g_{\mathrm{m4}}}{2\pi \left(2c_{\mathrm{gs4}} + c_{\mathrm{gd3}}\right)} \tag{6.5}$$

So by regulating the bias current I_{B}, the loop speed and thereby the settling time constant can be set.

The proposed transconductor can be extended to a fully differential version, by duplicating the transconductor core for the negative input signal polarity and adding a CMFB circuit. This is shown in fig. 6.4. The CMFB circuit is realized by replicating the core transconductor and using the dual input port for regulating the output voltage. In fig. 6.4, the common-mode signal is sensed by M1C and M2C. Any common-mode voltage variation will result in current being injected into the differential transconductor output nodes through the 2:1 current mirrors, and thus regulate the output voltage to the desired common-mode level.

By large signal analysis, we can deduct the resulting G_{m} for the differential transconductor and the output common-mode level. In the following, consider a large signal input applied to one of the input

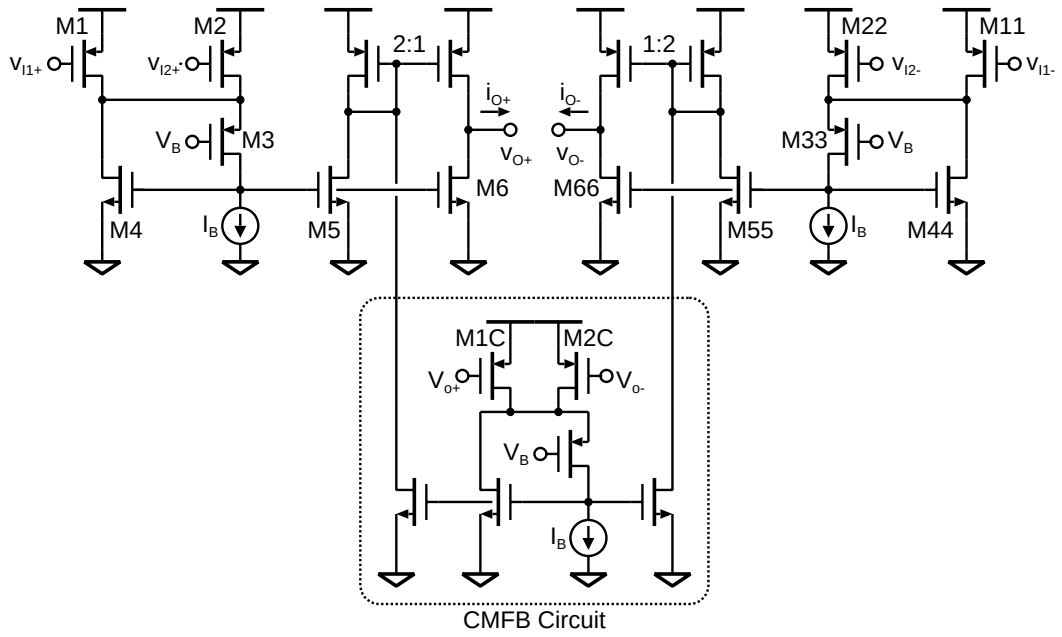


Figure 6.4: Differential transconductor.

ports, M1 or M2 for the positive phase, and M11 or M22 for the negative phase. Let the differential input be given by:

$$v_{\mathrm{I}+} = v_{\mathrm{I,CM}} + 1/2 v_{\mathrm{I,DIF}} \tag{6.6}$$

$$v_{\mathrm{I}-} = v_{\mathrm{I,CM}} - 1/2 v_{\mathrm{I,DIF}} \tag{6.7}$$

where $v_{\mathrm{I,CM}}$ is the input common-mode voltage and $v_{\mathrm{I,DIF}}$ is the input differential voltage. The differential output is equivalently given by:

$$v_{\mathrm{O}+} = v_{\mathrm{O,CM}} + 1/2 v_{\mathrm{O,DIF}} \tag{6.8}$$

$$v_{\mathrm{O}-} = v_{\mathrm{O,CM}} - 1/2 v_{\mathrm{O,DIF}} \tag{6.9}$$

where $v_{\mathrm{O,CM}}$ is the output common-mode voltage and $v_{\mathrm{O,DIF}}$ is the output differential voltage. As M1C and M2C have their drains shorted, the differential output voltage will not cause any current to injected or drawn from the transconductor output nodes. If all of the transconductor MOSTs have the same G_{m}, the resulting differential output current for the given inputs, can be found to be:

$$i_{\mathrm{O}+} = -1/2 \left(G_{\mathrm{m}} V_{\mathrm{I,CM}} + G_{\mathrm{m}} V_{\mathrm{O,CM}} + 1/2 G_{\mathrm{m}} V_{\mathrm{I,DIF}} \right) \tag{6.10}$$

$$i_{\mathrm{O}-} = -1/2 \left(G_{\mathrm{m}} V_{\mathrm{I,CM}} + G_{\mathrm{m}} V_{\mathrm{O,CM}} - 1/2 G_{\mathrm{m}} V_{\mathrm{I,DIF}} \right) \tag{6.11}$$

From the resulting differential output currents, the differential transconductance can be deducted:

$$G_{\mathrm{m,dif}} = \frac{\partial i_{\mathrm{O}}}{\partial v_{\mathrm{I,DIF}}} = 1/2 G_{\mathrm{m,single}} = 1/2 g_{\mathrm{m1}} \tag{6.12}$$

which is seen to be equal to half the single-ended transconductance. From eq. (6.10) and (6.11), can furthermore be seen, that a common-mode current will flow until the input- and output common-mode voltages match each other. The use of a replica transconductor core for the CMFB circuit, thus ensures that the input and output common-mode levels are the same. This enables cascading of the transconductors.

A prerequisite of the previous analysis was that all the transconductor MOSTs had the same G_{m}, i.e. the same aspect ratio. Normally, different aspect ratios will be implemented for the two input ports as a transconductance ratio is desired. The analysis for the CMFB however still holds, as long as the aspect ratio of M1C and M2C are set equal to half the aspect ratio sum of M1 and M2. So we have $W_{1\mathrm{C}} = W_{2\mathrm{C}} = (W_1 + W_2)/2$, if all transconductor MOSTs have the same length.

6.2.2 Noise Analysis

A figure of merit often used for transconductors is the noise excess factor (NEF). The noise excess factor is defined as the ratio between the total output thermal noise conductance over the transconductance [70]:

$$\mathrm{NEF} = \frac{G_{\mathrm{Th,out}}}{G_{\mathrm{m}}} \tag{6.13}$$

The NEF can be used for referring the overall transconductor noise to the input by [71]:

$$v_{\mathrm{ni}}^2(f) = 4kTcG_{\mathrm{m}}^{-1}\mathrm{NEF} \tag{6.14}$$

Consider a transconductor consisting of a single saturated MOST. From eq. (6.14), we see that in this case we have: $\mathrm{NEF} = 1$. For a single MOST working in the triode region, we find:

$$\mathrm{NEF} = \frac{V_{\mathrm{G}} - V_{\mathrm{th}} - nV_{\mathrm{S}}}{V_{\mathrm{D}} - V_{\mathrm{S}}} \tag{6.15}$$

As it is a requirement that $|V_{\mathrm{G}} - V_{\mathrm{th}} - nV_{\mathrm{S}}| > |V_{\mathrm{D}} - V_{\mathrm{S}}|$, for a MOST working in triode, we see that the further we bias a MOST in the triode region, the larger the NEF, i.e. the worse the thermal noise suppression we have. This points toward a trade-off between high linearity, i.e. deep triode region operating point, vs. low NEF, i.e. biasing close to the saturated region.

For the proposed transconductor core, the NEF can be found to be [70]:

$$\mathrm{NEF} = \frac{c_1 g_{\mathrm{ds1}}}{g_{\mathrm{m1}}}\left[1 + \frac{c_2 g_{\mathrm{ds2}}}{c_1 g_{\mathrm{ds1}}} + \frac{c_3 g_{\mathrm{ds1}}}{c_1 g_{\mathrm{m3}}} + \frac{c_4 g_{\mathrm{m4}}}{c_1 g_{\mathrm{ds1}}} + \frac{c_5 g_{\mathrm{m5}}}{c_1 g_{\mathrm{ds1}}} + c_{\mathrm{IB}}\left(1 + \frac{g_{\mathrm{ds1}}}{g_{\mathrm{m3}}}\right)^2 \frac{g_{\mathrm{m,IB}}}{g_{\mathrm{ds1}}}\right] \tag{6.16}$$

The term outside the bracket is the NEF contribution from M1. All the terms inside the bracket are thus to be minimized, for an optimized NEF. The 'c' factors are inversion dependent (see appendix A) and for M1 and M2, $c_1 = c_2 = 1$ as M1 and M2 are biased in the triode region. M3 is biased in weak inversion and $c_3 = 1/2$. As the remaining MOSTs are biased in saturated strong inversion, for those we have $c = 2/3$. M4 and M5 constitute a unity current mirror and thus have the same parameters. The current source I_{B}, is a cascoded MOST in strong inversion. Using these figures and assuming that M2

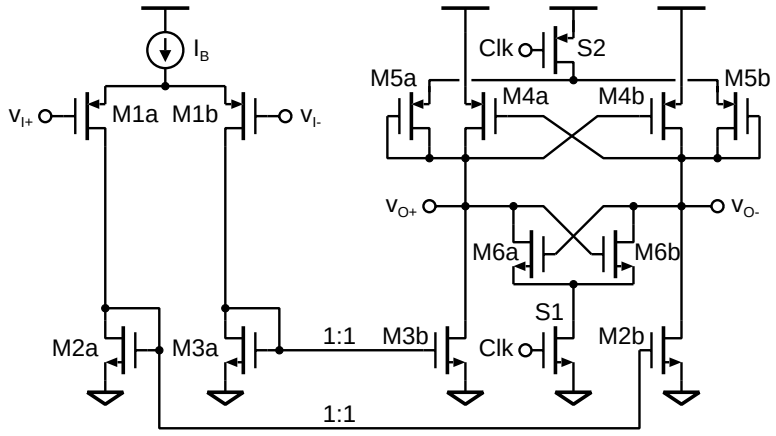


Figure 6.5: Quantizer schematic.

is a replica of M1, eq. (6.16) NEF can be simplified to:

$$\mathrm{NEF} = \frac{g_{\mathrm{ds1}}}{g_{\mathrm{m1}}}\left[2 + \frac{1}{2}\frac{g_{\mathrm{ds1}}}{g_{\mathrm{m3}}} + \frac{4}{3}\frac{c_4 g_{\mathrm{m4}}}{g_{\mathrm{ds1}}} + \frac{2}{3}\left(1 + \frac{g_{\mathrm{ds1}}}{g_{\mathrm{m3}}}\right)^2 \frac{g_{\mathrm{m,IB}}}{g_{\mathrm{ds1}}}\right] \tag{6.17}$$

The resulting inband thermal noise can be found by integrating the input referred noise over the signal bandwidth. For a given input sine with an amplitude *A*, the SNR is found:

$$\mathrm{SNR} = 10\log\left(\frac{A/2}{BW \cdot 4kTcG_{\mathrm{m}}^{-1}\mathrm{NEF}}\right) \tag{6.18}$$

This SNR, which expresses the analog thermal noise level, should always be designed with a good margin to the SQNR, which is expresses the quantization noise level and is chosen at the architecture design level. This however, only holds true for the first integrator, whereas the internal noise shaping relaxes the noise requirements of the remaining integrators.

6.2.3 Quantizer

A track-and-latch comparator was used for the 1-bit quantizer in the modulator. A schematic of the comparator is shown in fig. 6.5. The M1 input pair serves as a preamplifier in conjunction with the M4-M5 pair load. During the 'low' state of the clock, the M6 positive feedback pair is turned off, whereas the diode-coupled M5 pair is on leaving the comparator in track mode. Though the M4 pair constitutes a positive feedback pair, the diode-coupled M5 MOSTs ensures that the loop gain is less than one during track mode. The resulting track gain is:

$$A_{\mathrm{V}} = \frac{v_{\mathrm{o}}}{v_{\mathrm{i}}} = g_{\mathrm{m1}}\frac{g_{\mathrm{m4}} + g_{\mathrm{m5}}}{g_{\mathrm{m5}}^2 - g_{\mathrm{m4}}^2} \approx \frac{g_{\mathrm{m1}}}{g_{\mathrm{m5}}} \tag{6.19}$$

As the clock goes high, the M5 pair effectively 'disappears' and the loop gain around the M4 pair is now unimpeded. For faster latching, the M6 positive feedback pair is switched on as well on the high clock phase.

6.2.4 Capacitor Dimensioning

In section 4.5.2, we found that the total quantization noise could be estimated at the quantizer equilibrium point. Assuming a white noise distribution of the quantization noise, the resulting PSD is given by:

$$S_{\mathrm{q}} = \frac{\sigma_{\mathrm{q}}^2}{f_{\mathrm{s}}} = \frac{1}{A(K_{\mathrm{eq}})f_{\mathrm{s}}} \tag{6.20}$$

where S_{q} is normalized to 1. For a given equilibrium, the NTF is also determined as K_{eq} is given and the inband quantization noise can be calculated. Referring the quantization noise PSD to a noise reference voltage V_{ref}, the resulting inband quantization noise power can be found:

$$P_{\mathrm{q,inband}} = \frac{V_{\mathrm{ref}}^2}{A(K_{\mathrm{eq}})f_{\mathrm{s}}} \int_0^{f_{\mathrm{b}}} \left| \frac{1}{1 + K_{\mathrm{eq}}\widehat{H}\,(f)} \right|^2 df \tag{6.21}$$

where $\widehat{H}(f)$ is the combined feedback filter, i.e. $\widehat{H}(f) = H(f)G_{\mathrm{DAC}}(f)$. For an integrating capacitor C_{int}, connected to the output of a transconductor, the total accumulated thermal noise is:

$$P_{\mathrm{th}} = \frac{\mathrm{NEF} \cdot kT}{C_{\mathrm{int}}} \tag{6.22}$$

where the NEF includes the circuit induced excess noise. The total amount of tolerable noise for an integrator, can be found from the number of bits which is specified, and from the maximum signal energy. In section 4.4, the maximum stable amplitude (MSA) for a given ΣΔ-modulator was introduced. Referring the MSA to the reference voltage, the signal energy is given by: $P_{\mathrm{s}} = (\mathrm{MSA} \cdot V_{\mathrm{ref}})^2$. For a dynamic range equivalent to B bits, the total amount of tolerable noise can be found:

$$P_{\mathrm{N,tot}} = \frac{(\mathrm{MSA} \cdot V_{\mathrm{ref}})^2}{2 \cdot 10^{(6.02B+1.76)/10}} \tag{6.23}$$

By combining equations (6.21-6.23), we can solve for the minimum allowable integrator capacitor size:

$$C_{\mathrm{int}} = \mathrm{NEF} \cdot kT \left[\frac{(\mathrm{MSA} \cdot V_{\mathrm{ref}})^2}{2 \cdot 10^{(6.02b+1.76)/10}} - P_{\mathrm{q,inband}} \right]^{-1} \tag{6.24}$$

Once the necessary capacitor size is found, the integrator transconductance can be derived for the desired integrator unity-gain bandwidth.

6.3 Power Optimization of CT ΣΔ-Modulators

In section 4.5 we saw that the MSA is inversely proportional to the NTF cutoff frequency f_{c}, i.e. the more aggressively we chose to filter the quantization noise, the lower the MSA we could expect. However increasing the NTF cutoff frequency and the resulting increased noise suppression, provided a still growing SQNR even in the light of the decreasing MSA. Hence the maximum SQNR was achieved at a low MSA of approximately 0.3 to the full reference level. Choosing a ΣΔ-modulator with a cutoff

frequency at the very peak of the SQNR curve would be a rather poor choice for the following reasons:

- The maximum of the SQNR curves are seen to have a rather abrupt drop-off for even a slightly increased f_{c}. For analog implementations which may have parameter deviances of up till 20%, this could prove fatal as a sharply deteriorated SQNR would result.

- The low MSA implies high analog power consumption. In order to retain sufficient dynamic range to honor the overall ΣΔ ADC specifications, high biasing currents would be needed to suppress analog noise source when faced with low input signal energy.

These considerations point to a trade-off between analog power consumption, the MSA and the SQNR. In the following, we will examine this trade-off in order to choose the optimum loopfilter order, OSR and NTF cutoff frequency to fulfill a given set of requirements.

6.3.1 Analog Power Estimation

As more noise can be tolerated from the inner integrators, the integrator capacitors can be downscaled. The transfer functions for the noise-sources are within the signal bandwidth seen to be equivalent to differentiators of increasing order. A conservative estimate of the scaling factors, can be achieved by taking the ratio of the integrated STF power over the integrated power of the differentiators of order *n*, for a bandwidth from zero to the differentiator unity-gain frequency f_0. Using this simple approximation, for node *x* in fig. 6.1, we find:

$$k_{\mathrm{x}} = \frac{\int_0^{f_0} \mathrm{STF}^2(f) df}{\int_0^{f_0} \left(\frac{f}{f_0}\right)^{2x} df} = \frac{1}{2x+1} \tag{6.25}$$

which is seen to give a simple scaling along the cascade of integrators given by: $1:3:5:7:\cdots$

Assuming that all the capacitor noise is inband, the minimum sized capacitor for the first integrator can be determined from eq. (6.24). All other capacitors along the cascade of integrators are then given by the simple scaling rule.

For the first integrator, the power consumption is assumed to be set by the necessary GBW for the loop regulating the virtual ground node impedance Z_{x}. We set the required GBW of the loop to $10 f_{\mathrm{clk}}$, for good settling. Once the necessary bias current has been determined, the dimensions of the transconductor MOSTs can be found, as they are biased by the same current and since the size of the integrating capacitor is known.

The subsequent transconductors are subject to the same scaling as the capacitors. This is due to the fact that all integrators have the same unity gain frequency ω_0, so the G_{m} needs to scale by the same amount as the capacitor. As the MOST producing the G_{m}, also carries the transconductor bias current, the bias current is also downscaled by the same amount.

For the quantizer, let P_{c} denote the power consumption per MHz. The quantizer power consumption is assumed to be linearly dependent on the required speed, i.e. the sampling frequency.

Collecting the power consumption of the different modulator components we can deduct an estimate for the analog power for a N^{th} order loopfilter:

$$P_{\text{an}} \approx \underbrace{4V_{\text{ref}} f_{\text{s}} \sum_{\text{n}=1}^{N} C_{\text{int,n}}}_{\texttt{I}} + \underbrace{mV_{\text{DD}} \sum_{\text{n}=1}^{N} I_{\text{B,n}}}_{\texttt{II}} + \underbrace{P_{\text{c}} f_{\text{s}}}_{\texttt{III}}$$

where the contributions are:

`I:` Dynamic power consumption at a sampling frequency f_{s}, and a reference voltage V_{ref}. The summation indicate the capacitor scaling.

`II:` Static power consumption at supply voltage V_{DD}. The proposed transconductor operates in class A, and m designates the number of current branches in the transconductor. The bias current is scaled in the cascade of integrators.

`III:` Quantizer power consumption which is assumed to be linearly dependent on the sampling frequency f_{s}.

6.3.2 Digital Power Estimation

As the output of the modulator is a single-bit oversampled signal, some digital post-processing is necessary to extract the Nyquist rate quantized signal. The first step of the post-processing is a decimation filter which downsamples the signal. A combfilter is typically employed as this type of filter can be implemented using only registers and adders thus minimizing the power consumption. The combfilter transfer function is given by [72]:

$$H_{\text{comb}}(z) = \left(\frac{1 - z^{-D}}{1 - z^{-1}} \right)^{k} = \prod_{n=0}^{\log_2 D - 1} \left(1 + z^{-2^n} \right)^{k} \tag{6.26}$$

where *D* is the decimation factor and *k* is the order of the filter. A decimation filter order of $k = N + 1$ is sufficient for out-of-band quantization noise suppression [73]. The second form of eq. (6.26) can be implemented by a series of FIR filters each decimating by a factor of two as shown in figure 6.6:

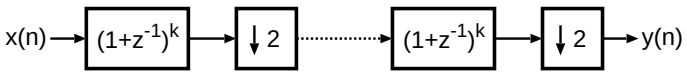


Figure 6.6: FIR2 digital decimation filter.

The word-length of the registers at the output of each FIR stage is equal to $(W_0 + ki)$ bits, where W_0 is the number bits at the filter input. Hence we have a filter configuration where the full sampling frequency is only applied to the first stage, which only has a word-length of a few bits. Whereas the other stages, with increasing word-length, are clocked at decimated frequencies giving a power-efficient decimation filter.

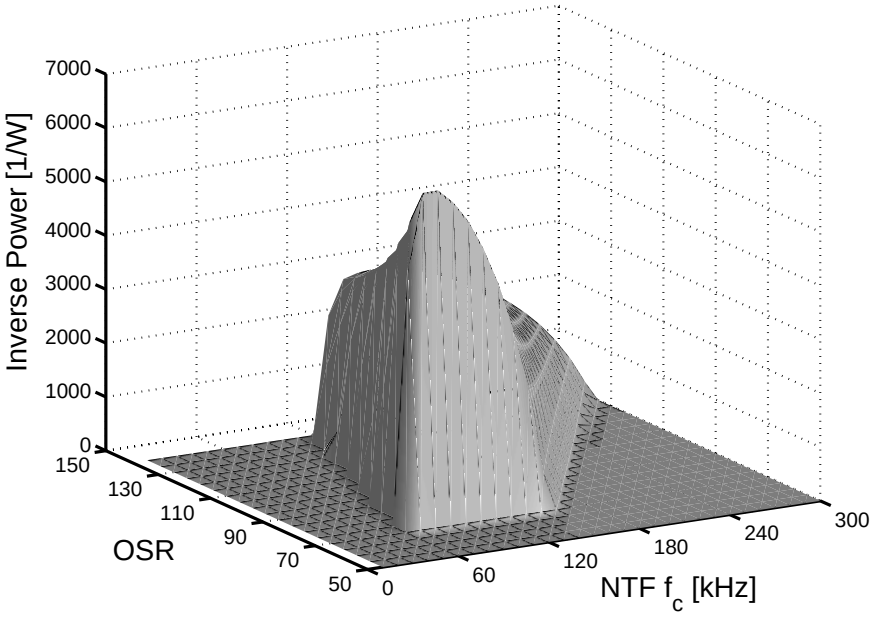


Figure 6.7: Inverse power consumption of a 3^{rd} order CT $\Sigma\Delta$-modulator versus OSR and NTF cutoff frequency f_{c}.

The digital circuitry when using the FIR2 decimation filter structure is confined to D-flip flops, half- and full-adders. For a sampling frequency f_{s}, we can estimate the digital power consumption:

$$P_{\mathrm{dig}} \approx \alpha \sum_{i=1}^{\log_2 D-1} \frac{f_{\mathrm{s}}}{2^{i-1}} \left(P_{\mathrm{fa}} + P_{\mathrm{ha}} + P_{\mathrm{dff}}\right) \tag{6.27}$$

where α is the activity factor. The frequency dependent power consumption of the digital blocks can be found in the datasheets of the employed technology if standard cells are employed for the implementation.

6.3.3 Choosing Loopfilter Order, OSR and Cutoff Frequency

We have thus far obtained estimates of the analog and the digital power consumption of an interpolative $\Sigma\Delta$-modulator of order *N*. The total modulator power consumption is thus given by:

$$P_{\mathrm{tot}} = P_{\mathrm{an}} + P_{\mathrm{dig}} \tag{6.28}$$

An example surface depicting the inverse power consumption of a third order modulator versus the OSR and the NTF cutoff frequency is shown in fig. 6.7. The OSR here, is relative to a signal bandwidth of 10 kHz. The plotted surface reveals a clear peak, indicating that an optimum point with regards to power efficiency exists. Hence for a set of given specifications, an optimum sampling frequency f_{s}, NTF cutoff frequency f_{c}, and OSR can be extracted for the given modulator order. The solution

space where the surface is zero, is where the parameter set resulted in either an unstable modulator or a modulator which does not fulfill the given requirements.

Some properties can be observed from the power surface. At low OSRs, no modulator for any NTF cut-off frequency will be able to fulfill the specifications. This results in sharp rise for the surface as the specifications are met when the sampling frequency is increased. The surface thereafter slopes downward for further increased sampling frequency as the power consumption increases.

The first stage transconductor which results from this optimization process, should be crosschecked to ensure that it fulfills the analog noise specifications, i.e. the NEF should be extracted, and the resulting SNR due to thermal noise should remain well below the SQNR.

From chapter 5, we have that the input signal to the ADC will be in the range of $100 - 200\,\mathrm{mV_{pp}}$, over a bandwidth of 4-10 kHz. A conservative specification of the required ADC could be to set the number of bits B, to 10, and set the reference voltage $V_{\mathrm{ref}} = 200\,\mathrm{mV}$. Producing the inverse power surface over a range of sampling- and cutoff frequencies, for loopfilter orders 2 through 5 gives us the projected minimum power consumptions, summarized in table 6.1:

N	2	3	4	5
P $[\mu\mathrm{W}]$	308	164	195	257

Table 6.1: Projected power consumption of CT ΣΔ-modulators for orders 2 through 5.

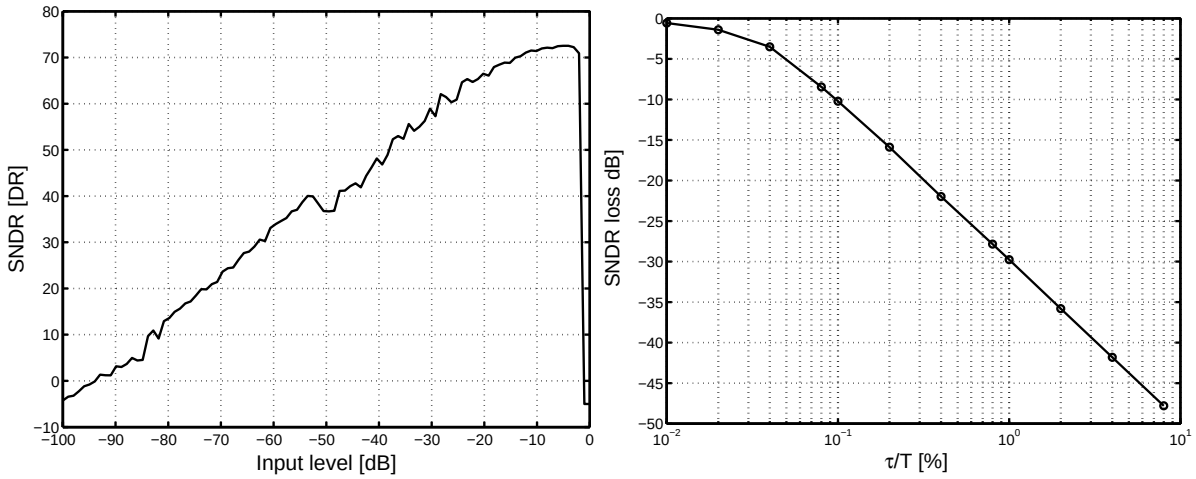
The second order loopfilter is limited by necessitating a fairly high sampling frequency, whereas for higher orders, the low MSA is the main constraint. Hence, a third order loopfilter is chosen, the cutoff frequency was set to $80\,\mathrm{kHz}$ and the sampling frequency to $1.4\,\mathrm{MHz}$, equivalent to an OSR of 70 for a bandwidth of $10\,\mathrm{kHz}$.

6.4 Simulations

In order to verify the CT ΣΔ-modulator performance prior to implementation, a series of simulations were performed using a MATLAB simulink model for ideal behavioral simulation. In fig. 6.8 (a), the SNDR curve resulting from applying a sweep of 100 input amplitudes over the entire input range is shown. A peak value SNDR of approximately 72 dB results at a maximum input amplitude of 0.71. After the MSA is surpassed, the SNDR is seen to drop abruptly as the modulator goes unstable for these simulations.

The MSA projected by the Gaussian stability criterion was 0.72, showing good correspondence with the simulated MSA value. The maximum SQNR estimated by eq. (4.45) however, overestimated the SQNR by approximately 10 dB giving a value of 82.5 dB for the given modulator configuration. This estimation is still much better than the estimation obtained from the simple estimation formula given in eq. (4.38), which projects a maximum SQNR of 109.5 dB for this configuration.

To investigate the effect of intersymbol interference, slew-rate limiting was induced in the feedback waveform as shown in fig. 6.9, where T is the sampling period.



(a) SNDR for the ideal modulator.

(b) Intersymbol interference performance loss.

Figure 6.8: Simulated SNDR over the $\Sigma\Delta$-modular input range and sensitivity to intersymbol interference.

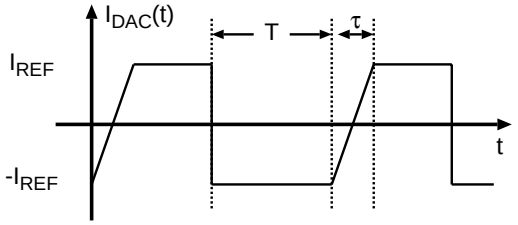


Figure 6.9: Non-equal rise and fall time in DAC feedback waveform.

A sweep of values for the τ/T-ratio was done, resulting in the curve shown in fig. 6.8 (b). This curve shows that the modulator is highly sensitive toward any intersymbol interference, indicating that τ/T-ratios smaller than 0.03 % should be guaranteed for a performance degradation of less than 3 dB.

To get an estimate for the maximum acceptable clock jitter, eq. (6.29) is used, which is replicated below for convenience:

$$\mathrm{SNR}_{\mathrm{NRZ}} = 10\log\left(\frac{\mathrm{OSR}\cdot A^2}{2\sigma^2_{\Delta_\mathrm{y}}\left(\sigma_j/T\right)^2}\right) \tag{6.29}$$

For 10 bits performance, we have $\mathrm{SNR}_{\mathrm{NRZ}} = 62\,\mathrm{dB}$. The variance of the output difference signal: $\sigma^2_{\Delta_\mathrm{y}} = y(n) - y(n-1)$, was found by simulation to be approximately 2.1. Thus for a clock frequency of 1.4 MHz, i.e. OSR=70, and an MSA of 0.72 we find that rms jitter should remain below 1.7 ns.

For achieving less than 3 dB performance degradation due to finite integrator gain, eq. (4.66) gives us that the gain should exceed 21 dB, which is easily fulfilled. The 3 dB SNDR degradation due to integrator leakage is verified by simulation.

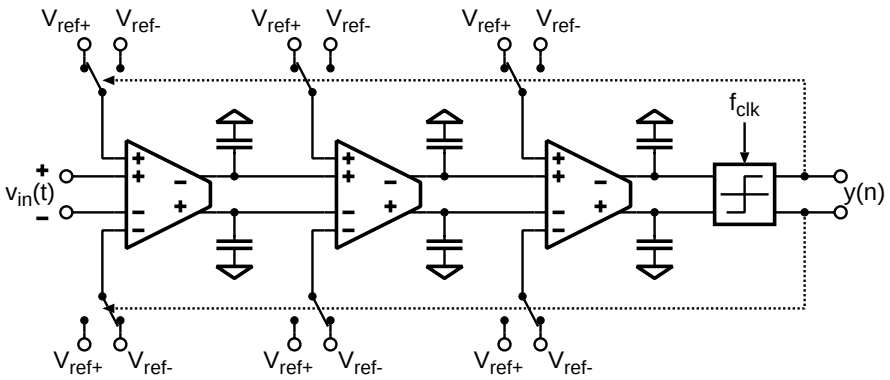


Figure 6.10: Differential third order CT ΣΔ-modulator employing $G_{\mathrm{m}} - C$ integrators.

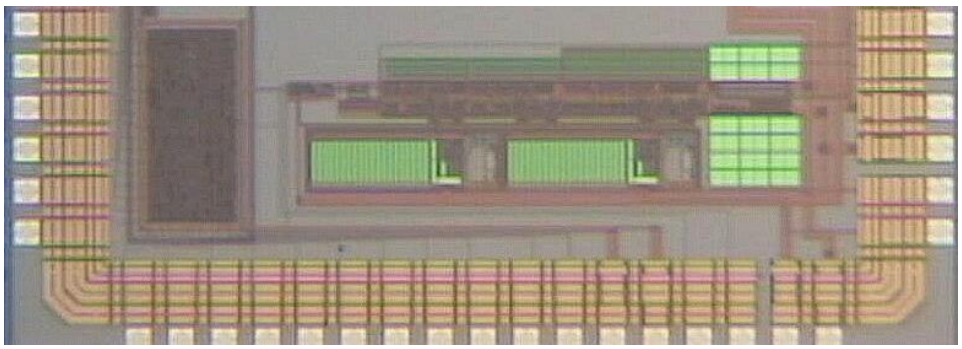


Figure 6.11: Chip microphotograph of the implemented ΣΔ-modulator.

6.5 Implementation

All the integrators are implemented using the fully differential transconductor of fig. 6.4. A block diagram of the resulting third order CT ΣΔ-modulator is shown in fig. 6.10. A couple of standard library D-flip-flops, which are not shown in the block diagram, captures the comparator output, prior to feeding it back to the switches controlling the feedback voltage. All feedback switches are implemented as NMOSTs, which are scaled according to the integrator feedback G_{m} scaling. A slightly delayed version of the clock is used for the D-flip-flops to ensure that the comparator output has settled before capture. The slight delay was set to $25\,\mathrm{ns}$, and is implemented using standard library delay cells. The reference voltages are generated off-chip and buffered using standard cell opamps on-chip. A chip microphotograph of the implemented third order CT ΣΔ-modulator is shown in fig. 6.11. The implementation includes a digital decimation filter, which is the square block in the left of the photo, which remains to be tested at the time of writing. The technology chosen for the implementation was the AMS $0.35\,\mu\mathrm{m}$ N-well, CMOS process, which features poly-poly capacitors and four metal layers.

The integrating capacitors were created using poly-poly capacitors. The different feedback coefficients were synthesized by scaling the aspect ratio of the two transconductor input MOSTs. A summary of key parameters for the implemented integrators is given in table 6.2.

The preamplifier stage in the comparator was biased with a tail current of $2\,\mu\mathrm{A}$. The output latch

	Int1	Int2	Int3
Reg. loop GBW [MHz]	10	4	2.8
Integrator DC gain [dB]	70	70	58
Branch bias current I_{B} [μA]	3	1.25	0.75
Capacitor size [pF]	7.5	1.75	1.25
Power consumption [μW]	64.8	27	20.7
NEF	12	16.5	16.3

Table 6.2: Integrators summary.

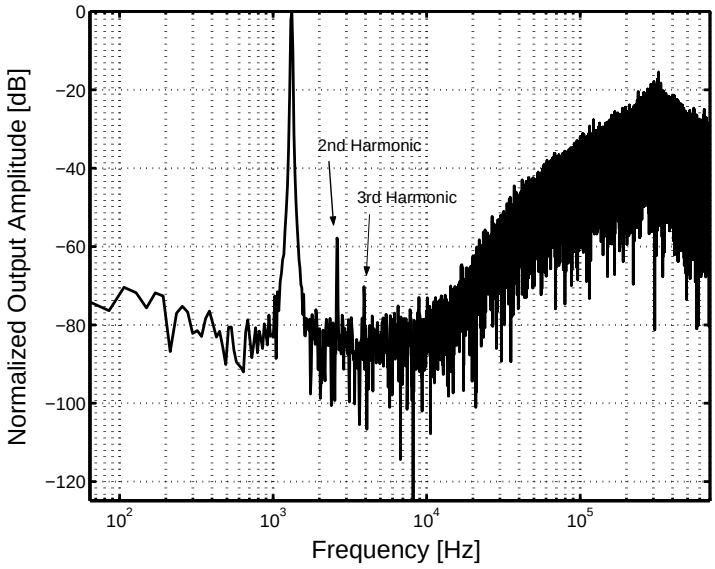


Figure 6.12: Typical $\Sigma\Delta$-modulator output spectrum.

stage of the comparator was biased at $V_{\mathrm{DD}} = 3.3\,\mathrm{V}$, in order to bring the comparator output level to the native full digital levels of the technology. A slightly higher comparator power consumption of $10.2\,\mu\mathrm{W}$ results.

From the power consumption reported, it is noticed that in excess of 50 % of the power consumption stems from the first integrator.

6.6 Measurements

The digital output stream from the CT $\Sigma\Delta$-modulator was captured using a Hewlett-Packard digital oscilloscope, model HP54645D. For clock generation an HP33120 signal generator was used. Each of the reference voltages were generated on the test board.

A typical output spectrum from the modulator is shown in fig. 6.12. The harmonic performance is seen to be dominated by the second harmonic in the spectrum. As the circuit is fully differential, it would be expected that the third order harmonic would dominate. This dominance of the second harmonic was seen across the entire set of test chips. At the time of writing, the cause of the second

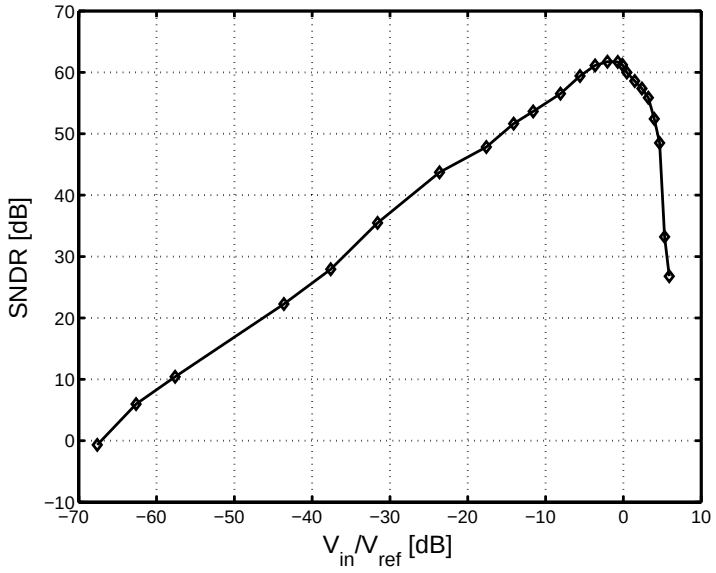


Figure 6.13: Typical SNDR curve for the ΣΔ-modulator.

harmonic is unknown. A possible cause for the second harmonic could be mismatches in the input transconductor MOSTs. As these MOSTs for each polarity of the input can *not* be laid out in finger-based common-centroid fashion, but rather are laid out separately with close proximity to each other, they are more prone to mismatch. It can be noticed that the same dominance is reported in [70].

The initial specification for the ΣΔ-modulator was to achieve a SNDR equivalent to 10 bits, i.e. 62 dB, over a 10 kHz bandwidth. The average maximum SNDR achieved over this bandwidth was 57.7 dB for an 340 Hz input tone and 55.4 dB for an 1860 Hz input tone. By reducing the bandwidth to the nerve signal bandwidth, i.e. 400 Hz - 4 kHz, the average maximum SNDR rose to 61.2 dB and 58.4 dB for the 340 Hz tone and the 1860 Hz tone respectively. The applied reference voltages were set at +/- 70 mV, and the input signal magnitude was set to $120\,\mathrm{mV_{pp}}$.

The measured SNDR over a range of input amplitudes is shown in fig. 6.13 for one of the test chips. The applied test tone frequency for this curve was 340 Hz and the SNDR was evaluated over a frequency range of 400 Hz to 4kHz. The SNDR slope attains it's maximum in the range [-3; -1] dB from full scale, whereas the calculated MSA was at 0.72, i.e. -2.85 dB. The input dynamic range is found to be approximately 66 dB. The power consumption of the ADC was measured to be $108\,\mu$W excluding bias circuitry and input/output buffer circuitry.

6.7 Summary

The subject of this chapter was the implementation of a continuous-time ΣΔ-modulator suitable for quantization of pre-amplified nerve signals. Using the performance estimation methods derived in chapter 6 and a few simple scaling rules for the internal sizing of the integrator capacitors and transconductors, a method for finding the most power-efficient configuration of a CT ΣΔ-modulator is shown.

As we have specified a desired performance in terms of SNDR of 10 bits, a third order modulator

is found to be suitable for our needs. Although 10 bits is a high specification for the modulator input signal, this medium resolution also serves to show the ability of $\Sigma\Delta$-modulators to achieve medium-to-high resolution without harsh requirements imposed on the analog building blocks.

From the simulated behavior, high sensitivity toward the feedback waveform from the quantizer output is found. Indeed, the need for a very low jitter clock could prove to be the bottleneck in the implementation of a high resolution system using a NRZ feedback waveform. The effect of intersymbol interference can be eliminated by the use of a RZ feedback waveform and the jitter sensitivity can be alleviated by changing the feedback waveform [74], however potentially at the cost of a higher integrator power consumption.

From the measurements of the implemented circuit, we see that 10 bits SNDR performance is not quite achieved over the initially specified 10 kHz bandwidth, but only by evaluating the noise and distortion over the nerve signal bandwidth itself do we achieve this performance. The main limitation is the second harmonic which is clearly dominating the harmonic performance though the designed integrator circuit is fully differential. The observed noise floor is also clearly flattened at inband frequencies, yielding a poorer SNR. As a resolution of less than 8 bits would probably suffice in the quantization of nerve signals, the required 10 bits in the specification gives us a good margin for allowing non-ideal effects.

Low power consumption is achieved in the prototype ADC as only $108\,\mu\mathrm{W}$ is consumed. This was however without on-chip voltage references which would need to be generated and buffered for a fully integrated solution. Also, unless a high-quality clock can be guaranteed to be supplied from the outside world, a highly jitter-free on-chip clock generator would need to be implemented, which may prove to be a non-trivial task to design and increase power consumption significantly.

CHAPTER 7

A CMOS CURRENT STEERING DAC

The subject of this chapter is the presentation of an implemented DAC. For the project at hand, the focus was aimed at implementing a generic high-resolution, high-speed DAC. The specifications of this work is to obtain 14 bits of static accuracy, and to obtain a high SFDR at a maximum update rate of 200 MS/s.

Such a design is not directly applicable to implantable circuits, as extremely high update rates are seldom needed and since the power consumption of such a high-speed device is not compatible with telemetry-powered units. However, the current steering principle is easily downscaled to fulfill other less stringent requirements, with resulting lower power consumption. The current steering principle is well suited for driving cuff electrodes as these require pulsed currents. Hence a specific driver can be omitted in an actuator design, when a current steering DAC is employed for the D/A conversion of the stimulus pulse control signal [7].

For a greater number of channels in stimulators, higher DAC operating speeds could be justified. Rather than implementing large arrays of DACs, a single high-speed DAC could be utilized for driving the stimulators in a multiplexed manner. For such a DAC, spectral purity is essential, i.e. a high SFDR is required. An example of an implantable multichannel stimulator for muscle control is given in [8]. Here, up to 32 channels need to be driven with 8-bit accuracy at a $1\,\mu\mathrm{s}$ resolution. A neural stimulator capable of 64 channel operation is presented in [75]. In this design, micro-machined silicon probes are used for neural stimulation, where one of the applicable areas is cochlear stimulation. Furthermore, retinal prosthesis could be mentioned, as these types of implants require a very high number of stimulators for sufficient resolution.

7.1 DAC Architecture

A schematic of the overall DAC architecture is shown in fig. 7.1. The DAC is three-way segmented, where the 5 MSB bits are implemented by thermometer decoded trimmed current sources to achieve the overall 14 bit accuracy. The remaining 9 LSB bits are split two-way into yet another thermometer decoded DAC, which implements the upper 4 LSBs, whereas the remaining 5 LSBs are implemented in a binary array.

The employed technology is the ST $0.18\,\mu\mathrm{m}$ CMOS process. The chosen process has a limited supply voltage of 1.8 V, hence feeding the output resistors directly from the DAC current sources is not feasible. Instead, a current folding output stage is used for ensuring sufficient headroom for the

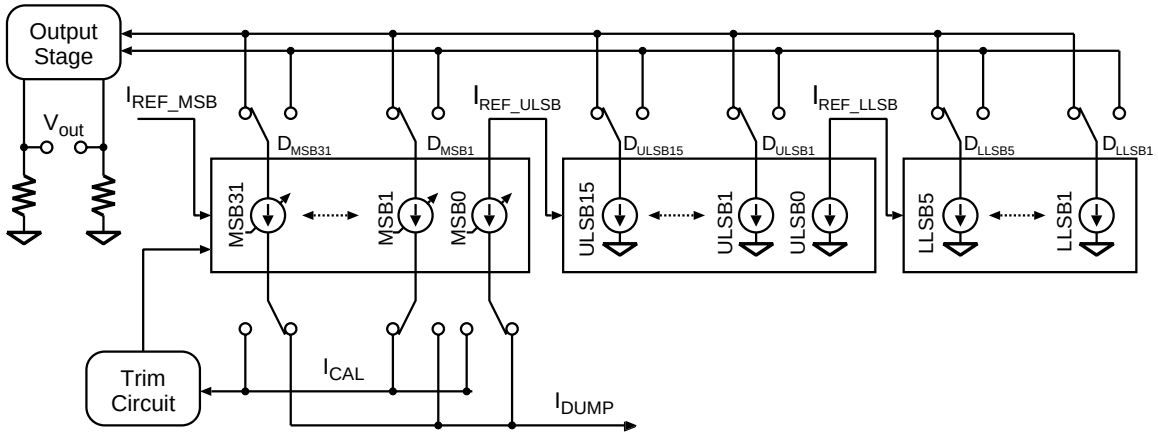


Figure 7.1: 14 bit DAC architecture.

desired voltage swing. Furthermore, the output stage provides a virtual ground node suitable for current summing at the DAC current source output terminals. The requirements for the output impedance of the individual current cells can be relaxed as the current cell output voltage is fixed by the output stage.

Achieving 9 bits of intrinsic accuracy is within the capability of the technology without requiring excessively large MOSTs, which is the reason for the DAC partitioning into the calibrated current sources for the MSBs. Each of the MSBs carry a nominal current of $500\,\mu\mathrm{A}$. The accuracy of the MSB current sources is ensured by sequentially calibrating each of the sources. This is done fully in the background, as the MSB sources are implemented as floating sources and thus have their tail currents available for the calibration circuit without affecting the overall DAC operation. When a MSB source is not being calibrated, the tail current is dumped to a current sink. To ensure monotonicity, the last of the calibrated sources in the array, MSB0, is used as a reference for LSB DAC.

The LSB DAC is implemented by having the coarse upper 4 LSBs (ULSBs) implemented as a segmented DAC in order to reduce static nonlinearity. Each of the ULSB current sources carry 1/16 of a MSB cell current, i.e. $31.25\,\mu\mathrm{A}$. The last of the matched sources in the ULSB array serves as a reference for the remaining 5 lower LSBs (LLSBs). The LLSBs are binary decoded and implemented by having a number of unit current cells in parallel for each bit to ensure good matching.

The DAC full scale output current is 16 mA, which leads to an LSB current given by: $I_{\mathrm{LSB}} = 16\,\mathrm{mA}/2^{14} \approx 977\,\mathrm{nA}$. Each of the DAC current outputs is loaded by an external $25\,\Omega$ resistor to produce the output voltage. The full output voltage swing is thus 400 mV.

7.2 Circuit Description

The implementation of a trimmable floating current source proved to be the most complex task of the overall DAC design. The majority of this section is therefore dedicated to the description of the MSB array and the necessary auxiliary circuitry.

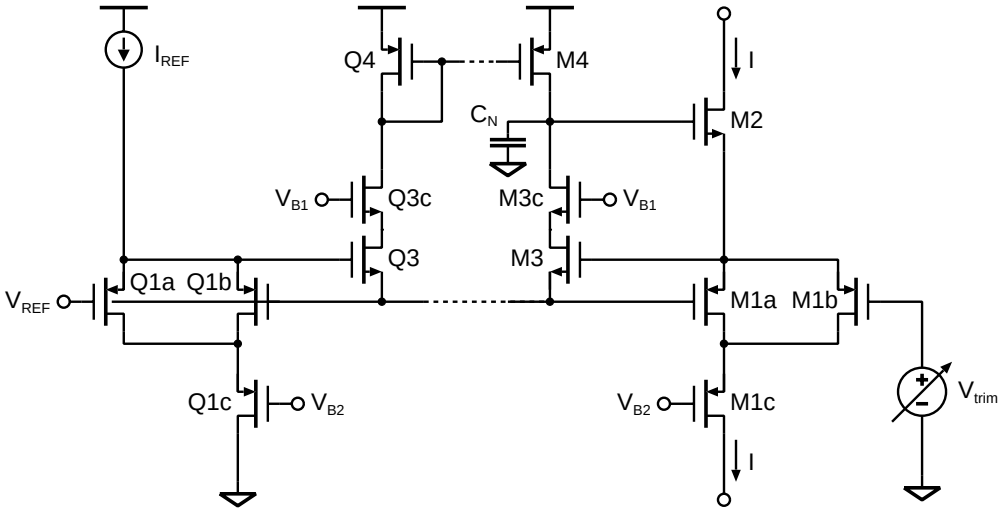


Figure 7.2: Floating current source.

7.2.1 MSB Cell

The schematic of the floating current source is shown in fig. 7.2. Transistors denoted by 'Q' are used for biasing, whereas the actual current source transistors are denoted by 'M'. For a given reference current I_{ref}, a voltage is developed relative to the biasing voltage V_{ref}: $|V_{\mathrm{GS,Q1}}| = |V_{\mathrm{GS,Q3}}|$, which yields a secondary current $I_{\mathrm{DS,Q3}}$ which is mirrored by Q4-M4. The negative feedback loop consisting of M2-M3 hereby establishes $V_{\mathrm{GS,M3}}$. For equal dimensions we then have that $|V_{\mathrm{GS,Q1a}}| = |V_{\mathrm{GS,M3a}}|$, yielding the same current flowing in transistors Q1a and M1a.

The trimming feature of the current source is achieved by making the V_{GS} of transistor M1b trimmable. By controlling the gate voltage V_{trim}, we can trim a given fraction of the total output current to the desired value. The current partitioning between M1a and M1b is determined by matching performance and allowable swing for the trim voltage. For a trim voltage swing of +/- 100 mV, letting 16 % of the MSB current flow in M1b was found to be sufficient, i.e. the M1b MOST carries a nominal bias current of $80\,\mu\mathrm{A}$.

Besides from establishing $V_{\mathrm{GS,M3}}$, M2 furthermore serves to boost the top output impedance of the current cell. Acting as a regulated cascode, M2 not only offers an output resistance of r_{ds2}, but is boosted by the gain of the loop consisting of M1a, M3, M3c, M4 and M2. An approximate expression for the top DC output impedance can be shown to be:

$$Z_{\mathrm{top}} = r_{\mathrm{ds2}} \left(1 + \frac{g_{\mathrm{m2}} g_{\mathrm{m3}} r_{\mathrm{ds4}}}{g_{\mathrm{m1}}} \right) \tag{7.1}$$

which attains a simulated value in excess of $1\,\mathrm{M}\Omega$ over all process corners. The high DC output impedance ensures the static performance to be within the specifications. A trade-off between high output impedance and low noise performance exists for higher frequencies. For noise reduction purposes, the C_{N} capacitor was inserted in the MSB cell regulating loop, thus reducing the noise bandwidth. As a result, the MSB top output impedance will exhibit roll-off at higher frequencies as the dominant pole

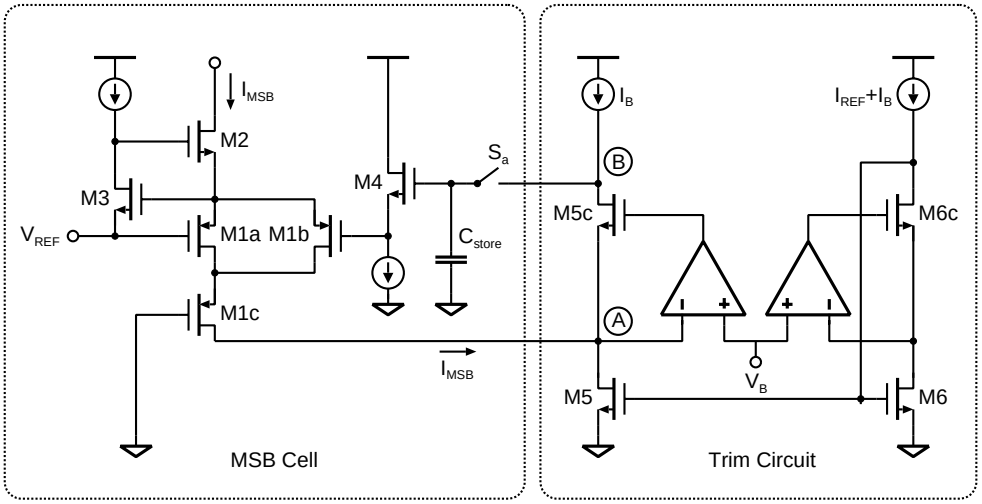


Figure 7.3: Analog trimming loop employed for MSB cell calibration.

in the loop is encountered. This dominant pole is approximately given by:

$$f_{\mathrm{p}} = \frac{1}{2\pi C_{\mathrm{N}} r_{\mathrm{ds4}}} \tag{7.2}$$

As low-noise performance was desired, the C_{N} capacitor was set to a value of 10 pF, resulting in a simulated pole frequency of 100 kHz. Thus for high DAC update rates, higher sensitivity to the switch transients should be expected due to the limited MSB cell recovery time.

The floating current source tail node is used by the trim circuit which operates at low speeds compared to the overall DAC update rate. By cascoding the current source pair, a high tail output impedance is guaranteed, ensuring low sensitivity to the continuous background calibration process.

7.2.2 Trim circuit

The concept of self-calibration in the DAC MSB current sources is introduced in [22], where a non-floating current source is calibrated by diode-coupling the current source MOST to a reference current, and the MOST gate-source capacitance is used for holding the charge during operation. As the trim voltage holding capacitance C_{hold}, is independent in our proposed floating current cell, calibration by diode-coupling is not possible. In [25], a $\Sigma\Delta$-modulator is used for quantizing the MSB cell output current value. The trimming value is then computed in the digital domain, where-after a trim voltage is generated using a separate DAC dedicated for this task. As such a scheme seems unnecessarily complicated, we propose an analog-only trim circuit which is shown in fig. 7.3. The trim circuit is continuously applied to the individual MSB cells through a switching scheme. In order to retain the trim voltage for the individual cell, the storage capacitor C_{store} is included in the MSB cell. The employed opamps are identical folded cascode type OTAs.

The $I_{\mathrm{REF}} + I_{\mathrm{B}}$ current flowing in the right branch of the trim circuit is mirrored by M5-M6. At node A, the MSB cell current is summed with the reference current and any discrepancy $\Delta I = I_{\mathrm{REF}} - I_{\mathrm{MSB}}$,

will flow to node B and produce the trim circuit output regulating voltage. This regulating voltage is then applied via the source-follower M4, to the trimmable part of the MSB cell M1b, when switch S_{a} is closed. The reason for utilizing the source follower M4, is to be able to employ a MOST for the C_{store} capacitor as the gate oxide has the largest sheet-capacitance available in process, hence minimizing capacitor area.

The trim circuit basically consists of an opamp regulated cascode current mirror. The reason for using OTAs in the trim circuit is two-fold:

- The trim circuit current input node A, is at a virtual ground node suitable for current summing. The input impedance is given by:

$$r_{\mathrm{in}}(s) = \frac{1}{g_{\mathrm{m5c}}\left(1 + A(s)\right)}\left(1 + \frac{R_{\mathrm{L}}}{r_{\mathrm{ds5c}}}\right) \tag{7.3}$$

  where $A(s)$ is the gain of the opamp and R_{L} is the finite output resistance of the biasing current source.

- The trim circuit output node B, producing the regulating voltage is at a high impedance approximately given by:

$$r_{\mathrm{out}} \approx R_{\mathrm{L}} \parallel r_{\mathrm{ds5}} g_{\mathrm{m5c}} r_{\mathrm{ds5c}} A(s) \tag{7.4}$$

As the trimmable part of the current source M1b, only carries 16 % of the MSB current, the loop gain requirements are modest. If we assume the residual errors of each of the 32 MSB current cells to be uncorrelated, the necessary minimum loop gain is approximately:

$$A_{\mathrm{loop,min}} > \sqrt{32}\frac{0.32 I_{\mathrm{MSB}}}{I_{\mathrm{LSB}}} \approx 60\,\mathrm{dB} \tag{7.5}$$

The simulated DC gain of the loop was found to be approximately 85 dB, providing a good margin and ensuring low residual offsets.

The switch S_{a} in fig. 7.3 is implemented as a single NMOST. As the load of switch is the gate of source follower MOST M4, the switch MOST can be minimum size. During the off-state of the switch, some leakage current $I_{\mathrm{leak},N}$, will inevitably flow through the reverse-biased junction diode connected to C_{store}. As shown in fig. 7.4, this effect can be partially alleviated by employing a dummy PMOST which is always in the off-state. The PMOST junction leakage current will then compensate the NMOST leaking resulting in the residue leakage given by: $I_{\mathrm{leak}} = I_{\mathrm{leak},N} - I_{\mathrm{leak},P}$. As qualitatively shown in fig. 7.4, this leakage current will then in turn modulate the MSB cell output current as the gate voltage of M1b changes. The maximum change in the MSB cell current is then given by the leak time which is dependent on the number of MSB cells which are to be trimmed: $T_{\mathrm{leak}} = 32 \cdot T_{\mathrm{trim}}$, where T_{trim} is the trim time for the individual MSB source. Hence we have:

$$\Delta I_{\mathrm{MSB,max}} = g_{\mathrm{m1b}} \frac{I_{\mathrm{leak}} T_{\mathrm{leak}}}{C_{\mathrm{store}}} \tag{7.6}$$

The resulting rms error power is given by:

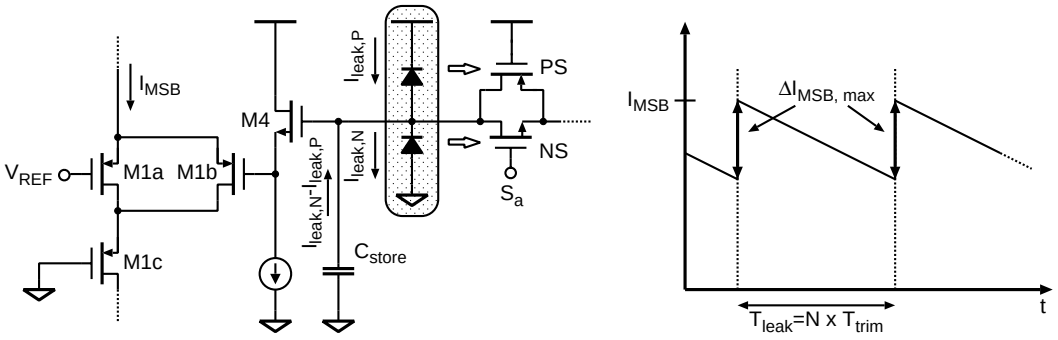


Figure 7.4: Diode leakage from C_{store}.

$$\Delta I^2_{\mathrm{MSB,rms}} = \frac{1}{T_{\mathrm{leak}}} \int_{-T_{\mathrm{leak}}/2}^{T_{\mathrm{leak}}/2} \left(\frac{g_{\mathrm{m1b}} I_{\mathrm{leak}}}{C_{\mathrm{store}}} t \right)^2 dt = \frac{\Delta I^2_{\mathrm{MSB,max}}}{12} \tag{7.7}$$

Assuming that all the error powers resulting from the 32 individual MSB cells are uncorrelated, we have the resulting RMS error due to junction leakage:

$$I_{\mathrm{leak,error(rms)}} = \sqrt{32 \Delta I^2_{\mathrm{MSB}}} = \sqrt{\frac{32}{12}} \Delta I_{\mathrm{MSB,max}} \tag{7.8}$$

We require that this resulting total error due to leakage should be much smaller than the quantization error in the MSB cell:

$$\sqrt{\frac{32}{12}} \Delta I_{\mathrm{MSB,max}} \ll \frac{I_{\mathrm{LSB}}}{\sqrt{12}} \tag{7.9}$$

Inserting the expression in eq. (7.6) in eq. (7.9) we can derive an expression for the minimum allowable storage capacitor C_{store}:

$$C_{\mathrm{store}} \geq \frac{\sqrt{32} g_{\mathrm{m1b}} I_{\mathrm{leak}} T_{\mathrm{leak}}}{I_{\mathrm{LSB}}} \tag{7.10}$$

Estimation of the leak current was done by simulation and a suitable storage capacitor size of 10 pF was chosen.

The trim loop GBW is to a first order determined by the M1b transconductance and the storage capacitor:

$$\mathrm{GBW}_{\mathrm{loop}} = \frac{g_{\mathrm{m1b}}}{C_{\mathrm{store}}} \tag{7.11}$$

which was found to be approximately 10 MHz in the simulations. The GBW of the regulating OTAs was set to 100 MHz to provide a wide margin from the trim loop speed. The DC gain of the OTAs is 60 dB.

7.2.3 MSB Cell Trim Sequence

In fig. (7.5) the MSB trim signals and a simplified schematic is shown for a trim sequence of two MSB cells. For each MSB cell, the trim cycle consists of five phases during which trim control transfer, trim setup and actual trimming is performed. A walk-through of the trimming of the two shown MSB

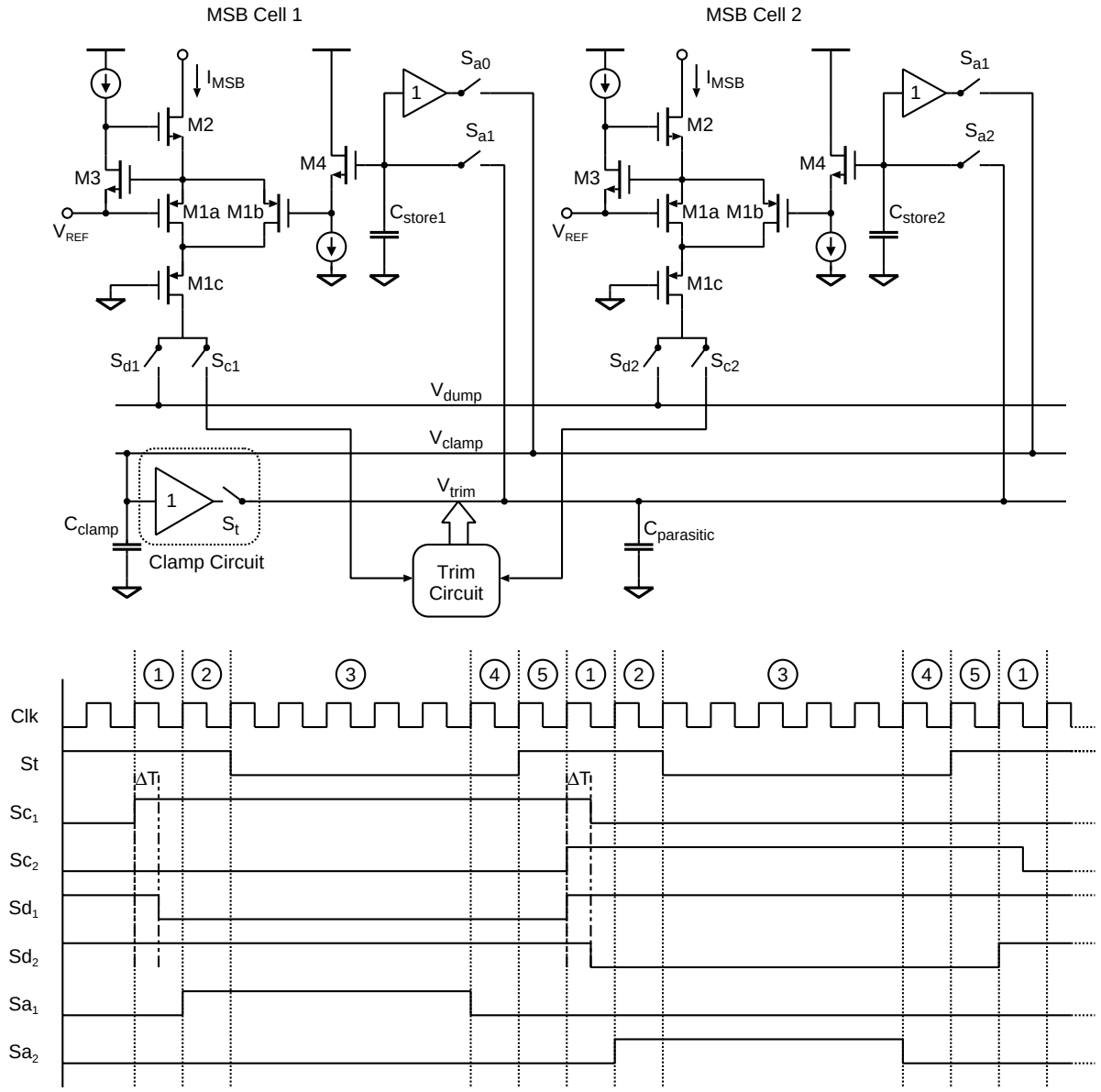


Figure 7.5: Trim sequence of MSB cells.

sources starting from the trim phase ③ of MSB cell 1 will be given in the following:

1. Starting from phase ③ of fig. (7.5), trimming of the first MSB cell is done. The calibrate switch $S_{\mathrm{c1}}$is on whereas the dump switch S_{d1} is low directing all the MSB cell current I_{MSB}, to the trim circuit. As the apply switch S_{a1} is on, the trim loop for MSB cell 1 is closed as depicted in fig. 7.3 charging the storage capacitor C_{store1}. Simultaneously as MSB cell 1 is being trimmed, the prior calibration voltage of MSB cell 2 is being transferred via a buffer to C_{clamp} in the clamp circuit in the lower left of the schematic.

2. The trimming of MSB cell 1 ends as S_{a1} goes low in phase ④. Simultaneously, the charging of C_{clamp} ends.

3. Phase ⑤ starts the transfer of the trim circuit from MSB cell 1 to MSB cell 2. As switch S_{t} is enabled, the clamping circuit effectively clamps the output of the trim circuit to the value stored on C_{clamp}, i.e. the calibration voltage of MSB cell 2. Saturation of the output of the trim circuit due to transient currents during trim transfer is thereby avoided.

4. In phase ①, the output currents of the MSB cells are redirected to V_{dump} and the trim circuit respectively. Hence for MSB cell 1, S_{d1} goes high and S_{c1} low and vice-versa for MSB cells 2. For each S_{d}-S_{c} pair, a small time overlap ΔT is inserted. The time-overlap serves to ensure that there always exists a path for the current in order to prevent voltage excursion at the M1 pair drain nodes. The current input node of the trim circuit and the current dump node are both low-impedance nodes kept at the same voltage. Hence the sum of currents flowing to V_{dump} and the trim circuit input is always constant.

5. For the final step of the trim transfer, the MSB cell 2 trim voltage apply switch S_{a2} is turned on in phase ②. However, as the clamp switch S_{t} is still enabled, the calibration voltage of MSB cell 2 is set to the value previously stored on C_{clamp}, i.e. the prior MSB cell 2 calibration voltage. This shows the second advantage of the clamp circuit: Voltage excursions on the MSB cell trimming node which would otherwise cause glitching in the MSB cell output current, are avoided.

6. As S_{t} goes low and we enter phase ③ for MSB cell 2, clamping of the trim circuit output ceases and the trim loop is closed for MSB cell 2.

In the trim sequence, we saw that the clamping circuit has a two-fold justification: To hold the output of the trim circuit at a known value during transfer of trim control to avoid saturation of the high-impedance trim circuit output. Also, as a significant amount of parasitic capacitance may be present on the V_{trim} line as shown in the figure, charge sharing between the storage capacitor C_{store} of the MSB cell and $C_{\mathrm{parasitic}}$ as the trim loop is closed could cause a severe glitch at the MSB cell output current with a maximum peak of:

$$i_{\mathrm{glitch,\,max}} = \frac{C_{\mathrm{parasitic}}}{C_{\mathrm{parasitic}} + C_{\mathrm{store}}} V_{\mathrm{trim}} g_{\mathrm{m1b}} \tag{7.12}$$

Any such glitching would be visible as frequency spectrum spurs located at some fraction of the trim clock frequency. The clamping circuit unity gain buffer is implemented as a single folded cascode OTA with a GBW of 70 MHz and a DC gain of 60 dB.

The number of clock cycles available for the actual trim phase, i.e. phase ③, was made configurable by two bits controlled off-chip. This feature enables us to accommodate differing leakage currents and to ensure that incomplete calibration of a source does not occur. The total number of clock cycles per full trim cycle and the resulting MSB refresh rate for a system clock frequency of 200 MHz is summarized in table 7.1. The trim clock is derived from the system clock in order to be able to synchronize the trim switching with the output switch activity.

Delay [0:1]	00	01	10	11
Number of clock-cycles	25	73	265	1033
MSB refresh rate [kHz]	250	85.6	23.6	6.05

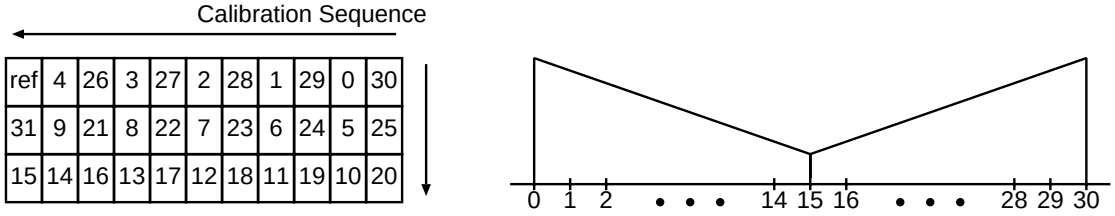
Table 7.1: Delay settings and resulting MSB refresh rate.

7.2.4 MSB Array Switching/Layout Considerations

The thermometer decoded MSB array was laid out in three rows and eleven columns. The MSB sources are calibrated starting at the upper right of the array, and then done from right to left for each of the rows as indicated in fig. 7.6(a). An exception is source number 31, as this current source is the starting point of the calibration sequence.

Located in the far left column, are the current cells which have special tasks. The current reference cell is located in the top left. Cell number 31 is used as a reference for the succeeding ULSB array and cell number 15 constitute the mid-code cell.

For the given calibration sequence, the MSB error due to leakage will constitute a symmetrical error around the mid-code cell number 15 as shown in fig. 7.6(b). As each current source is trimmed, the static linearity is guaranteed. The output switching sequence for the MSB array is shown in fig.7.6(a).



(a) MSB array switching sequence

(b) MSB array error distribution

Figure 7.6: MSB current cells switching sequence and calibration error distribution.

For a signal centered around mid-code, both half-phases of the signal will see the same error due to the error symmetry ensured by the calibration sequence. Thus the pair-wise grouping of the thermometer decoded current cells will minimize even order harmonics.

7.2.5 ULSB and LLSB Arrays

Both the ULSB and LLSB current source arrays were implemented as simple cascoded MOST current sources as shown in fig. 7.7. The last of the trimmed current sources in the MSB array, i.e. MSB source number 31, is used as a reference for the ULSB array. As the ULSB array is thermometer decoded, all of the current sources carry the same unit current.

The last of the ULSB current sources is used as a reference for the LLSB array. As the LLSB array is binary decoded, the current carried by the sources is bisected for each downward bit-step toward the LSB, as indicated in the figure.

Each of the unit ULSB sources consists of four unit MOSTs. For the layout of the ULSB array, a

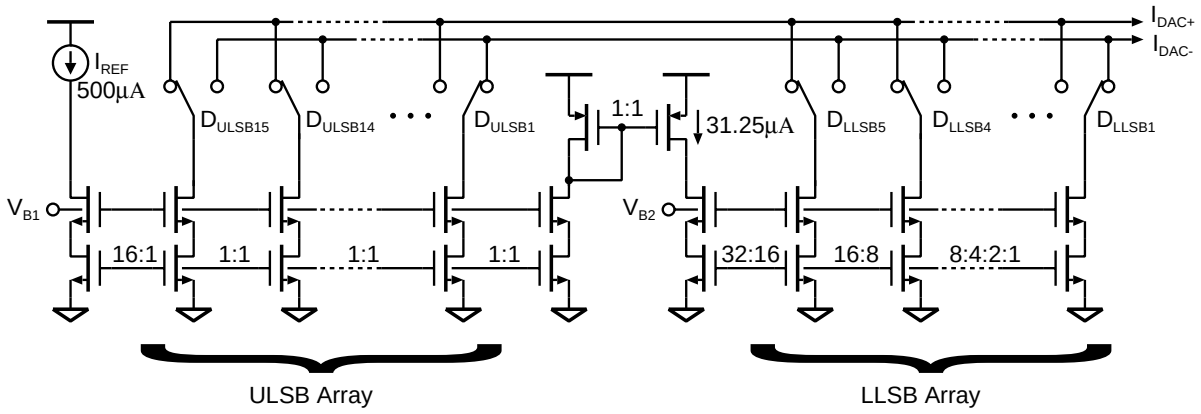


Figure 7.7: ULSB and LLSB array implementation.

unit MOST was placed in each quadrant in the array. As each unit current source is distributed across the array in a symmetrical manner, a common centroid layout for the ULSB array is achieved.

For the LLSB array, all the binary current sources are constructed from a number of unit MOSTs. The smallest binary current source, i.e. the LSB source, is composed of two such unit MOSTs. To achieve a common centroid layout for the unequally sized LLSB sources, the unit MOSTs were placed in a chessboard configuration ensuring equal distribution over the LLSB array [31].

The size of the unit MOST in each array was determined by matching considerations, and was set to $20\,\mu\mathrm{m}/9.3\,\mu\mathrm{m}$ for the ULSB and $2\,\mu\mathrm{m}/8\,\mu\mathrm{m}$ for the LLSB unit MOST respectively.

No special effort was put into the switching scheme of neither the ULSB or the LLSB array, and straight sequential switching is used in both arrays.

7.2.6 Output Switch Banks

All current source output switches were implemented as PMOST pairs. A schematic showing the output switch and the digital control circuitry is given in fig. 7.8. The size of the switch MOSTs is determined by the allowable voltage drop. In our case, the switch MOSTs were designed for a 50 mV voltage drop, which yields switch MOSTs with an aspect ratio of $40\,\mu\mathrm{m}/0.18\,\mu\mathrm{m}$ for the MSB cells. The succeeding switch banks for the two sub-DACs were scaled according to the current ratio.

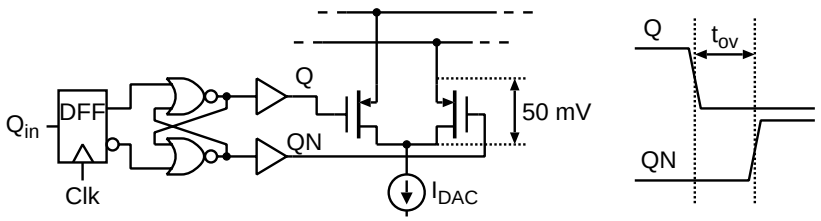


Figure 7.8: Current source output switch and digital control circuitry.

The switch timing control was implemented to ensure that a current path always exists. Using the simple digital circuit in fig. 7.8, a small overlap period t_{ov}, is ensured in the switching waveforms. By

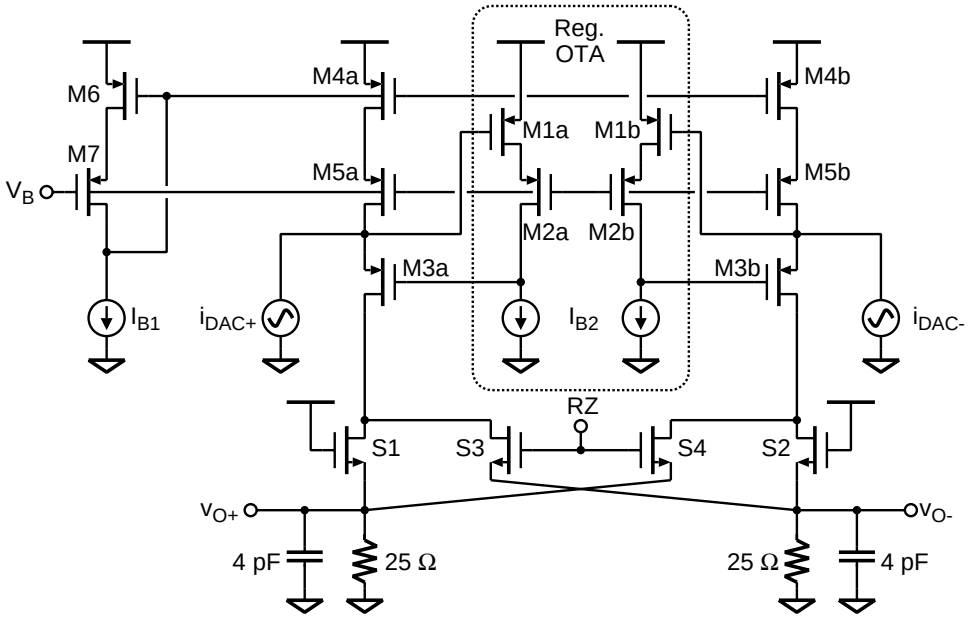


Figure 7.9: DAC current-folding output stage.

controlling the load driven by each of the NOR gates, the length of the overlap period can be controlled. In our case, the overlap period was set to 250 ps.

To match the different loads seen by the digital circuitry, digital buffers scaled according to switch size were inserted to drive the switch gates.

7.2.7 Output stage

A schematic of the output stage is shown in fig. 7.9. As explained earlier, the low voltage overhead necessitates the current folding stage to provide sufficient headroom for the output voltage swing. The M3 transistors act as folded cascodes for the DAC signal currents, which are represented by the $i_{\mathrm{DAC}+/-}$ sources. These cascodes are regulated by a very simple OTA consisting of two common-source stages in parallel. The primary merit of such a simple OTA is the ability to attain high speed. The resulting input impedance seen from the DAC is thus given by:

$$Z_{\mathrm{in}}(s) \approx \frac{1}{g_{\mathrm{m3}}\left(1 + A(s)\right)} \tag{7.13}$$

where $A(s)$ is the regulating OTA transfer function.

As each MSB current cell carries a current of $500\,\mu\mathrm{A}$, the total current swing of $16\,\mathrm{mA}$ sets the required bias current I_{B1}. For a bias current of $18\,\mathrm{mA}$, the large signal currents in the M3 MOSTs will vary between $2\,\mathrm{mA}$ and $20\,\mathrm{mA}$. This will cause a large variation of the M3 g_{m}, and cause the settling behavior to be non-equal for the positive and the negative outputs. Furthermore as this settling behavior is code dependent, harmonic distortion will result and grow progressively worse with higher input signal frequencies, as the DAC signal currents will change at an increasingly rapid rate.

To counter the nonlinear settling, a RZ scheme has been included in the output stage. In fig. 7.9, the

S1-S4 MOSTs constitute the RZ circuit. As all switch MOSTs are of equal dimension, the DAC signal current will be split equally into the cross-coupled branches and cancel the signal current of opposite phase. Another major reason for including a RZ scheme, is to mask any glitching stemming from the DAC during switching.

The frequency domain magnitude response of the RZ scheme can be written as [11]:

$$|\mathrm{RZ}(f)| = \frac{\tau}{T}\left|\frac{\sin\left(\pi f \frac{\tau}{T}\right)}{\pi f \frac{\tau}{T}}\right| \tag{7.14}$$

where the duty cycle is given by $\tau/T \cdot 100\%$.

From eq. (7.14), we see that the price paid for employing the RZ scheme, is that signal power directly proportional to the RZ signal duty-cycle will be lost, i.e. half the signal energy for a 50 % duty cycle. An added benefit is that since the output signal is of zero-order-hold type, the frequency domain roll-off improves in an inverse proportional manner with the duty cycle.

The duty cycle of the RZ also sets the minimum required GBW of the regulating OTA. If we assume a first order settling behavior, the necessary GBW for settling within *B* bits and a 50 % duty cycle is given by:

$$\mathrm{GBW} \geq \frac{1}{2\pi}\left(B+1\right)\ln\left(2\right) f_{\mathrm{s}} \tag{7.15}$$

where f_{s} is the DAC update frequency. For 14 bits accuracy and a DAC update frequency of 200 MHz, eq. (7.15) resolves to a minimum GBW of 331 MHz. The implemented regulating OTA features a simulated GBW of 600 MHz and a DC gain of 70 dB, providing a good safety margin for the requirements.

7.3 Measurements

A microphotograph of the implemented chip is shown in fig. 7.10, where the major blocks are highlighted. The employed $0.18\,\mu\mathrm{m}$ process, features 6 metal layers and a single poly layer.

7.3.1 Static Measurements

The static performance of the DAC was measured by stepping through all 14 bit input codes and recording the DAC output. As a very precise reference current source is needed for the static measurements, a HP4156A Precision Semiconductor Parameter Analyzer was used for sourcing the reference. The resulting INL and DNL are shown in fig. 7.11. Here, the INL is based on a best-fit line to the observed data. Spikes are clearly discernible in the DNL plot. These stem from each time one of the thermometer decoded MSB sources is switched, i.e. every 512 steps. These are also visible as the main ripple in the INL plot. The maximum integral and differential nonlinearity observed over the 16384 codes are:

$$\begin{aligned} \mathrm{INL}_{\mathrm{MAX}} &= 0.65 \cdot V_{\mathrm{LSB}} \\ \mathrm{DNL}_{\mathrm{MAX}} &= 0.55 \cdot V_{\mathrm{LSB}} \end{aligned} \tag{7.16}$$

A zoomed view of the INL is shown in fig. 7.12. Here, the DAC segmentation is clearly visible. The jump visible in the middle of the zoomed view is the switching of one of the MSB sources, whereas the

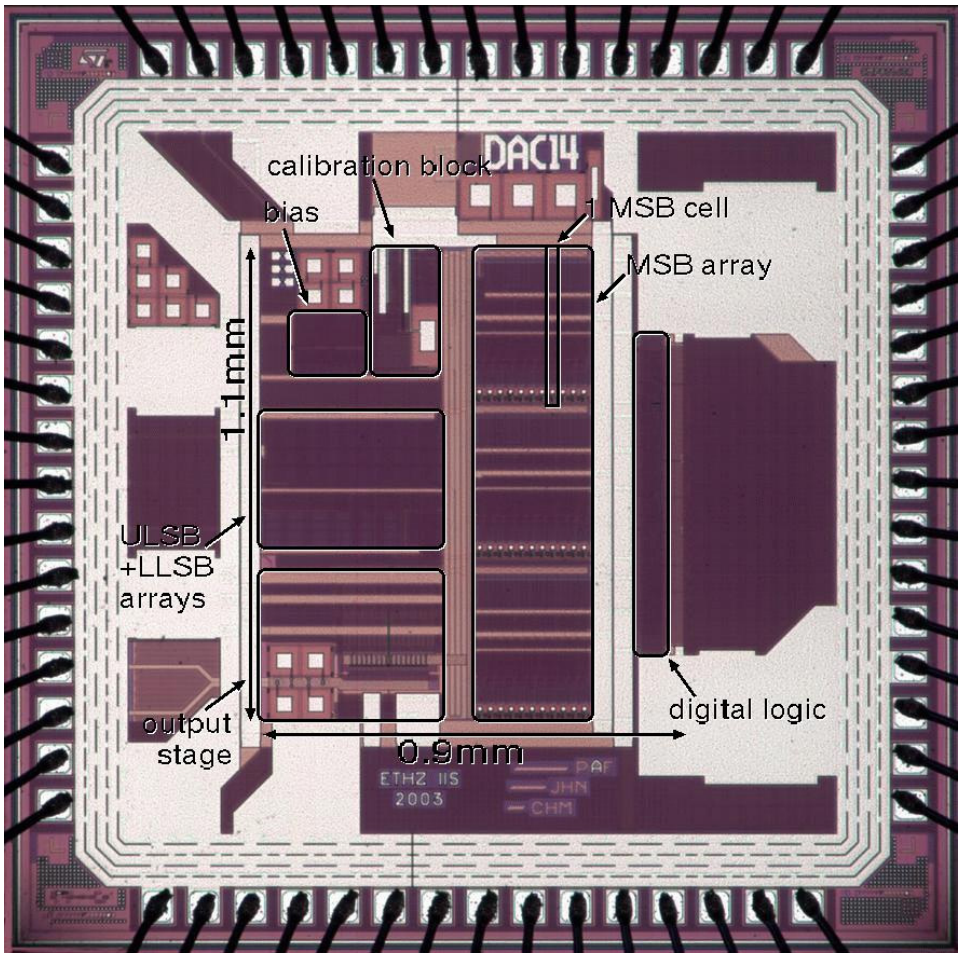


Figure 7.10: 200 MS/s 14 bit DAC microphotograph.

smaller jumps every 32 steps result from the switching of an ULSB source.

The maximum INL is slightly above the prescribed 1/2 LSB [12]. The transition that causes the slightly elevated INL value is due to the switching of the MSB sources. The main reason is believed to be the current mirror responsible for feeding the calibrated MSB current to the ULSBs. This current mirror has a 10 bit accuracy requirement imposed upon it, which may not be achieved in the fabricated circuit.

7.3.2 Dynamic Measurements

For generating the digital input data in the dynamic test setup, a HP16522A pattern generator was used. This pattern generator features a memory depth of 250 K vectors, at a maximum frequency of 200 MHz. The digital input sine was generated such that phase continuity was ensured for a cyclically repeated sequence. The dynamic performance is limited by the static performance at low signal frequencies. As the the signal frequency is increased, dynamic error effects start to dominate and typically appear as harmonic spurs in the output spectrum.

For the following measurements, the configurable trim time for the individual MSB cells was set to

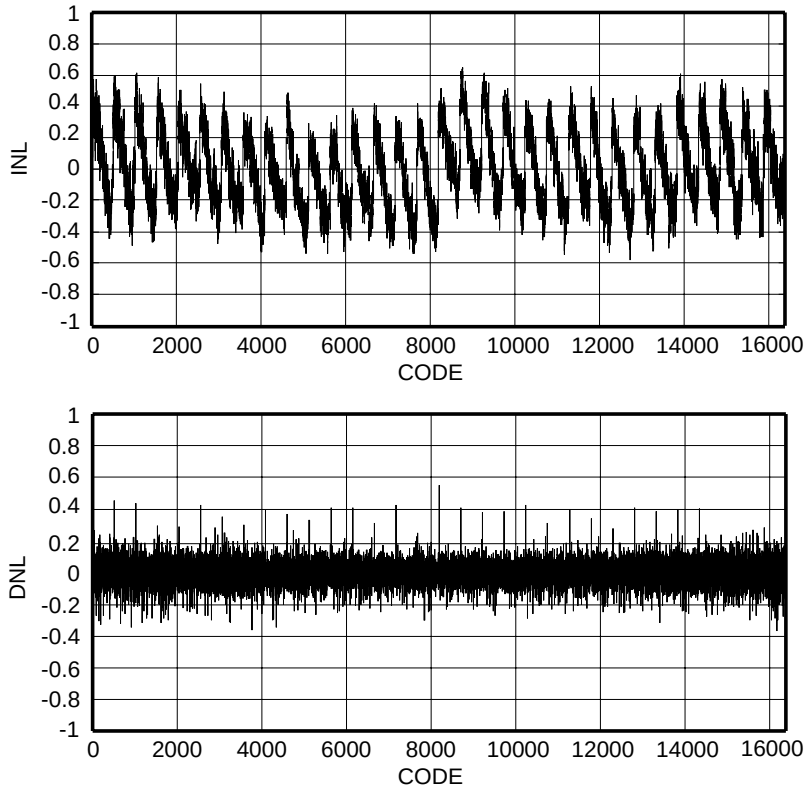


Figure 7.11: Integral nonlinearity and differential nonlinearity of the 14 bit DAC.

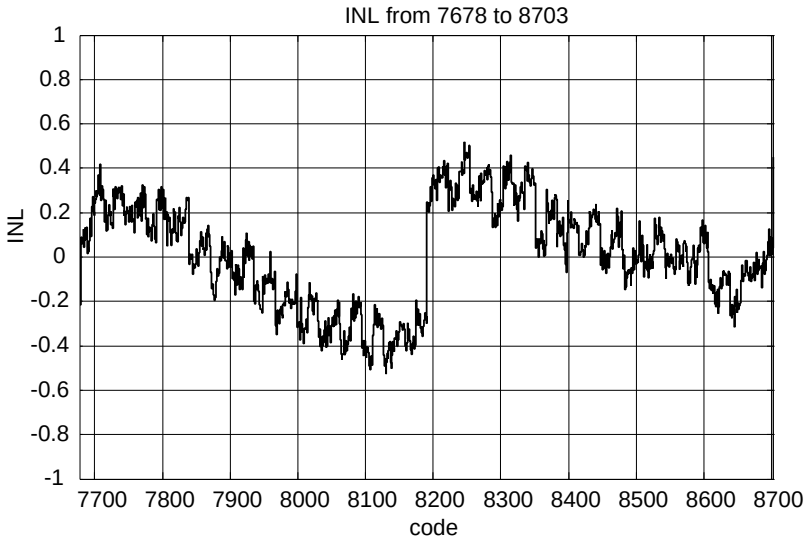
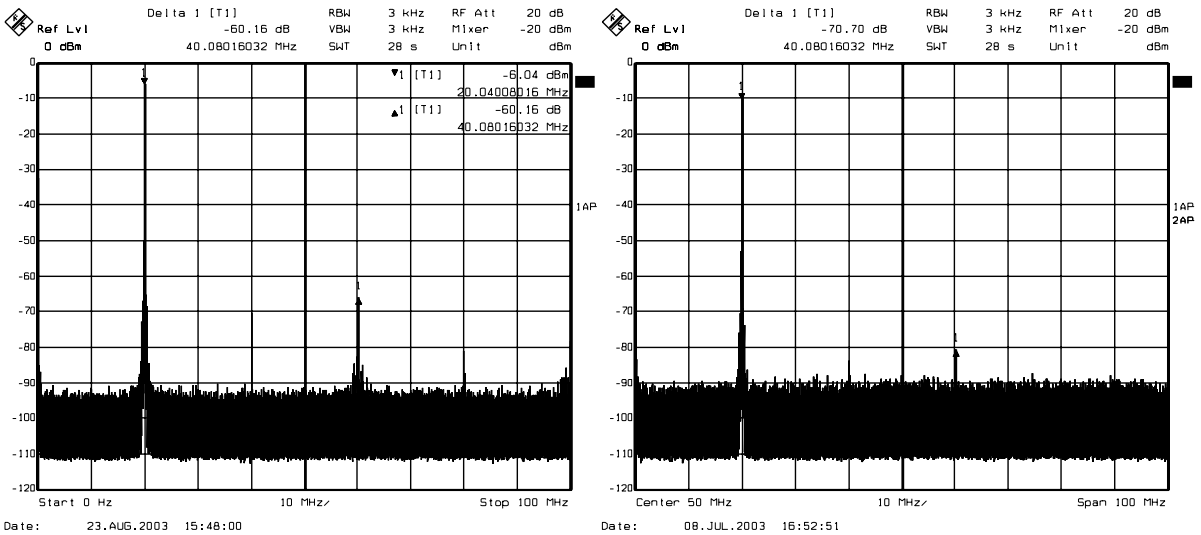


Figure 7.12: Zoomed view of the integral nonlinearity.

it's minimum setting, i.e. 21 clock cycles.

To illustrate the effect of the RZ-scheme, two example DAC output spectra are shown in fig. 7.13. The spectra were recorded using a Rohde and Schwarz model 1066.3010.30 spectrum analyzer. Here,



(a) RZ scheme disabled.

(b) RZ scheme enabled.

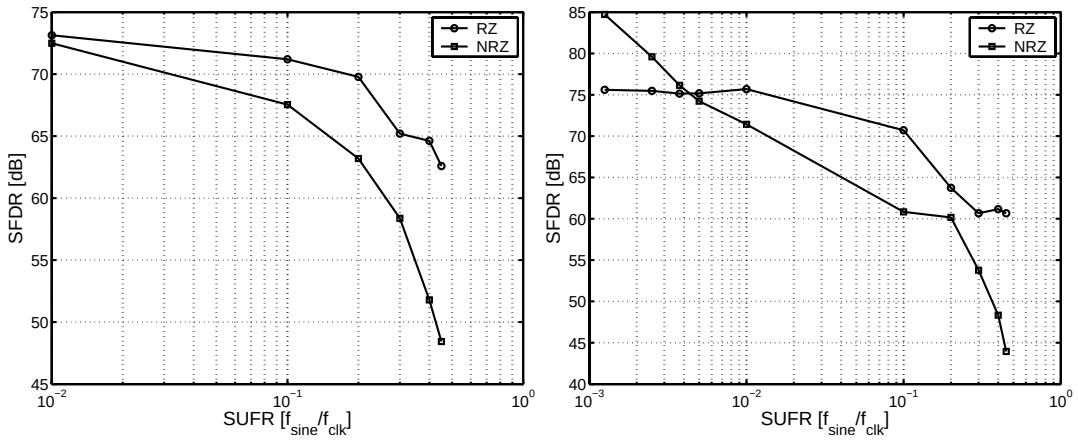
Figure 7.13: Example output spectra with a tone at 20 MHz and an update frequency at 200 MHZ, i.e. SUFR=0.1.

a signal frequency of 20 MHz is applied at an update rate of 200 MHz. For both spectra, the SFDR is dominated by the third harmonic. It is seen that SFDR is improved by approximately 10 dB when the RZ-scheme is applied. The RZ duty cycle was set to 33 %, equivalent to a signal loss of 3.5 dB which is also clear from the drop in tone energy seen from the spectra.

To investigate the general SFDR performance for the DAC over input frequency range, a full scale sine was applied at increasing tone frequencies for two DAC update frequencies of 100 MHz and 200 MHz. For each input frequency, the SFDR both with and without the RZ-scheme enabled was recorded. For the low update frequency, the lowest signal-to-update-frequency (SUFR) investigated was 1 %, whereas the lowest input tone frequencies at the 200 MHz update rate was set to SUFR=0.125%, i.e. $f_{\mathrm{sine}} \approx 250\,\mathrm{kHz}$. The resulting SFDR graphs for each update frequency are shown in fig. 7.14(a) and (b) respectively.

From both graphs, it is clear that the RZ-scheme improves the SFDR for increasing signal frequencies. However, for lower input frequencies, fig. 7.14(b) reveals that the dynamic nonlinearities imposed by the RZ-switches themselves will dominate the spectrum. Hence, the advantage of the RZ-scheme is lost for low input frequencies. The cross point of the SFDR curves in fig. 7.14(b) is seen to be approximately at SUFR=0.5%, i.e. 1 MHz at a 200 MHz update rate.

Dependent on the delay setting, spurs resulting from the trimming of the MSB array will appear in the DAC output spectrum. For the [00] delay setting, the rate at which the MSB sources are being exchanged for trimming is given by: $f_{\mathrm{MSB,exh}} = f_{\mathrm{clk}}/21 \approx 9.5\,\mathrm{MHz}$, for an update rate of 200 MHz. The other delay settings allocate more clock cycles for the trim phase and will thus result in spurs at lower frequencies. The DAC output power noise spectrum is shown in fig. 7.15. Just below 10 MHz, the calibration spurs are identified. As the maximum spur energy is at -97 dBm, the spurs will not limit



(a) Update frequency = 100 MHz.

(b) Update frequency = 200 MHz.

Figure 7.14: SFDR for various input tone frequencies and two different DAC update frequencies.

the overall SFDR at any SUFR. As the MSB exchange rate for trimming is highest at the [00] delay

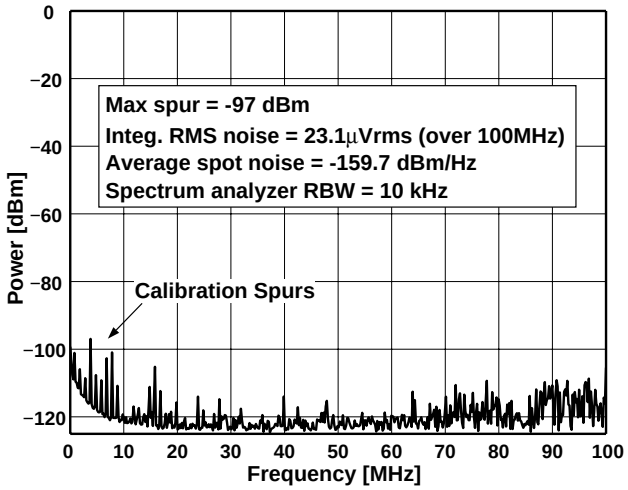


Figure 7.15: DAC noise floor and spurs stemming from the trimming of the MSB array.

setting, this also represents a worst-case condition as the MSB refresh rate is highest at this setting. Testing spurious performance at the other delay settings acknowledges this fact as the spurs recede and become indiscernible from the DAC noise floor.

The other parameter which was also extracted from the noise floor measurements, is the average noise floor which was found to be -159.7 dBm/Hz, equivalent to $2.3\,\mathrm{nV}\sqrt{\mathrm{Hz}}$.

7.4 Summary

The measured DAC performance and key specifications are summarized in table 7.2.

Supply Voltage (Analog and Digital)	1.8 V
Max. clock frequency	200 MHz
Analog supply current	48.7 mA
Digital supply current	5.3 mA
Power consumption	97 mW
Resolution	14 bits
Max. differential output current	16 mA
Max. INL	0.65
Max. DNL	0.55
SFDR @ $f_{\mathrm{signal}} = 250\,\mathrm{kHz}$	84 dB
Output noise floor	$-159\,\mathrm{dBm/Hz} \Rightarrow 2.3\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$
DAC core area	$0.98\mathrm{mm}^2$

Table 7.2: DAC key specifications and performance summary.

From the static measurements we have shown that the trimmed current source technique is applicable to high resolution DACs. The achieved static linearity is slightly above 1/2 LSB. It is conjectured that the main contributor to the slightly degraded static linearity, is the current mirror which serves to convey the last trimmed MSB source current to the ULSB array. This is corroborated by inspection of the INL curve, which features relatively large jumps at each 512 steps, i.e. each time an MSB source is switched. By increasing the number of MSBs to be trimmed, the matching requirements of the ULSB reference mirror would be relaxed and better static performance would be expected. However, increasing the number of trimmed MSB by one, comes at a cost of doubling the number of MSB current cells to be calibrated.

In the dynamic measurements, the linearity was seen to deteriorate rapidly as the input signal frequency was increased. To mitigate the deteriorated SFDR, a RZ-scheme was employed in the output stage for masking the transient behavior of the DAC. The RZ-scheme was seen to enhance the SFDR significantly at higher signal frequencies, whereas the nonlinear settling introduced by the RZ switches themselves, was seen to dominate at low signal frequencies. The price paid for introducing the RZ-scheme, is that a fraction of signal energy proportional to the RZ signal duty cycle is lost.

In [25], a similar RZ-scheme is employed albeit without significant SFDR loss at low input signal frequencies, thus demonstrating the feasibility of a RZ-scheme also at low input frequencies.

By utilizing floating current sources, a fully background trimming technique could be utilized for the MSB current sources, thus enabling minimum dynamic impact on the DAC output. This was confirmed by the measurements, where the spurs inevitably introduced by the background switching were seen to have insignificant power levels.

An area for improvement would be the current cell output switch control signals. These control signals are presently distributed from a central register to the individual switch pairs. For better synchronization, a latch should be included for each switch pair and a good clock tree should be employed to ensure low clock skew. The output switch on-time overlap presently used, could also be improved. In [24], a switch driver which shifts the cross-point of the switch control signals is proposed. Thereby the simultaneous switch on-time is minimized, while ensuring a current path at all time.

CHAPTER 8
CONCLUSION

In this thesis, circuitry for the interfacing of sensors and actuators has been treated. Although the entry-point of this treatise was in the field of sensors and actuators for implantable biomedical circuitry, the solutions presented can be applied to a wide range of applications. Many sensors are characterized by weak output signals and relatively low bandwidths as is the case for the nerve cuff electrode sensor. Another design requirement such as low-power, is a necessity in most any portable device imaginable.

For the sensor interface, we have treated weak signal amplification and analog to digital data conversion of the sensor output. Two prototype cuff-recorded nerve signal instrumentation amplifiers have been developed during this study. Both versions consist of two individual stages to achieve a high gain in a controlled manner. The first version is characterized by a relatively simple structure enabling high current efficiency and low noise levels, as very few transistors are employed for the signal amplification. To nullify the effects of input offset, a continuous-time offset compensation scheme by current steering is utilized. The 1/f-noise which typically dominates low-frequency noise performance in CMOS circuits, is dealt with simply by maximizing the amplifier input transistors. Though this first version amplifier fulfills the specifications set forth, some drawbacks in the design can be mentioned. The very large input transistors will inherently lead to poor PSRR at high frequencies as large capacitive coupling results. The biasing scheme used requires dual supplies and finally the large time constant necessitated by the continuous-time offset compensation scheme dictates the use of large capacitors, in the range of 1 nF. Such large-valued capacitors are expensive to implement on-chip as they require large area.

To remedy these drawbacks, a second version of the nerve signal amplifier is proposed. Single supply is achieved in this design, eliminating the need for a mid-supply pin. Rather than relying on raw sizing of the input transistors for 1/f-noise suppression, this amplifier uses the well-known chopper modulation technique for 1/f-noise suppression. The theoretical considerations showed that the chopper modulated amplifier gain is quite sensitive toward phase shift between the amplifier input and output. Since a relatively high biasing current was shown to be needed for thermal noise suppression in the amplifier input stage, a high bandwidth of the amplifier first stage results. The high bandwidth of the first stage makes it suitable for chopping, as the chopping frequency could be set well beyond the signal band without significant phase shift in the amplifier. A current-steering scheme was also employed for offset compensation in the second amplifier version. To avoid the use of large on-chip capacitors, a discrete-time scheme was used instead. An auxiliary OTA was added to each stage, which is used in a negative feedback manner to set the necessary steering voltage on a storage capacitor connected to the main amplifier offset nulling port. This discrete-time scheme has the disadvantage that the steering

voltage needs to be refreshed at an interval, necessitating the instrumentation amplifier to be brought off-line during a special calibration phase.

In the measurements, the first amplifier version is seen to correspond well with expected behavior. The more sophisticated second version had some problems in the first stage offset compensation scheme, leaving an excessively high residual offset. This error was tracked down to the first stage auxiliary OTA, as this OTA also featured an auto-zero scheme which failed to operate correctly. The second stage offset compensation was seen to operate correctly thus verifying the validity of the scheme. By disabling the auto-zero scheme, the functionality of the overall chopped amplifier could be verified. It was seen that dynamic offset had a significant impact on overall performance. An external low-pass filter was used in the measurement setup for eliminating this effect. Though the instrumentation amplifier second version remedies the mentioned drawbacks of the first design, some problems were apparent. Nevertheless, these could be mitigated without great effort in a potential redesign. Both amplifiers operate at sub-$300\,\mu\mathrm{W}$ power consumption, however excluding bias circuitry. As the second version operates at a lower supply voltage than the first design, lower power consumption is achieved in this circuit though more current is drawn. A slightly higher power consumption in a redesign of the second version should be expected, as an inter-stage low-pass filter is needed for mitigating dynamic chopper effects.

For post-treatment of the recorded neural signals, the pre-amplified signals are to be quantized. Numerous architectures exist for A/D conversion. In this study, a continuous-time $\Sigma\Delta$-modulator ADC was chosen for implementation. Theoretical analysis of this class of ADCs showed that their modeling varies significantly from Nyquist-rate data converters. For this reason, effort was put into presenting the modeling of these converter types. By theoretical estimation of expected $\Sigma\Delta$ ADC performance, loopfilter order, sampling frequency and loopfilter cut-off frequency could be extracted from a power-efficiency requirement.

For nerve stimulation, current pulses of differing magnitude and shape are used for emulating the natural nerve signal. For this purpose, the current steering DAC is well suited due to it's ability to generate the output current directly from the digital input code. The theory of this DAC architecture and possible implementations are treated with emphasis on minimizing systematic on-chip mismatch errors using different techniques. A current-steering DAC using full background calibration for high accuracy is implemented and measured. This DAC was intended as a generic demonstrator of a very high-speed and high accuracy DAC, with emphasis on high SFDR. Although the specifications for the implemented design are not suitable for implantable actuators, the current-steering DAC is easily downscaled in both speed and output current for other purposes. Hence the developed ideas remain valid for current-steering DACs with less stringent performance requirements imposed upon them, making them more suitable for biomedical purposes. High speed DACs could however be envisioned for future biomedical implants where very large numbers of actuators are present, e.g. visual retinal prosthesis.

In summary, the focus of this work has been on implementing highly integrated sensor/actuator interface circuitry with a minimal use of off-chip components. Several prototype demonstrators have been built to verify the developed solutions. Although some problems were encountered in the realized circuits, none proved fatal for the developed ideas.

Part III

Appendices

APPENDIX A

MOS TRANSISTOR FUNDAMENTALS

As all the implemented circuits described in this thesis are implemented in standard CMOS processes, a brief review of the fundamentals of this transistor type is in order. Several of the circuits employ the MOST in various regions of operation, from weak inversion to strong inversion. In order to be able to model the transistor correctly in all regions of operation, the EKV MOST model is employed [76]. In the following, all references and derivations are done for an NMOST. For PMOST, negative signs should be put in front of all large-signal voltages.

A.1 EKV Model Definitions

In order to maintain the symmetry inherent to the MOST, all voltages in the EKV model are referred to the bulk terminal as defined in fig. A.1 (a).

From the given terminal voltage definitions, the channel pinch-off voltage is defined:

$$V_{\mathrm{P}} \equiv V_{\mathrm{G}} - V_{\mathrm{th},0} - \gamma \left[V_{\mathrm{G}} - V_{\mathrm{th},0} + \left(\sqrt{\Psi_0} + \frac{\gamma}{2} \right)^2 - \left(\sqrt{\Psi_0} + \frac{\gamma}{2} \right) \right] \tag{A.1}$$

where the different parameters are:

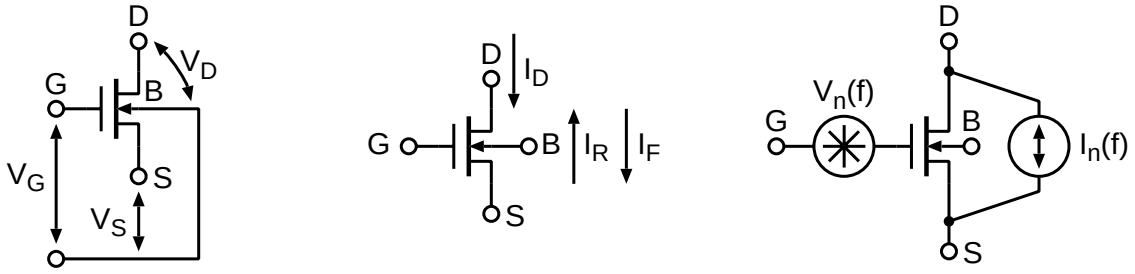
- $V_{\mathrm{th},0}$ - MOST native threshold voltage.
- γ - body effect factor.
- Ψ_0 - semiconductor surface potential.

The pinch-off voltage is defined as the voltage difference from the gate to either the drain or source terminal at which point the channel inversion charge becomes zero [77]. Hence if the source/drain voltage exceeds this value, the channel is effectively “pinched off” and the MOST enters a satured state. The pinch-off voltage is also termed the gate-overdrive or effective voltage.

From the definition of the pinch-off voltage, the socalled slope factor *n* is defined:

$$n \equiv \frac{\partial V_{\mathrm{G}}}{\partial V_{\mathrm{P}}} \tag{A.2}$$

The slope factor is necessary for calculating several small signal parameters and can be more easily



(a) MOST voltage definitions (b) MOST current definitions (c) MOST noise sources

Figure A.1: MOST voltage and current definitions with noise sources.

evaluated using the following expression where n is defined as a function of the gate voltage V_{G}:

$$\frac{1}{n} = \frac{\partial V_{\mathrm{P}}}{\partial V_{\mathrm{G}}} = 1 - \frac{\gamma}{2\sqrt{V_{\mathrm{G}} - V_{\mathrm{th},0} + \left(\sqrt{\Psi_0} + \frac{\gamma}{2}\right)^2}} \tag{A.3}$$

Typical values for the slope factor for the standard CMOS processes employed for the circuits implemented is $n = 1 \sim 1.5$.

A.2 Modes of Operation

A.2.1 Large Signal

If both V_{D} and V_{S} are below V_{P}, the channel is in strong inversion from source to drain and the transistor operates in the linear triode region. If the drain voltage exceeds the pinch-off voltage, the channel becomes pinched off at the drain terminal and the transistor enters saturation.

If both V_{D} and V_{S} are above V_{P}, the entire channel is pinched off and the transistor operates in weak inversion, as long as one of the terminal voltages stays close to V_{P}. For a drain-source voltage difference $V_{\mathrm{D}} - V_{\mathrm{S}} \gg V_{\mathrm{T}}$, the transistor enters a bipolar mode. V_{T} is the thermal voltage given by:

$$V_{\mathrm{T}} = \frac{kT}{q} \tag{A.4}$$

where k is the Boltzmann constant, T is the absolute temperature and q is the electronic charge.

The drain current in all modes of operation can be split up into a forward and a reverse current as shown in fig. A.1 (b):

$$I_{\mathrm{D}} = I_{\mathrm{F}} - I_{\mathrm{R}} \tag{A.5}$$

A table giving the forward and reverse current for weak and strong inversion operation is given below:

	Strong inversion		Weak inversion
	$V_{\mathrm{S(D)}} < V_{\mathrm{P}}$	$V_{\mathrm{S(D)}} \geq V_{\mathrm{P}}$	$V_{\mathrm{S}} > V_{\mathrm{P}}$ and $V_{\mathrm{D}} > V_{\mathrm{P}}$
$I_{\mathrm{F(R)}}$	$\frac{n\mu C_{\mathrm{ox}}(W/L)}{2}\left(V_{\mathrm{P}} - V_{\mathrm{S(D)}}\right)^2$	0	$2n\mu C_{\mathrm{ox}}\,(W/L)\,V_{\mathrm{T}}^2 e^{\frac{V_{\mathrm{P}} - V_{\mathrm{S(D)}}}{V_{\mathrm{T}}}}$

	Weak Inversion	Strong Inversion
Triode	$2nK'(W/L)V_{\mathrm{T}}^2 e^{V_{\mathrm{P}}/V_{\mathrm{T}}} \cdot \left[e^{-V_{\mathrm{S}}/V_{\mathrm{T}}} - e^{V_{\mathrm{D}}/V_{\mathrm{T}}}\right]$ for $\begin{cases} V_{\mathrm{S}} > V_{\mathrm{P}} \\ V_{\mathrm{D}} > V_{\mathrm{P}} \\ V_{\mathrm{S}} \simeq V_{\mathrm{D}} \end{cases}$	$nK'\left(\frac{W}{L}\right)\left[V_{\mathrm{P}} - \frac{(V_{\mathrm{S}}+V_{\mathrm{D}})}{2}\right] \cdot (V_{\mathrm{D}} - V_{\mathrm{S}})$ for $\begin{cases} V_{\mathrm{S}} \leq V_{\mathrm{P}} \\ V_{\mathrm{D}} \leq V_{\mathrm{P}} \end{cases}$
Saturated	$2nK'(W/L)V_{\mathrm{T}}^2 \cdot e^{(V_{\mathrm{P}}-V_{\mathrm{S}})/V_{\mathrm{T}}}$ for $\begin{cases} V_{\mathrm{S}} > V_{\mathrm{P}} \\ V_{\mathrm{D}} > V_{\mathrm{P}} \\ V_{\mathrm{S}} - V_{\mathrm{D}} \gg V_{\mathrm{T}} \end{cases}$	$\frac{1}{2}nK'\left(\frac{W}{L}\right)(V_{\mathrm{P}} - V_{\mathrm{S}})^2$ for $\begin{cases} V_{\mathrm{S}} \leq V_{\mathrm{P}} \\ V_{\mathrm{D}} > V_{\mathrm{P}} \end{cases}$

Table A.1: MOST drain current in strong and weak inversion when operated in triode and saturation.

where the previously undefined parameters are:

- μ - carrier mobility, μ_{n} for NMOST and μ_{p} for PMOST respectively.
- C_{ox} - gate oxide capacitance per unit area.
- W/L - transistor aspect ratio.

A typically employed shorthand for μC_{ox} is K'. Henceforth K'_{N} and K'_{P} for NMOST and PMOST respectively.

Evaluating the resulting drain current for the different operating modes of the MOST is summarized in table. A.1. In the weak inversion region, an often appearing term, which is designated the specific current [55], is given by: $I_{\mathrm{S}} = 2nK'(W/L)V_{\mathrm{T}}^2$. The magnitude of the specific current versus the drain current can be used as a simple measure for the level of inversion of the channel. I.e. if $I_{\mathrm{D}}/I_{\mathrm{S}} = 1$, the channel is said to be in moderate inversion, whereas for diminishing $I_{\mathrm{D}}/I_{\mathrm{S}}$, the channel approaches weak inversion and for an increasing ratio, the channel becomes strongly inverted. This current ratio is therefore designated the inversion coefficient (IC) [78]: $IC = I_{\mathrm{D}}/I_{\mathrm{S}}$. This gives a simple rule of thumb, which is summarized in the table below, which can be used for determining the operating inversion level of a MOST:

Inversion Coefficient	Channel inversion level
IC<0.1	Weak inversion
0.1<IC<10	Moderate inversion
10<IC	Strong inversion

A.2.2 Small Signal

From the large signal drain current, the small signal transconductance referring all voltages to the source can be extracted:

$$g_{\mathrm{m}} = \frac{\partial I_{\mathrm{D}}}{\partial V_{\mathrm{GS}}} \tag{A.6}$$

Equivalently, the small signal output conductance of the MOST can be found from:

$$g_{\mathrm{ds}} = \frac{\partial I_{\mathrm{D}}}{\partial V_{\mathrm{DS}}} \tag{A.7}$$

	Weak Inversion		Strong Inversion	
	Triode	Saturated	Triode	Saturated
g_{m}	$\frac{I_{\mathrm{D}}}{nV_{\mathrm{T}}} = 2\beta V_{\mathrm{T}}\left[e^{\frac{V_{\mathrm{G}}-V_{\mathrm{th}}-nV_{\mathrm{S}}}{nV_{\mathrm{T}}}} - e^{\frac{V_{\mathrm{G}}-V_{\mathrm{th}}+nV_{\mathrm{D}}}{nV_{\mathrm{T}}}}\right]$	$\frac{I_{\mathrm{D}}}{nV_{\mathrm{T}}} = 2\beta V_{\mathrm{T}} e^{\frac{V_{\mathrm{G}}-V_{\mathrm{th}}-nV_{\mathrm{S}}}{nV_{\mathrm{T}}}}$	$\frac{I_{\mathrm{D}}}{V_{\mathrm{G}}-V_{\mathrm{th}}-n/2(V_{\mathrm{S}}+V_{\mathrm{D}})} = \beta\left(V_{\mathrm{D}}-V_{\mathrm{S}}\right)$	$\frac{2I_{\mathrm{D}}}{V_{\mathrm{G}}-V_{\mathrm{th}}-nV_{\mathrm{S}}} = \sqrt{\frac{2\beta I_{\mathrm{D}}}{n}} = \frac{\beta}{n}\cdot (V_{\mathrm{G}}-V_{\mathrm{th}}-nV_{\mathrm{S}})$
g_{ds}	$\frac{I_{\mathrm{F}}}{V_{\mathrm{T}}} = 2n\beta V_{\mathrm{T}} e^{\frac{V_{\mathrm{G}}-V_{\mathrm{th}}-nV_{\mathrm{S}}}{nV_{\mathrm{T}}}}$	$\frac{I_{\mathrm{R}}}{V_{\mathrm{T}}} = 2n\beta V_{\mathrm{T}} e^{\frac{V_{\mathrm{G}}-V_{\mathrm{th}}-nV_{\mathrm{D}}}{nV_{\mathrm{T}}}}$	$\sqrt{2n\beta I_{\mathrm{R}}} = \beta\left(V_{\mathrm{G}}-V_{\mathrm{th}}-nV_{\mathrm{S}}\right)$	$\frac{k_\lambda I_{\mathrm{D}}}{L}$

Table A.2: Small signal transconductance and output conductance.

However, using the simple expressions for the drain current of table A.1 gives us an output conductance $g_{\mathrm{ds}} = 0$, for both the weak- and strong inversion saturated state, hence not rendering a usable model. Instead for the satured strong inversion region, we have [13]:

$$g_{\mathrm{ds,si}} = \frac{k_\lambda I_{\mathrm{D}}}{L} \tag{A.8}$$

where k_λ is a technology dependent parameter. For the satured weak inversion region we have [76]:

$$g_{\mathrm{ds,wi}} = \frac{I_{\mathrm{R}}}{V_{\mathrm{T}}} \tag{A.9}$$

The resulting small signal conductances for the different MOST operating modes are summarized in table A.2. For the expressions in table A.2, an approximation for the pinch-off voltage has been used [79]: $V_{\mathrm{P}} = 1/n(V_{\mathrm{G}} - V_{\mathrm{th}})$. A shorthand notation is furthermore used: $\beta = K'(W/L)$. From table A.2, we see that the transconductance is maximized by letting it enter weak inversion as a bipolar mode is imposed upon it. The bipolar transistor transconductance is given by [13]:

$$g_{\mathrm{m,bipolar}} = \frac{I_{\mathrm{C}}}{V_{\mathrm{T}}} \tag{A.10}$$

where I_{C} is the collector current. Since the MOST transconductance is scaled by the slope factor n, and the weak inversion transconductance is an asymptotic value, we see that the MOST transconductance will always be lower than the corresponding bipolar transconductance.

A.3 MOST Noise

The noise of a MOST can be modelled employing only two noise sources as shown in fig. A.1 (c). The drain current noise is the thermal noise and the power spectral density in saturation is given by [76]:

$$S_{\mathrm{th}}(f) = c4kTg_{\mathrm{m}} \tag{A.11}$$

where the scaling factor c can be shown to be $\frac{2}{3}n$ and $\frac{n}{2}$ for strong inversion and weak inversion operation respectively.

A simple model for the drain-referred flicker (1/f) noise is given in [80]:

$$S_{1/\mathrm{f}}(f) = \frac{K_{\mathrm{F}} I_{\mathrm{D}}^{\mathrm{AF}}}{L^2 C_{\mathrm{ox}} f^{\mathrm{ef}}} \tag{A.12}$$

where K_{F} is a technology dependent constant and the exponent 'ef' typically varies between 0.7-1 [81]. Dividing eq. (A.12) by g_{m}^2 gives us the gate-referred noise expression:

$$S_{1/\mathrm{f,\,gate}}(f) = \frac{1}{g_{\mathrm{m}}^2} \frac{K_{\mathrm{F}} I_{\mathrm{D}}^{\mathrm{AF}}}{L^2 C_{\mathrm{ox}} f^{\mathrm{ef}}} = \frac{n K_{\mathrm{F}} I_{\mathrm{D}}^{\mathrm{AF}-1}}{K' W L C_{\mathrm{ox}} f^{\mathrm{ef}}} \approx \frac{K'_{\mathrm{F}}}{W L C_{\mathrm{ox}} f} \tag{A.13}$$

In the last simplified expression, we approximated $n = \mathrm{ef} = \mathrm{AF} = 1$, and collected the technology constants $K'$and K_{F} into the single constant K'_{F}. For the transconductance, the expression for saturated strong inversion mode was used. As g_{m} reaches it's maximum in weak inversion, the strong inversion g_{m} can be used for a worst-case approximation of the 1/f-noise.

The last expression shows that the 1/f-noise is approximately bias independent. Usually, the 1/f-noise is modelled as a gate-referred noise voltage as shown in fig. A.1 (a).

The resulting gate-referred noise source for a saturated MOST can thus be written:

$$v_{\mathrm{n}}^2(f) = \frac{c4kT}{g_{\mathrm{m}}} + \frac{K'_{\mathrm{F}}}{W L C_{\mathrm{ox}} f} \tag{A.14}$$

The frequency at which the two noise types have equal power density is termed the noise corner and typically lies in the range of several kHz for MOSTs. Hence for low-frequency signal processing, the 1/f-noise typically dominates.

APPENDIX B

Z-TRANSFORMATION OF SAMPLED CONTINUOUS-TIME FILTERS

The subject of this appendix is to derive the necessary mathematical tools for obtaining equivalent discrete-time filters for sampled continuous-time filters. This proves useful in the design of continuous-time $\Sigma\Delta$-modulators.

B.1 Laplace Transform of a Sampled Signal

Consider the CT signal $y(t)$ for $t > 0$. The corresponding sampled DT signal $y(n)$, can be obtained from the CT signal:

$$y(n) \quad = \quad \sum_{k=0}^{\infty} y(kT)\delta(t-kT) = y(t)\sum_{k=0}^{\infty}\delta(t-kT) \tag{B.1}$$

where $\delta(t)$ is the Dirac delta function and T is the sampling period. So the sampled version of the continuous-time signal is obtained by multiplying it by a pulse train of delta functions with a distance of T between each pulse and can be viewed as a function resulting from the multiplication of two independent signals, namely $f_1(t) = y(t)$ and $f_2(t) = \sum_{k=0}^{\infty}\delta(t-kT)$.

The Laplace transform $Y_{\mathrm{s}}(s)$, of the sampled signal $y(n)$, which can be expressed as the product of two independent signals where both have well defined Laplace transforms, can be resolved by using the complex convolution theorem as derived in [82]. This theorem states:

$$F(s) = \mathcal{L}\left\{f_1(t)f_2(t)\right\} = \frac{1}{2\pi j}\int_{a-j\infty}^{a+j\infty} F_1(w)F_2(s-w)dw = \frac{1}{2\pi j}\oint_{\mathrm{C}} F_1(w)F_2(s-w)dw \tag{B.2}$$

where $\mathcal{L}\left\{\cdot\right\}$ is the Laplace transform operator. Let σ_1 and σ_2 denote the largest real part of the poles in $F_1(w)$ and $F_2(w)$ respectively, the Laplace transform defined in eq. (B.2) exists for:

$$\mathrm{Re}(s) > \sigma_1 + \sigma_2 \tag{B.3}$$

and when the real part a, in the complex line integral is chosen such that:

$$\sigma_1 < a < \mathrm{Re}(s) - \sigma_2 \tag{B.4}$$

The last equality in eq. (B.2) is valid for a closed contour C, which encloses all poles of $F_1(w)$ and where the integrated function $F(w) = F_1(w)F_2(s-w)$, has one excess pole [83].

When the given constraints above are fulfilled the complex contour integral of eq. (B.2) can be resolved using residues [83], which results in the Laplace transform being given by:

$$F(s) = \sum_{\text{poles in F}_1\text{(w)}} \text{Res}\left\{F_1(w)F_2(s-w)\right\} \tag{B.5}$$

where $\text{Res}\left\{\cdot\right\}$ denotes the residue taken at a particular singularity.

Recall the expression for the sampled signal $y(n)$, given by the CT signal $y(t)$, defined in eq. (B.1). The two corresponding time-signals of eq. (B.2) can now be identified as:

$$f_1(t) = y(t) \tag{B.6}$$

$$f_2(t) = \sum_{k=0}^{\infty} \delta(t-kT) \tag{B.7}$$

The corresponding Laplace transform for the input signal $y(t)$, is simply:

$$F_1(s) = \mathcal{L}\left\{y(t)\right\} = Y(s) \tag{B.8}$$

Whereas for the pulse train of delta functions, the Laplace transform can be shown to be [83]:

$$F_2(s) = \mathcal{L}\left\{\sum_{k=0}^{\infty} \delta(t-kT)\right\} = \frac{1}{1-e^{-sT}},\ \text{Re}(s) > 0 = \sigma_2 \tag{B.9}$$

where the last constraint is necessary for the existence of the Laplace transform of $f_2(t)$ as the series otherwise would diverge. Inserting the expressions found in eq. (B.8) and eq. (B.9) into the formula for the Laplace transform of the product function given in (B.5) we have:

$$Y_s(s) = \sum_{\text{poles in Y(w)}} \text{Res}\left\{Y(w)\frac{1}{1-e^{-(s-w)T}}\right\} \tag{B.10}$$

As the poles of $F_2(s)$ all have a real part $\sigma_2 = 0$, the constraint for choosing s when using eq. (B.10) simplifies to:

$$\text{Re}(s) > \sigma_{\text{y}} \tag{B.11}$$

where σ_{y} is the largest real part of the poles in $Y(s)$. Similarly, the constraint for choosing the real part in the complex line integral of eq. (B.2) simplifies to:

$$\sigma_{\text{y}} < a < \text{Re}(s) \tag{B.12}$$

which effectively supersedes the prior requirement given in (B.11). Furthermore, as $\text{Re}(s) > 0$ in order for the Laplace transform of the train of delta functions to exist, eq. (B.10) is valid if the following requirements are met:

1. At least one excess pole should be present in $Y(w)$.

2. The largest real part of a pole in $Y(w)$, denoted σ_{y}, and the real part a of the resulting complex line integral must fulfill: $\sigma_{\mathrm{y}} < a \leq 0$.

B.2 Continuous-Time Filter Z-Transformation

When the spectral properties of DT signals are to be investigated, the DT equivalent of the Laplace Transform is employed, i.e. the z-transform. The common notation for the z-transform is: $\mathcal{Z}\{\cdot\}$ By definition, the z-transform of a sampled signal is found by taking the Laplace transform of the sampled signal at each of the sampling instances. Applying the Laplace Transform to each of the sampled values in the infinite series given in eq. (B.1) we have:

$$Y_s(s) = \sum_{k=0}^{\infty} y(kT)e^{-ksT} \tag{B.13}$$

where the term e^{sT} stems from the definition of the Laplace transform. For the equivalent DT signal $y(k)$, the z-transform is given by [82]:

$$Y(z) = \sum_{k=0}^{\infty} y(k)z^{-k} \tag{B.14}$$

Eq. (B.13) is recognized as the z-transform of $y(n)$ if we replace e^{sT} by z. Hence the relationship between the Laplace transform variable s, and the z-transform variable z is given by [82]:

$$z = e^{sT} \tag{B.15}$$

$$s = T^{-1}\ln(z) \tag{B.16}$$

So the z-transform can be viewed as a mere shorthand notation for the Laplace transform of a DT signal. From the expressions derived above, we see that we can find the z-transform of an equivalent DT signal if we can obtain the Laplace transform of the sampled CT signal. In the previous section we derived a formula for finding the Laplace transform of a sampled signal by using residues. Inserting the relationship given in eq. (B.15), we can express eq. (B.10) in terms of z:

$$Y(z) = \sum_{\text{poles in Y(w)}} \mathrm{Res}\left\{Y(w)\frac{1}{1-e^{wT}z^{-1}}\right\} \tag{B.17}$$

For the feedback filter of the CT $\Sigma\Delta$-modulator, the effective combined filter is the product: $\widehat{H(s)} = H(s)G_{\mathrm{DAC}}(s)$. If the Laplace transform of the feedback DAC is known, an equivalent digital feedback filter, $H(z)$, can be found by using eq. (B.17) where $Y(w) = H(w)G_{\mathrm{DAC}}(w)$.

If we assume a zero-order hold function for the feedback DAC, the transfer function can be shown

to be: $G_{\mathrm{DAC}}(s) = \left(1 - e^{-sT}\right)/s$. Inserting the definition of z in the DAC transfer function yields:

$$G_{\mathrm{DAC}}(s) = \frac{1 - z^{-1}}{s} \tag{B.18}$$

Finally, the corresponding discrete-time filter $H(z)$, is found by inserting the combined feedback filter $H(s)G_{\mathrm{DAC}}(s)$, into eq. (B.17):

$$H(z) = \left(1 - z^{-1}\right) \sum_{\text{poles in } \frac{\mathrm{H(w)}}{\mathrm{w}}} \mathrm{Res}\left\{\frac{H(w)}{w}\frac{1}{1 - e^{wT}z^{-1}}\right\} \tag{B.19}$$

If the feedback filter is a lowpass type, it is seen that the combined feedback filter $\widehat{H(s)}$, will always have at least one excess pole, regardless of the filter order. Also, all filter poles will be located left to the imaginary axis for stability. Hence we see that both requirements given in subsection B.1 for employing eq. (B.19) are fulfilled. Numerous formulas for taking residues of a great variety of functions can be found in [82, 83]. For lowpass functions a particularly useful formula for taking residues of multiple poles is replicated below:

For the multiple pole $w = w_{\mathrm{p}}$ in transfer function $F(w)$, the residue c, is determined by the following formula [82]:

$$c = Res(F(w))|_{w=w_{\mathrm{p}}} \tag{B.20}$$

$$= \frac{1}{(k-1)!}\lim_{w \to w_{\mathrm{p}}} \frac{d^{(k-1)}\left[(w - w_{\mathrm{p}})^k F(w)\right]}{dw^{(k-1)}} \tag{B.21}$$

where k is the pole order.

APPENDIX C

THE DIGITAL ΣΔ-MODULATOR

This appendix presents the classic interpolative digital ΣΔ-modulator structure and determines the theoretical maximum signal to noise ratio for this structure.

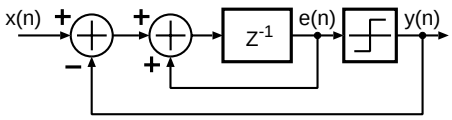


Figure C.1: Digital 1st. order ΣΔ-modulator.

Consider the first order digital ΣΔ-modulator scheme shown in fig. C.1. The z^{-1} block is a discrete delay and forms a discrete-time integrator in conjunction with the positive local feedback. The signal transfer function is given by:

$$\mathrm{STF}(z) = \frac{Y(z)}{X(z)} = \frac{1/(z-1)}{1+1/(z-1)} = z^{-1} \tag{C.1}$$

And the noise transfer function with regards to the quantization noise produced, is seen to be:

$$\mathrm{NTF}(z) = \frac{Y(z)}{Q(z)} = \frac{1}{1+1/(z-1)} = \left(1-z^{-1}\right) \tag{C.2}$$

To find the magnitude of the noise transfer function in the frequency domain, the identity: $z = e^{sT} = e^{j2\pi\frac{f}{f_s}}$ is inserted into the $\mathrm{NTF}(z)$expression given in (C.2):

$$\mathrm{NTF}(f) = 1 - e^{-j2\pi\frac{f}{f_s}} = \sin\left(\frac{\pi f}{f_\mathrm{s}}\right)\cdot 2j\cdot e^{-j\pi\frac{f}{f_s}} \tag{C.3}$$

Taking the magnitude of equation(C.3) thus yields:

$$|\mathrm{NTF}(f)| = 2\sin\left(\frac{\pi f}{f_\mathrm{s}}\right) \tag{C.4}$$

Cascading the structure of figure C.1 to any order n, as depicted in figure C.3, and repeating the above

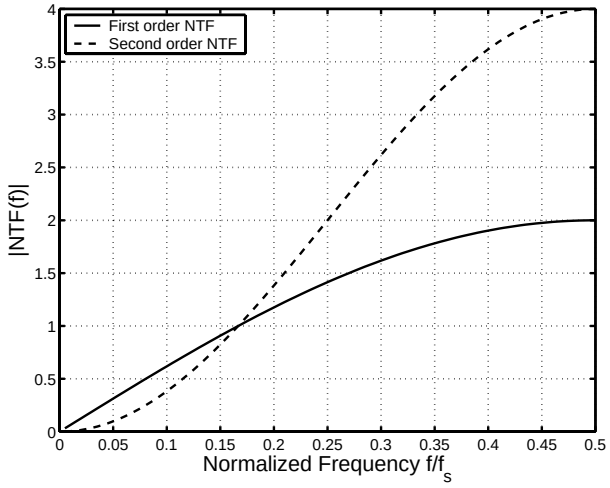


Figure C.2: Noise transfer functions for 1st. and 2nd. order digital ΣΔ-modulators.

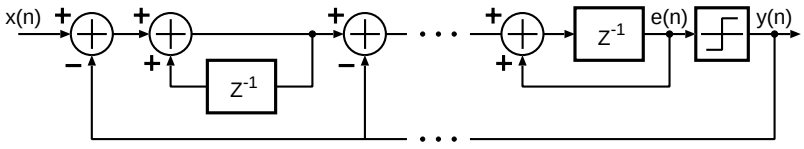


Figure C.3: Digital n-th. order ΣΔ-modulator.

analysis yields a noise transfer function with magnitude given by:

$$|\mathrm{NTF}(f)| = \left[2\sin\left(\frac{\pi f}{f_\mathrm{s}}\right)\right]^n \tag{C.5}$$

The resulting noise transfer functions are depicted for orders $n = 1, 2$ in figure C.2, up to half the sampling frequency. The noise shaping action of the modulator is clearly seen from these curves and it is evident that the second order modulator provides better performance than the first order modulator for $f \ll f_s$. Assuming the quantization noise to be white with a total noise power: $P_e = \frac{\Delta^2}{12}$, where Δ is the quantization level. As was previously stated, the white noise assumption yields a spectral density with the same height at all frequencies: $S_e(f) = \frac{\Delta}{\sqrt{12 \cdot f_s}}$.

The quantization noise power within the signal band, where the signal band is given by f_b, for a digital ΣΔ-modulator is given by the following integral, where $|\mathrm{NTF}(f)_n|$ denotes the frequency magnitude of a n^{th} order noise transfer function:

$$P_e = \int_{-f_b}^{f_b} S_e^2(f)\,|NTF(f)_n|^2\,df \tag{C.6}$$

$$= \int_{-f_b}^{f_b} \frac{\Delta^2}{12}\frac{1}{f_s}\left(2\sin\left(\frac{\pi f}{fs}\right)\right)^{2n} df \tag{C.7}$$

Since the signal band frequencies are much lower than the sampling frequency, we can approximate

$\sin\left(\frac{\pi f}{fs}\right)$ to be $\frac{\pi f}{fs}$, we have:

$$P_{\mathrm{e}} = \frac{\Delta^2}{12}\frac{1}{f_s}\int_{-f_b}^{f_b}\left(2\frac{\pi f}{f_s}\right)^{2n} df \tag{C.8}$$

$$= \frac{\Delta^2}{12}\frac{1}{f_s}\left(\frac{2\pi}{f_s}\right)^{2n}\frac{1}{2n+1}\left[f^{(2n+1)}\right]_{-f_b}^{f_b} \tag{C.9}$$

$$= \frac{\Delta^2}{12}\frac{\pi^{2n}}{2n+1}\left(\frac{2f_b}{f_s}\right)^{2n+1} \tag{C.10}$$

If we assume that we have a sine wave input at maximum peak value $\frac{\Delta}{2}$, it's power is given by: $P_s = \frac{\Delta^2}{8}$. The maximum signal to quantization noise ratio (SQNR) is thus given by:

$$\mathrm{SQNR}_{\max} = 10\log\left(\frac{P_s}{P_e}\right) \tag{C.11}$$

$$= 10\log\left(\frac{\frac{\Delta^2}{8}}{\frac{\Delta^2}{12}\frac{\pi^{2n}}{2n+1}\left(\frac{2f_b}{f_s}\right)^{2n+1}}\right) \tag{C.12}$$

$$= 10\log\left(\frac{3}{2}\left(\mathrm{OSR}\right)^{2n+1}\cdot\frac{2n+1}{\pi^{2n}}\right) \tag{C.13}$$

where it has been used that the oversampling ratio $\mathrm{OSR} = \frac{f_s}{2f_b}$.

BIBLIOGRAPHY

[1] C. I. Menolfi, *Low Noise CMOS Chopper Instrumentation Amplifiers for Thermoelectric Microsensors*. PhD thesis, Integrated Systems Laboratory, Dept. of Information Technology and Electrical Engineering, Swiss Federal Institute of Technology Zurich (ETHZ), March 2000.

[2] M. Thomsen, *Characterisation and Optimisation of Whole Nerve Recording Cuff Electrodes*. PhD thesis, Center For Sensory-Motor Interaction, University of Aalborg, 1998.

[3] S. Boyer, M. Sawan, M. Abdel-Gawad, S. Robin, and M. Elhilali, "Implantable selective stimulator to improve bladder voiding: design and chronic experiments in dogs," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, pp. 464–470, December 2000.

[4] A. Lickel, *Restoration of Lateral Hand Grasp in a Tetraplegic Patient Applying Natural Sensory Feedback*. PhD thesis, Center For Sensory-Motor Interaction, University of Aalborg, 1998.

[5] G. Gudnason, J. H. Nielsen, E. Bruun, and M. Haugland, "A distributed transducer system for functional electrical stimulation," in *Proceedings - International Conference on Electronics, Circuits and Systems, ICECS'01*, vol. 1, pp. 397–400, IEEE Circuit and Systems Society, September 2001.

[6] G. Gudnason, *CMOS Circuit Design for Biomedical Telemetry*. PhD thesis, Ørsted-DTU, Technical University of Denmark, August 2001.

[7] G. Gudnason, E. Bruun, and M. Haugland, "A chip for an implantable neural stimulator," *Analog Integrated Circuits and Signal Processing, Kluwer Academic Publishers*, vol. 22, pp. 81–89, January 2000.

[8] B. Smith and et al., "An externally powered, multichannel, implantable stimulator-telemeter for control of paralyzed muscle," *IEEE Transactions on Biomedical Engineering*, vol. 45, pp. 463–475, April 1998.

[9] C. C. Enz and G. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: Autozeroing, correlated double sampling, and chopper stabilization," *Proceedings of the IEEE*, vol. 84, pp. 1584–1614, November 1996.

[10] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE Journal of Solid-State Circuits*, vol. 24, pp. 1433–1440, October 1989.

[11] S. Pedersen, "Introduktion til signalanalyse." Notes in Danish for course at The Technical University of Denmark, February 1992.

[12] R. van de Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters*. Kluwer Academic Publishers, 1994.

[13] D. A. Johns and K. Martin, *Analog Integrated Circuit Design*. John Wiley and Sons, Inc., 1997.

[14] J. Bastos, M. Steyaert, and W. Sansen, "A high yield 12-bit 250-MS/s CMOS D/A converter," in *Proceedings - Custom Integrated Cicuits Conference, CICC'96*, vol. 4, pp. 431–434, IEEE Solid State Circuits Society, May 1996.

[15] A. van den Bosch, M. Steyaert, and W. Sansen, "The impact of statistical design on the area of high performance CMOS current-steering D/A converters," *Analog Integrated Circuits and Signal Processing, Kluwer Academic Publishers*, vol. 28, pp. 141–148, August 2001.

[16] J. Bastos, A. M. Marques, M. Steyaert, and W. Sansen, "A 12-bit intrinsic accuracy high-speed CMOS DAC," *IEEE Journal of Solid-State Circuits*, vol. 33, pp. 1959–1969, December 1998.

[17] T. Miki, Y. Nakamura, M. Nakaya, S. Asai, Y. Akasaka, and Y. Horiba, "An 80-MHz 8-bit CMOS D/A converter," *IEEE Journal of Solid-State Circuits*, vol. 21, pp. 983–988, December 1986.

[18] H. J. Schouwenaars, D. W. J. Groeneveld, and H. A. H. Termeer, "A low-power stereo 16-bit CMOS D/A converter for digital audio," *IEEE Journal of Solid-State Circuits*, vol. 23, pp. 1290–1297, December 1988.

[19] Y. Nakamura, T. Miki, A. Maeda, H. Kondoh, and N. Yazawa, "A 10-b 70-MS/s CMOS D/A converter," *IEEE Journal of Solid-State Circuits*, vol. 26, pp. 637–642, April 1991.

[20] G. van der Plas, J. Vandenbussche, W. Sansen, M. Steyaert, and G. Gielen, "A 14-bit intrinsic accuracy Q^2 random walk CMOS DAC," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 1708–1718, December 1999.

[21] C. Hammerschmied, *CMOS A/D Converters Using MOSFET-Only R-2R Ladders*. PhD thesis, Integrated Systems Laboratory, Dept. of Information Technology and Electrical Engineering, Swiss Federal Institute of Technology Zurich (ETHZ), 2001.

[22] D. W. J. Groeneveld, H. J. Schouwenaars, H. A. H. Termeer, and C. A. A. Bastiaansen, "A self-calibration technique for monolithic high-resolution D/A converters," *IEEE Journal of Solid-State Circuits*, vol. 24, pp. 1517–1522, December 1989.

[23] A. R. Bugeja, B.-S. Song, P. L. Lakers, and S. F. Gillig, "A 14-b, 100-MS/s CMOS DAC designed for spectral performance," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 1719–1732, December 1999.

[24] K. Falakshahi, C.-K. K. Yang, and B. A. Wooley, "A 14-bit 10-Msamples/s D/A converter using multibit $\Sigma\Delta$ modulation," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 607–615, May 1999.

[25] A. R. Bugeja and B.-S. Song, "A self-trimming 14-b, 100-MS/s CMOS DAC," *IEEE Journal of Solid-State Circuits*, vol. 35, pp. 1841–1852, December 2000.

[26] C. Lin and K. Bult, "A 10-b, 500 MSample/s CMOS DAC in 0.6 mm^2," *IEEE Journal of Solid-State Circuits*, vol. 33, pp. 1948–1958, December 1998.

[27] M. Hauser, "Principles of oversampling A/D conversion," *Journal of the Audio Engineering Society*, vol. 39, pp. 3–26, January/February 1991.

[28] T. Ritoniemi, T. Karema, and H. Tenhunen, "Design of stable high order 1-bit sigma-delta modulators," in *Proceedings - International Symposium on Circuits and Systems, ISCAS'90*, vol. 4, pp. 3267–3270, IEEE Circuit and Systems Society, May 1990.

[29] E. Stikvoort, "Some remarks on the stability and performance of the noise shaper or sigma-delta modulator," *IEEE Transactions on Communications*, vol. 36, pp. 1157–1162, October 1988.

[30] L. Risbo, $\Sigma\Delta$ *A/D Modulators - Stability Analysis and Optimization*. PhD thesis, Electronics Institute, Technical University of Denmark, June 1994.

[31] I. H. H. Jørgensen, *Current Mode Data Converters for Sensor Systems*. PhD thesis, Dept. of Information Technology, Technical University of Denmark, April 1997.

[32] J. H. Nielsen, "Sigma-delta modulator for miniature microphones," Master's thesis, Dept. of Information Technology, Technical University of Denmark, April 1999.

[33] P. C. Maulik, M. S. Chadha, W. L. Lee, and P. J. Crawley, "A 16-bit 250-kHz delta-sigma modulator and decimation filter," *IEEE Journal of Solid-State Circuits*, vol. 35, pp. 458–467, April 2000.

[34] S. Rabii and B. A. Wooley, "A 1.8-V digital-audio sigma-delta modulator in 0.8-μm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 783–796, June 1997.

[35] J. H. Nielsen and E. Bruun, "A design methodology for power-efficient continuous-time $\Sigma\Delta$ A/D converters," in *Proceedings - International Symposium on Circuits and Systems, ISCAS'03*, vol. 1, pp. 1069–1072, IEEE Circuit and Systems Society, May 2003.

[36] O. Oliaei, "Design of continuous-time sigma-delta modulators with arbitrary feedback waveform," *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 50, pp. 437–444, August 2003.

[37] E. J. van der Zwan and E. C. Dijkmans, "A 0.2mW CMOS $\Sigma\Delta$ modulator for speech coding with 80 dB dynamic range," *IEEE Journal of Solid-State Circuits*, vol. 31, pp. 1873–1880, December 1996.

[38] F. Gerfers, M. Ortmanns, and Y. Manoli, "A design strategy for low-voltage low-power continuous-time $\Sigma\Delta$ A/D converters," in *Proceedings - Design, Automation and Test Conference and Exhibition, DATE'01*, pp. 361–368, IEEE Computer Society, March 2001.

[39] F. Muñoz, A. P. VegaLeal, R. G. Carvajal, A. Torralba, J. Tombs, and J. Ramírez-Angulo, "A 1.1V low-power $\Sigma\Delta$ modulator for 14-b 16KHz A/D convesion," in *Proceedings - International Symposium on Circuits and Systems, ISCAS'01*, vol. 1, pp. 619–622, IEEE Circuit and Systems Society, May 2001.

[40] A. M. Abo and P. R. Gray, "A 1.5V, 10 -bit, 14.3-MS/s CMOS pipeline analog-to-digital converter," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 599–606, May 1999.

[41] V. Peluso, P. Vancorenland, A. M. Marques, M. Steyeart, and W. Sansen, "A 900-mV low-power $\Sigma\Delta$ A/D converter with 77-dB dynamic range," *IEEE Journal of Solid-State Circuits*, vol. 33, pp. 1887–1897, December 1998.

[42] J. A. Cherry and W. M. Snelgrove, "Excess loop delay in continuous-time delta-sigma modulators," *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 46, pp. 376–389, April 1999.

[43] S. H. Ardalan and J. J. Paulos, "An analysis of nonlinear behaviour in delta-sigma modulators," *IEEE Transactions on Circuits and Systems*, vol. 34, pp. 593–603, June 1987.

[44] L. Risbo, "Synthesis and simulations of stable high-order sigma-delta modulators," Master's thesis, Electronics Institute, Technical University of Denmark, August 1991.

[45] L. Risbo, "Stability predictions for high-order $\Sigma\Delta$ modulators based on quasilinear modeling," in *Proceedings - International Symposium on Circuits and Systems, ISCAS'94*, vol. 5, pp. 361–364, IEEE Circuit and Systems Society, May-June 1994.

[46] R. Schreier, "An empirical study of high-order single-bit delta-sigma modulators," *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 40, pp. 461–466, August 1993.

[47] A. Marques, V. Peluso, M. S. Steyaert, and W. M. Sansen, "Optimal parameters for $\Sigma\Delta$ modulator topologies," *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 45, pp. 1232–1241, September 1998.

[48] K. C. H. Chao, S. Nadeem, W. L. Less, and C. G. Sodini, "A higher order topology for interpolative modulators for oversampling A/D converters," *IEEE Transactions on Circuits and Systems*, vol. 37, pp. 309–318, March 1990.

[49] G. Fischer and A. J. Davis, "Alternative topologies for sigma-delta modulators - a comparative study," *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 44, pp. 789–797, October 1997.

[50] J. A. Cherry and W. M. Snelgrove, "Clock jitter and quantizer metastability in continuous-time delta-sigma modulators," *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 46, pp. 661–676, June 1999.

[51] J. Cherry and W. Snelgrove, *Continuous-Time Delta-Sigma Modulators for High-Speed A/D Conversion*. Kluwer Academic Publishers, 2000.

[52] B. Boser, "The design of sigma-delta modulation analog-to-digital converters," *IEEE Journal of Solid-State Circuits*, vol. 23, pp. 1298–1308, December 1988.

[53] H. Huang, S. Karthikeyan, and E. Lee, "A 1 V instrumentation amplifier," in *Proceedings - Midwest Symposium on Circuits and Systems, MWSCAS'99*, vol. 1, pp. 170–173, IEEE Circuit and Systems Society, August 1999.

[54] M. K. Haugland, *Natural Sensory Feedback for Closed-loop Control of Paralyzed Muscles*. PhD thesis, University of Aalborg, 1994.

[55] C. C. Enz, "The EKV model: a most model dedicated to low-current and low-voltage analogue circuit design and simulation," in *Low-Power HF Microelectronics A Unified Approach* (G. A. S. Machado, ed.), Circuits and Systems Series 8, ch. 7, pp. 247–299, IEE, 1996.

[56] J. H. Nielsen and T. Lehmann, "An implantable CMOS amplifier for nerve signals," *Analog Integrated Circuits and Signal Processing, Kluwer Academic Publishers*, vol. 36, pp. 153–164, July 2003.

[57] A. S. Sedra and K. C. Smith, *Microelectronic Circuits*. New York: Holt, Rinehart & Winston, 3rd ed., 1991.

[58] R. Rieger, J. Taylor, A. Demosthenous, N. Donaldson, and P. J. Langlois, "Design of a low-noise preamplifier for nerve cuff electrode recording," *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 1373–1379, August 2003.

[59] R. Adams, P. Ferguson, A. Ganesan, S. Vincelette, A. Volpe, and R. Libert, "Theory and practical implementation of a fifth-order sigma-delta A/D converter," *Journal of the Audio Engineering Society*, vol. 39, pp. 515–528, July 1991.

[60] Y. Matsuya, K. Uchimura, A. Iwata, T. Kobayashi, M. Ishikawa, and T. Yoshitome, "A 16-bit oversampling A-to-D conversion technology using triple-integration noise shaping," *IEEE Journal of Solid-State Circuits*, vol. 22, pp. 921–929, June 1987.

[61] D. Welland, P. del Signore, E. Swanson, T. Tanaka, K. Haruashita, S. Hara, and K. Takasuka, "A stereo 16-bit delta-sigma A/D converter for digital audio," *Journal of the Audio Engineering Society*, vol. 37, pp. 476–486, June 1989.

[62] M. Miller and C. Petrie, "A multi-bit sigma-delta adc for multi-mode receiver," in *Proceedings - Custom Integrated Cicuits Conference, CICC'02*, pp. 191–194, IEEE Solid State Circuits Society, May 2002.

[63] S. Gupta and V. Fong, "A 64-MHz clock-rate $\Sigma\Delta$ adc with 88-dB SNDR and -105-dB IM3 distortion at a 1.5-MHz signal frequency," *IEEE Journal of Solid-State Circuits*, vol. 37, pp. 1653–1661, December 2002.

[64] J. Grilo, I. Galton, K. Wang, and R. Montemayor, "A 12-mW ADC delta-sigma modulator with 80 dB of dynamic range integrated in a single-chip bluetooth transceiver," *IEEE Journal of Solid-State Circuits*, vol. 37, pp. 271–278, March 2002.

[65] A. Gerosa and A. Neviani, "A very low-power 8-bit $\Sigma\Delta$ converter in a 0.8 μm CMOS technology for the sensing chain of a cardiac pacemaker, operating down to 1.8V," in *Proceedings - International Symposium on Circuits and Systems, ISCAS'03*, vol. 5, pp. 49–52, IEEE Circuit and Systems Society, May 2003.

[66] V. Cheung, H. Luong, and M. Chan, "A 0.9-V 0.2μW CMOS single-opamp-based switched-opamp $\Sigma\Delta$ modulator for pacemaker applications," in *Proceedings - International Symposium on Circuits and Systems, ISCAS'02*, vol. 5, pp. 185–188, IEEE Circuit and Systems Society, May 2002.

[67] F. Henkel, U. Langemann, A. Hanke, S. Heinen, and E. Wagner, "A 1-MHz-bandwidth second-order continuous-time quadrature bandpass sigma-delta modulator for low-IF radio receivers," *IEEE Journal of Solid-State Circuits*, vol. 37, pp. 1628–1633, December 2002.

[68] R. van de Plassche, W. Sansen, and J. Huijsing, eds., *Analog Circuit Design*. Kluwer Academic Publishers, 1995.

[69] Q. Huang, "A MOSFET-only continuous-time bandpass filter," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 147–158, February 1997.

[70] U. Yodprasit and C. C. Enz, "A 1.5-V 75-dB dynamic range third order $G_m - C$ filter integrated in a 0.18-μm standard digital CMOS process," *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 1189–1197, July 2003.

[71] P. Van Peteghem, B. Haby, H. Fossati, and G. Rice, "A very-linear CMOS transconductance stage for OTA-C filters," in *Proceedings - Custom Integrated Cicuits Conference, CICC'89*, pp. 25.3/1–4, IEEE Solid State Circuits Society, May 1989.

[72] H. Aboushady, Y. D. M.-M. Louërat, and H. Mehrez, "Efficient polyphase decomposition of comb decimation filters in $\Sigma\Delta$ analog-to-digital converters," *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 48, pp. 898–903, October 2001.

[73] J. C. Candy, "Decimation for sigma delta modulation," *IEEE Transactions on Communications*, vol. 34, pp. 72–76, January 1986.

[74] M. Ortmanns, F. Gerfers, and Y. Manoli, "Clock jitter insensitive continuous-time $\Sigma\Delta$ modulators," in *Proceedings - International Conference on Electronics, Circuits and Systems, ICECS'01*, vol. 1, pp. 1049–1052, IEEE Circuit and Systems Society, September 2001.

[75] C. Kim and K. Wise, "A 64-site multishank CMOS low-profile neural stimulating probe," *IEEE Journal of Solid-State Circuits*, vol. 31, pp. 1230–1238, September 1996.

[76] C. Enz, F. Krummenacher, and E. A. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications," *Analog Integrated Circuits and Signal Processing, Kluwer Academic Publishers*, vol. 8, pp. 83–114, July 1995.

[77] M. Bucher, C. Lallement, C. C. Enz, and F. Krummenacher, "Accurate MOS modelling for analog circuit simulation using the EKV model," in *Proceedings - International Symposium on Circuits and Systems, ISCAS'96*, vol. 4, pp. 703–706, IEEE Circuit and Systems Society, May 1996.

[78] R. R. Harrison and C. Charles, "A low-power low-noise CMOS amplifier for neural recording applications," *IEEE Journal of Solid-State Circuits*, vol. 38, pp. 958–965, June 2003.

[79] G. A. S. Machado, C. C. Enz, and M. Bucher, "Estimating key parameters in the EKV MOST model for analogue design and simulation," in *Proceedings - International Symposium on Circuits and Systems, ISCAS'95*, vol. 3, pp. 1588–1591, IEEE Circuit and Systems Society, May 1995.

[80] W. Liu and et al., *BSIM3v3.2.2 MOSFET Model - Users Manual*. Dept. of Elec. Eng. and Computer Sciences, University of California, Berkeley, 1999.

[81] J. Chang, A. A. Abidi, and C. R. Viswanathan, "Flicker noise in CMOS transistors from subthreshold to strong inversion at various temperatures," *IEEE Transactions on Electron Devices*, vol. 41, no. 11, pp. 1965–71, 1994.

[82] E. I. Jury, *Theory and Application of the Z-Transform Method*. John Wiley and Sons, Inc., 1964.

[83] K. Andersen, "Samplede regulerings systemer." Notes in Danish for course at The Technical University of Denmark, August 1982.

APPENDIX D

PUBLICATIONS

During this study, the following journal and conference papers were presented:

1. Jannik H. Nielsen and Torsten Lehmann. “An implantable CMOS amplifier for nerve signals”. *Proc. 8th IEEE International Conference on Electronics, Circuits and Systems*, pp. 1183-1186, Malta, Sep. 2001.

2. Gunnar Gudnason, Jannik H. Nielsen, Erik Bruun and Morten Haugland. “A distributed transducer system for functional electrical stimulation”. *Proc. 8th IEEE International Conference on Electronics, Circuits and Systems*, pp. 397-400, Malta, Sep. 2001.

3. Jannik H. Nielsen and Erik Bruun. “A design methodology for power-efficient continuous-time $\Sigma\Delta$ A/D converters", *Proc. IEEE International Symposium on Circuits and Systems*, vol. 1, pp. 1069-1072, Thailand, May 2003.

4. Jannik H. Nielsen and Torsten Lehmann. “An implantable CMOS amplifier for nerve signals”. *Analog Integrated Circuits and Signal Processing*, 36(2), pp. 153-164, July, 2003.

5. Jannik H. Nielsen and Erik Bruun. "A CMOS chopper amplifier for implantable sensors". *Proc. 21^{st} NORCHIP Conference*, pp. 275-278, Latvia, Nov. 2003.

6. Quiting Huang, Pier-Andrea Francese, Chiara Martelli and Jannik Nielsen. “A 200MSPS, 14b, 97mW DAC in 0.18μm CMOS”, to appear in *Proc. IEEE International Solid-State Circuits Conference*, USA, Feb. 2004.

7. Jannik H. Nielsen and Erik Bruun. “A low-power 10-bit continuous-time CMOS $\Sigma\Delta$ A/D converter", to appear in *Proc. IEEE International Symposium on Circuits and Systems*, Canada, May 2004.

The papers are reproduced on the following pages with permission of their respective copyright holders.

AN IMPLANTABLE CMOS AMPLIFIER FOR NERVE SIGNALS

Jannik Hammel Nielsen[1], Torsten Lehmann[2]

[1]Ørsted-DTU, Technical University of Denmark, DK-2800 Kgs. Lyngby, jhn@oersted.dtu.dk
[2]Cochlear Ltd., 14 Mars Road, Lane Cove 2066 NSW Australia, tlehmann@cochlear.com.au

ABSTRACT: In this paper, a low noise high gain CMOS amplifier for minute nerve signals is presented. By using a mixture of weak- and strong inversion transistors, optimal noise suppression in the amplifier is achieved. A continuous-time offset-compensation technique is utilized in order to minimize impact on the amplifier input nodes. The method for signal recovery from noisy nerve signals is presented. A prototype amplifier is realized in a standard digital 0.5 μm CMOS single poly, n-well process. The prototype amplifier features a gain of 80 dB over a 3.6 kHz bandwidth, a CMRR of more than 87 dB and a PSRR greater than 84 dB. The equivalent input referred noise in the bandwidth of interest is $5\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$. The amplifier power consumption is 275 μW.

1 INTRODUCTION

Thousands of individuals sustain damages to the central nervous system, e.g. spinal chord injury, stroke, etc. which potentially results in paralyzed limbs. However, such damages usually leave the nerves and muscles in the limbs unaffected. In recent years, biomedical research has focused on retrieving nervous information from the natural sensors of the body and using this to activate paralyzed muscles by Functional Electrical Stimulation (FES) [1, 2]. Succesful trials have been conducted using external devices, correcting dropfoot in early schlyrosis patients [1], and restoring basic hand functions [2], by using natural sensor feedback.

Our involvement comes from developing implantable ASICs for both sensing neural information and for stimulating muscles through FES. An implantable stimulator has been developed and is reported in [3]. Developing implantable devices, reduces the risk for infection as skin continuity is ensured. Also, the need for external sensors which need calibration, may be bulky and are subject to mechanical stress limiting device life-span [1], is eliminated by using the natural sensors of the body.

This paper reports the implementation of the proposed MOSFET amplifier in [4], for amplification of minute nerve signals.

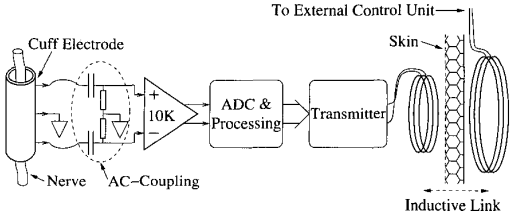


Fig. 1: Proposed nerve signal recovery implant.

2 PROPOSED SENSING IMPLANT

The overall sensing system for implantation to be developed is shown in fig. 1.

The electrical contact with the nerve is obtained using a so-called *cuff electrode* [1, 2]. The cuff electrode has a length of 10-20 mm and is placed around the nerve trunk. It has a contact resistance of about 5 kΩ and provides a differential nerve signal of about $\pm 10\,\mu$V. Control signals, data transmission and power supply are conveyed through the inductive link, dictating a very small supply current, on the order of a few hundred μA.

The signal from the cuff is AC-coupled to remove inherent DC offsets. Due to the very small signal amplitude, it is necessary to preamplify the signal prior to any processing. The preamplifier increases the input amplitude from $\pm 10\,\mu$V to ± 100 mV. As we are not interested in the absolute value of the signal, gain variation of 10%-15% can be tolerated. The nerve signals reside primarily in the bandwidth $400 < f_\mathrm{n} < 4000$ Hz [2], giving the necessary amplifier bandwidth.

After amplification the signal is anti-aliased and A/D converted. Some local processing to recover the nerve signal is then performed and the resulting signal is transmitted to the external control system through the inductive link.

3 SIGNAL RECOVERY

The technique used for signal recovery [1, 2], samples the nerve signal at $f_\mathrm{s} = 10$ kHz. After A/D conversion, the signal is rectified, integrated and downsampled to $f_\mathrm{r} = 20$ Hz, giving the signal envelope.

The nerve signal can be described as a zero mean

0-7803-7057-0/01/$10.00 ©2001 IEEE.

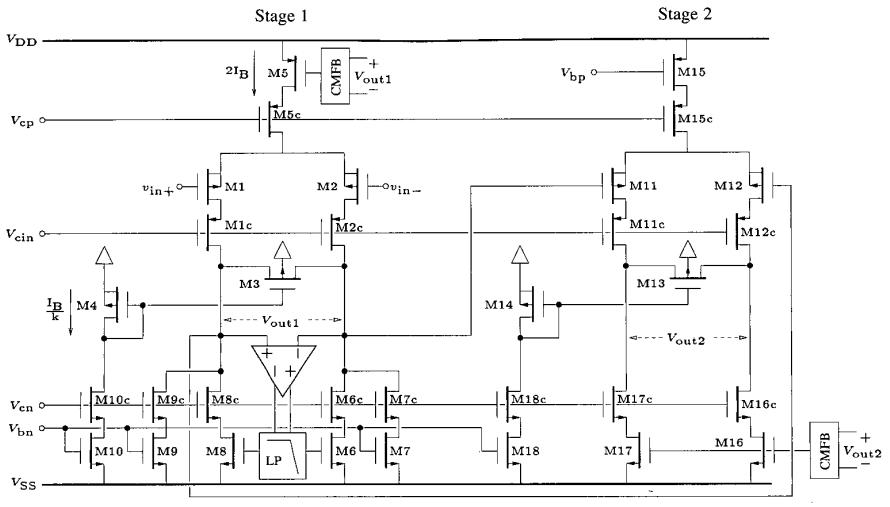


Fig. 2: Amplifier schematic.

stochastic variable V_n, with variance and mean:

$$\sigma_n^2 = \mathrm{E}\left\{V_n^2\right\} \text{ and } \mu_n = \mathrm{E}\left\{V_n\right\} = 0$$

The rectified signal V_r and its mean are given by:

$$V_r = \sqrt{V_n^2} \text{ and } \mu_r = \mathrm{E}\left\{\sqrt{V_n^2}\right\}$$

We can thus find the variance of the rectified signal:

$$\sigma_r^2 = \mathrm{E}\left\{\left(V_r - \mathrm{E}\left\{V_r\right\}\right)^2\right\} = \sigma_n^2 - \mu_r^2 \qquad (1)$$

Assuming a gaussian distribution of V_n, the mean of V_r can be found to be: $\mu_r = \sqrt{2/\pi}\,\sigma_n$. Inserting μ_r in eq. (1) yields the rectified signal variance: $\sigma_r^2 = \sigma_n^2\,(1 - 2/\pi)$. The variance of the downsampled signal is reduced by the ratio of sampling frequencies $k = f_s/f_r$. Thus the SNR of the downsampled signal is:

$$\mathrm{SNR} = 10\log\left(k\mu_r^2/\sigma_r^2\right) \simeq 29\,\mathrm{dB},\ f_s = 10\,\mathrm{kHz}$$

Though this SNR may seem low, it is sufficient to determine the state of a human nerve.

4 AMPLIFIER

A diagram of the designed amplifier is shown in fig. 2. Transistors denoted by a 'c' are used as cascodes. The amplifier consists of two amplification stages with similar topology, each has a gain of 100.

4.1 First stage

Due to the small amplitude of the input signal, it is of prime importance that the noise of the input transistors is kept to a minimum. The amplifier induced noise should remain below the thermal noise inherent to the cuff electrode. The input referred noise in a strong inversion MOSFET is given by [5]:

$$\frac{V_i^2}{\Delta f} = 4kT\left(\frac{2}{3}\right)\frac{1}{g_m} + \frac{K_f}{WLC_{ox}f} \qquad (2)$$

The first term of eq. (2) is the thermal noise in the MOSFET and the second term is the flicker (1/f) noise, where K_f is a process-dependent constant. In general, 1/f noise is found to be lower in P-MOS than in N-MOS transistors [6], thus P-MOS transistors are used in the input differential pair. Eq. (2) shows that the thermal noise is minimized by maximizing the transconductance g_m. Using the EKV-model, the g_m of the MOSFET in weak- and strong inversion can be shown to be [7]:

Weak inversion: $g_{m,\,weak} = I_D/nV_T$
Strong inversion: $g_{m,\,strong} = 2nI_D/V_{eff}$

Where $n \simeq 1.25$ is the slope factor and the thermal voltage $V_T \simeq 26.7\,\mathrm{mV}$ at 37°C. The ratio of the transconductances is: $g_{m,\,weak}/g_{m,\,strong} \simeq 5.6$ for a typical effective voltage of 300 mV. Thus, by biasing the input transistors in weak inversion, we can obtain maximum thermal noise suppression for a given I_D. So in order to obtain a low input referred noise, most of the amplifier current is drawn by M1 and M2.

The DC gain of stage 1 is given by:

$$A_{V1} = -\frac{g_{m1}}{2g_{ds3}} = \frac{(W/L)_1}{(W/L)_3}\cdot\frac{1}{2n} \qquad (3)$$

Provided M1, M2 and M3 in fig. 2 are biased in weak inversion. Thus the gain can set by device dimensions. The prerequisite of eq. (3) is that

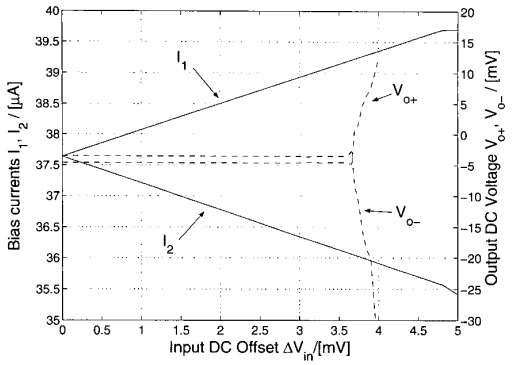


Fig. 3: Offset cancellation by current steering.

the gate-source voltages of M1, M2 and M3 are equal. This is ensured by letting a scaled current I_B/k, flow through the diode-connected M4 and using V_{G4} to bias M3. M4 is accordingly scaled $(W/L)_4 = 1/k\,(W/L)_1$, and will also be working in weak inversion. The source/drain voltage of M3 is set equal to zero by a common mode feedback circuit. All current source transistors (M6-M10) operate in strong inversion to minimize their noise contribution [5].

Gain variation due to device dimension mismatch can be minimized by layout techniques. The gain is however also dependent on the slope factor n. It can be shown that n is ultimately dependent on the substrate doping concentration N_{sub}. For variations in N_{sub} of ± 1 order of magnitude, the variation in n was found to be less than 10%, which is acceptable.

4.2 Offset cancellation

As the inherent threshold offset ΔV_{t0}, present in the input pair M1, M2 indeed may be orders of magnitude larger than the input signal, it will force the input stage out of weak inversion. Hence some scheme for offset cancellation is needed. To avoid the clock-feedthrough problems of switched methods, a continuous time scheme is utilized.

In fig. 2, the output of stage 1 is LP filtered and used to control source transistors M6 and M8 to match the current offset in M1 and M2. As we are not interested in DC, this scheme is applicable provided the time constant around the control loop is large enough. In fig. 3, the simulated bias currents I_1, I_2 and the measured output voltages of stage 1 are shown vs input DC-offset. Fig. 3 shows that offsets $< 3.5\,\mathrm{mV}$ will be compensated. Some residual offset will remain at the output due to device mismatch in the control loop and is $\simeq 1.1\,\mathrm{mV}$ in fig. 3.

Using this scheme will cause some bias current mismatch in M1 and M2. The operating point

Supply Voltage	3 V
Power Consumption	275 μW
SNR	36 dB
THD(@ f=1 kHz)	2.2 %
Typ. Input Offset (Stage 1)	$< 40\,\mu$V
CMRR	> 87 dB, $f < 100$ kHz
PSRR	> 84 dB, $f < 100$ kHz
Gain	$10000 \pm 10\%$
Equiv. input referred noise	$5\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$

Table 1: Measured Amplifier Performance.

transconductance of M1 and M2 is given by [7]:

$$g_m = 2nK_p \left(\frac{W}{L}\right) \frac{1}{nV_T} \exp\left(\frac{V_G - nV_S - V_{t0}}{nV_T}\right)$$

The current offset can be modelled by letting a small pertubation ΔV_{GS}, offset the operating point:

$$g_{mi} = g_m \exp\left(\pm \Delta V_{GS}/nV_T\right),\ i = 1,\ 2 \qquad (4)$$

As $|\Delta V_{GS}| \ll nV_T$, we can use a 1st order approximation for eq. (4). Thus the total transconductance of the differential pair can be expressed:

$$g_{m,tot} = \frac{g_{m1}g_{m2}}{g_{m1} + g_{m2}} = \frac{g_m}{2}\left[1 - (\Delta V_{GS}/nV_T)^2\right] \simeq \frac{g_m}{2}$$

Hence to a first order approximation, the gain of the stage will not be affected.

4.3 Second stage

The second amplification stage basically has the same topology as the first stage, only all transistors are now in strong inversion. Due to the magnitude of the signals in the second stage, weak inversion operation is not applicable here. Using the EKV model in strong inversion, we find the gain to be:

$$A_{V2} = -\frac{g_{m11}}{2g_{ds13}} = \frac{(W/L)_{11}}{(W/L)_{13}} \cdot \frac{1}{2n}$$

As the transistors are biased in strong inversion, the inherent offset can be tolerated as this will not bring the transistors out of saturation. However, AC-coupling the two stages was done in the experimental setup to minimize the contribution to offset from the first stage.

Thus by choosing the gain to 100 for both stages, a total gain of 80 dB is achieved.

5 EXPERIMENTAL RESULTS

A test chip with the proposed amplifier has been fabricated in a standard digital 0.5 μm, single poly, N-well CMOS process. Table 1 summarizes the measured performance of the amplifier. Fig. 4

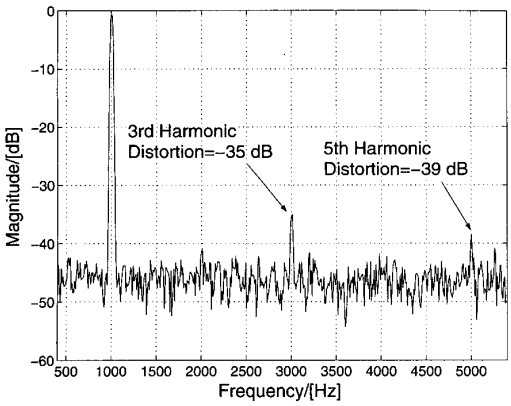


Fig. 4: Measured spectrum at the amplifier output.

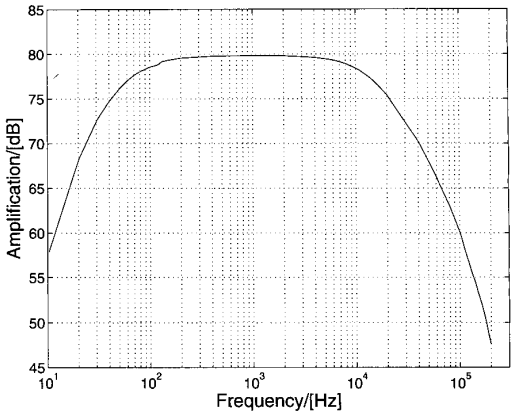


Fig. 5: Amplifier magnitude frequency response.

shows the output power spectrum for an input sine at 1 kHz and amplitude $v_{\text{in}} \simeq 22\,\mu\text{V}_{\text{rms}}$. The total distortion power is -33 dB below the fundamental, equivalent to a THD of 2.2 %, which is acceptable for for our application. The circuit draws 91.5 μA from a 3 V supply, thus consuming 275 μW. The measured output magnitude frequency response is shown in fig. 5. For frequencies $f < 100$ kHz, the amplifier has a CMMR > 87 dB and a PSSR > 84 dB. The equivalent input referred noise over the signal bandwidth is $5\,\text{nV}/\sqrt{\text{Hz}}$, equivalent to a maximum SNR of $\simeq$ 36 dB.

Some gain variation was observed in the test chips. These variations are mainly due to offsets in the common mode voltage of stage 1, which in turn modulates the gain as V_{GS} of M3 is altered.

The simple scheme of using a single fixed bias transistor as load, does not provide high linearity as seen from fig. 4. However, as the signal recovery method only provides an SNR of approx. 29 dB, the THD is within acceptable bounds.

6 CONCLUSIONS

In this paper an amplifier for minute nerve signals was presented. The utilization of weak inversion input transistors, with maximized $g_{\text{m}}/I_{\text{D}}$ to suppress inherent device noise, has been shown to be feasible for the very tight power consumption limits inherent to implantable devices.

For offset cancellation, a continuous-time scheme by tuning bias currents was used. This was done in order to avoid the dynamic offsets introduced by switching schemes, which potentially could be far larger than the minute input signal itself. A condition of this method is that only AC-information needs to be amplified as DC is filtered out.

Finally, the characteristics of a test chip with a prototype amplifier was presented. The prototype exhibits performance within the boundaries put forth. Some harmonic distortion is present, and the SNR is limited. However, for the used signal recovery method, these are within acceptable range.

7 ACKNOWLEDGMENT

This work was supported by the Danish Medical Research Council.

REFERENCES

[1] M. K. Haugland, *Natural Sensory Feedback for Closed-loop Control of Paralyzed Muscles*. Ph.D. thesis, University of Aalborg, 1994.

[2] A. Lickel, *Restoration of Lateral Hand Grasp in a Tetraplegic Patient Applying Natural Sensory Feedback*. Ph.D. thesis, University of Aalborg, 1998.

[3] G. Gudnason, E. Bruun and M. Haugland, "An implantable mixed analog/digital neural stimulator circuit", *IEEE ISCAS*, Orlando, FL., May 30-June 2 1999, pp. 375-8.

[4] K. Papathanasiou and T. Lehmann, "An implantable CMOS nerve signal conditioning system for recording nerve signals with cuff electrodes", *IEEE ISCAS*, Geneva, May 28-31, 2000, pp. 281-4.

[5] K. G. Lamb, S. J. Sanchez and W. T. Holman, "A low noise operational amplifier design using subthreshold operation", *IEEE Midwest Symp. Circuits Syst.*, Notre Dame, IN, August 9-12, 1998, pp. 35-8.

[6] J. Chang, A. A. Abidi and C. R. Viswanathan, "Flicker noise in CMOS transistors from subthreshold to strong inversion at various temperatures", *IEEE Trans. On Electron Devices*, vol. 41, 1994, pp. 1965-71.

[7] C. C. Enz, "The EKV Model: A Most Model Dedicated to Low-Current and Low-Voltage Analogue Circuit Design and Simulation", *Low-Power HF Microelectronics A Unified Approach* (G. A. S. Machado, ed.), IEE, 1996, pp. 247-99.

A distributed transducer system for functional electrical stimulation

Gunnar Gudnason, Jannik H. Nielsen and Erik Bruun
Ørsted-DTU, Technical University of Denmark, DK-2800 Lyngby
gg@oersted.dtu.dk

Morten Haugland
Center for Sensory-Motor Interaction, Aalborg University, DK-9220 Aalborg

Abstract

Implanted transducers for functional electrical stimulation (FES) powered by inductive links are subject to conflicting requirements arising from low link efficiency, a low power budget and the need for protection of the weak signals against strong RF electromagnetic fields.

We propose a solution to these problems by partitioning the RF transceiver and sensor/actuator functions onto separate integrated circuits. By amplifying measured neural signals directly at the measurements site and converting them into the digital domain before passing them to the transceiver, the signal integrity is less likely to be affected by the inductive link. Neural stimulators are affected to a lesser degree, but still benefit from the partitioning.

As a test case, we have designed a transceiver and a sensor chip which implement this partitioning policy. The transceiver is designed to operate in the 6.78MHz ISM band and consumes approximately 360μW. Both chips were implemented in a standard 0.5 μm CMOS technology, and use a 3 V supply voltage.

1. Introduction

The subject of this article is a solution for several problems which plague designers of inductively powered biomedical implants for functional electrical stimulation. Inductive links are commonly employed to power and control such implants since they provide a wireless connection, which is desirable since percutaneous wires provide an infection path into the body. Inductive links also eliminate the need for an implanted battery, which might eventually need to be replaced. The price which must be paid for these advantages is the relatively high power dissipation in the external (non-implanted) apparatus due to the low efficiency of the link, and a degradation of the signals under examination, because of the high electromagnetic fields. While the RF signals normally lie far outside the biological signal band, they can easily desensitize an amplifier designed for a 10–20μV signal range.

The signal amplitude quoted above is typical for signals obtained using cuff electrodes [1]. Stimulators connected to cuff electrodes generate much larger signals, so they are not affected to the same extent as sensors by the external RF field.

We propose therefore a physical partitioning of the signal processing functions, in order to limit the transmission distances for sensitive signals, and to allow optimal placement of critical functions of the system. The conflicting design criteria which can be accommodated to a large extent by partitioning are:

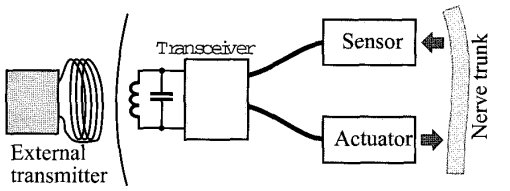


Figure 1. A partitioning example for a FES system containing a transceiver chip and sensor and actuator chips. The last two are not necessarily connected to the same nerve trunk.

- The RF part of the system should be close to the skin surface for better power transmission, and also for better link bandwidth. The proximity may also be an advantage for surgical access.
- The transducers should be placed close to the active sites, which can be relatively deep inside the body.
- Long routing of transducer signals should be avoided, as the strong electromagnetic field will induce an RF interferer overlaid on the desired signal. Physical separation of the transducer from the RF link also reduces the problems, since the field falls off rapidly with distance (as $1/r^3$ in the far-field limit).
- The system must contain a considerable amount of digital logic, for control, timing, and buffering of data. The logic will inevitably couple switching noise onto the supplies and into the substrate. Using low-noise logic solutions like current-steering logic [3] can eliminate the problem, but the static supply current of CSL and other ultra-low noise logic families makes them unsuitable as the system complexity passes a certain point. Placing the sensors on separate chips isolates them from the logic supply noise.

One option which was examined, and which alleviates some of the problems above, is to place only the receiver

0-7803-7057-0/01/$10.00 ©2001 IEEE.

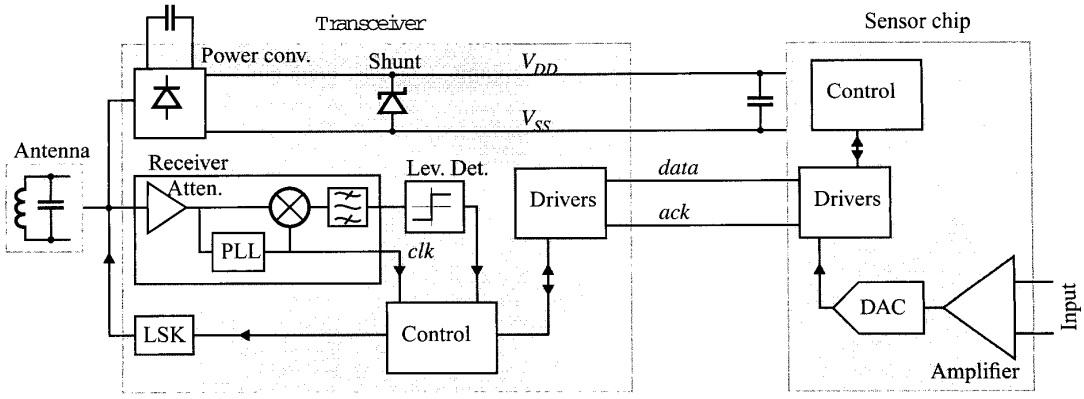


Figure 2. A simplified block diagram of the system, showing the main parts of the transceiver and sensor chips. Differential signals are shown with a single line for clarity.

coil close to the skin surface, and route the antenna signal to a single integrated transceiver/transducer. This is however not practical since biocompatible interconnect solutions like Cooper cable[1] have a high impedance at DC and RF.

We adopted a solution with simple digital communication protocols between the central transceiver and control chip, and the outlying transducer chips. The protocol was designed so that it could be accommodated along with the power supply wires in a simple 4-wire flexible cable. In this specific test case, we implemented a simple system with one transceiver chip and one sensor chip. The sensor chip includes an amplifier for cuff electrode signals, which is itself the subject of another article [2], an AD converter and bus interface logic. The transceiver includes a direct conversion receiver (actually a homodyne), a load modulation circuit for transmission of data out of the system, control logic, supply regulators and references.

2. System description

The following is a description of the transceiver chip and the aspects relating to communication between it and the other parts of the system. Since the main focus of this article is on the way in which partitioning of functions can solve some of the problems in an implanted system of this type, the sensor chip will only be described inasmuch as it relates to the top-level design of the system. The principle of partitioning applies to any transducer type that is relevant in an implanted system, so the internals of the sensor chip will not be emphasized. Figure 2 shows a top-level diagram of the system.

2.1. The link

The use of inductive links for power and data transmission is described in great detail elsewhere, see for example [5, 6].

By separating the transceiver chip and the antenna, it is possible to reduce the distance between the external transmitter antenna (a tuned LC circuit) and the internal one to 15–20mm. The coupling coefficient for normal coil geometries at this distance is on the order of 0.01–0.05. This permits far better power transfer than if the transceiver were placed together with the transducer, and makes load modulation a viable method to extract information from the system. Systems with smaller coupling coefficients must resort to active transmission of signals to the outside [4], which increases the power consumption.

The chosen modulation method is PAM with Manchester encoding and a modulation index of approximately 0.2. This is compatible with high-efficiency class D or E transmitter configurations [7, 8]. The carrier frequency is 6.78MHz which coincides with one of the ISM (industrial, scientific and medical) bands. The target bit rate is 50kbit/s.

2.2. The power supply

The power is extracted from the RF carrier by means of a full-wave bridge rectifier, which is implemented by using diode-connected P-channel MOS transistors in a common N-well, as shown in figure 3. The standard CMOS process does not offer high-quality floating diodes, so another solution must be chosen. The available elements are three types of p-n junction and diode-connected MOS transistors, but these all suffer from high substrate currents and/or parasitic elements which divert some of the current from its intended path. The PMOS rectifier bridge is accompanied by parasitic vertical PNP transistors which can divert some of the input current to V_{SS} instead of to V_{DD}.

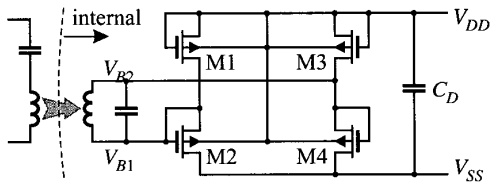


Figure 3. The power conversion circuit.

[1] Produced by Finetech Medical Ltd.

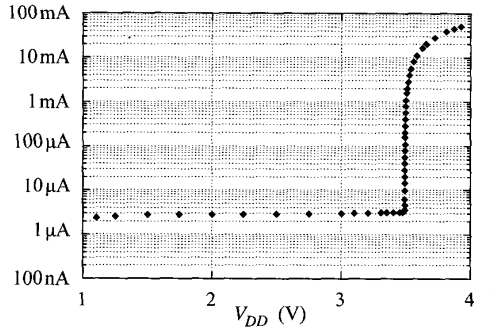


Figure 4. The measured shunt regulator current.

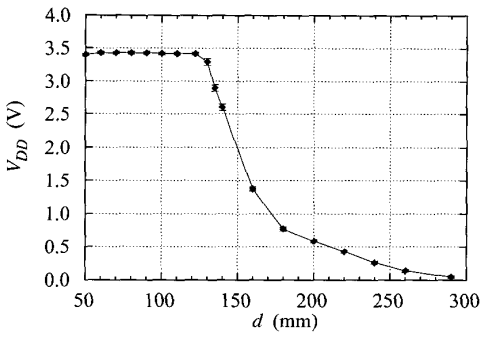


Figure 5. The measured on-chip supply voltage as a function of the separation between the transmitter and transceiver antennas.

We have solved this problem by dimensioning the PMOS transistors so that they are biased in weak inversion over the entire operating range. Since the nominal threshold voltage is about 0.6 V, there is a range of drain-gate voltages for M1 in figure 3 (corresponding to the emitter-base potential in the parasitic PNP) where the MOS current dominates the current through the bipolar transistor by a large factor. The current-handling capability of the rectifiers can be increased for a given maximum parasitic/main current ratio, by increasing the width (and area) of the transistors. The main penalty is an increased input capacitance. Our experimental data show that the parasitic current ratio is smaller than 0.005 for a system supply current of 200 μA and an input capacitance of 2 pF. All the transistors in figure 3 have $W/L = 800/0.5$ in micrometers.

The transducer chip contains an active shunt regulator which provides protection against excessive input power. The transmitted power is a very strong function of antenna spacing, so an unprotected circuit can easily be burned out. The shunt circuit is basically a pass transistor between V_{DD} and V_{SS}, and a feedback loop which compares V_{DD} to a bandgap reference voltage. As the supply voltage approaches the trigger point (which was set equal to the maximum supply voltage for the technology), the regulator begins shunting current from the supply, and the current changes by 4 decades over a short supply voltage interval (see figure 4).

The performance of the power supply conditioning system is shown in figure 5. The conversion of the RF power to a DC supply voltage and the overvoltage protection is carried out mostly on-chip, with the only external component being the energy storage capacitor C_D. The transmitter used in these measurements was running at a relatively high power level, witnessed by the fact that the full supply voltage is reached at a separation of 120 mm. Despite the strong dependence of the magnetic field amplitude on the separation d, the on-chip shunt succeeds in maintaining the supply voltage at the nominal level at all separations.

The transceiver chip provides the supply voltage for the other chips in the system. Two strands out of four in the Cooper cable have been assigned to V_{SS} and V_{DD}. Because of the helically wound construction of the cable, and the Pt-Ir composition, the impedance is relatively high at all frequencies. To provide a stable supply voltage for the outlying chips, it is therefore necessary to add decoupling capacitors to the supply lines at the transducer ends. A stimulator chip can require relatively large current pulses from the supply while delivering a stimulus.

2.3. Transceiver

The receiver part of the transceiver chip is a direct-conversion receiver. Since the input signal has a larger amplitude than the supply range, an attenuator with a gain of approximately 0.1 is inserted in the signal path. The input signal is tracked by a PLL which also provides the system clock (the block diagram does not show a clock divider circuit which divides the clock frequency down to 1.7 MHz).

The output of the PLL and the input signal are put through a mixer, and the mixer output is filtered in a 4th-order differential G_m–C filter. The filter output is directed to a mixed analog-digital level detector, which detects the sign of the transmitted data bits. Instead of using a partly analog level detector, the normal processing method would be to sample the filter output for further digital signal processing. This does however require more sophisticated digital circuits than we were willing to implement, and possibly requires more supply current.

The target bit rate for the system was 50kbit/s, and the receiver is designed to handle up to 100kbit/s.

Data is transmitted out of the system by load modulation. The reflected impedance seen by the external transmitter is varied by changing the load seen by the secondary LC circuit. A switch is connected between the terminals of the antenna, and by closing the switch, a maximum change in the reflected impedance is obtained. This simple scheme has the disadvantage of stopping power transfer to the system during load modulation, and other load modulation circuits have been designed to avoid this [9]. The duty cycle of the switch closure is however so low in our case that the reduction in average power transfer is small.

2.4. Interchip communication

The cable type that we used as a reference for our interconnects is, as mentioned before, a 4-strand biocompati-

ble cable type by the name of Cooper cable. In addition to being biocompatible, the cable is wound in a helical pattern and embedded in silicone, so it can be stretched. The stretching reduces the probability of damage to the surrounding tissue.

Because of the biological constraints that the cable must fulfill, it does not have very good electrical properties, and the use of the cable must be adjusted accordingly. We measured the series resistance at DC of a representative length of the cable, and found it to be 210 Ω/m. The capacitance between any two wires varies from 20–80 pF/m because of the lack of symmetry, and the capacitance of each wire to the surroundings is about 50 pF/m.

No clock signal is sent across the connection between the chips. Instead, we use an asynchronous handshake mechanism to control data transfer, and an internal oscillator in the sensor chip to provide a time base.

The line drivers are class AB circuits which have a quiescent current consumption of 5 μA each, and can slew the line voltages with a 50 μA current. This type of driver was chosen instead of faster types because this is more than sufficient for the purpose, and by limiting the slewing currents the supply transients are reduced. The high and low voltages on the signal lines are 1.0 V and 0.3 V respectively instead of the full supply range, again to reduce supply transients and power consumption.

3. Measurement results

Most of the basic functions of the transceiver and sensor chips have been measured, and they behave according to specifications.

The results for the power conversion circuit show that the use of PMOS transistors in weak inversion as rectifiers is an ideal solution for pure CMOS technologies. Current throught the substrate are eliminated by placing everything in an N-well, and the effect of parasitic bipolars are all but eliminated.

The active shunt regulator and other regulators and references on the chip are up to the design criteria. Specifically, the shunt regulator consumes negligible supply current within the normal voltage range, with a very sharp rise in current as the trigger point is exceeded. Previous systems use passive zener-based regulators, whose gradual I-V characteristic provides insufficient overvoltage protection for low-voltage CMOS technologies.

The data link to the transceiver chip was tested by using a Class D transmitter driving the inductive link, with a data rate of 50kbit/s and 20% ASK modulation and Manchester encoding. The data transfer functions according to the specifications, and higher data rates can easily be supported with minor modifications.

4. Conclusion

We have demonstrated a partitioning scheme for implanted sensor and actuator devices which places the signal processing functions where they are needed. By using this scheme, it is possible to provide better isolation of weak biological signals from strong external disturbances, while simultaneously reducing the overall power consumption by placing the transceiver closer to the external interface. The communication between separate parts of the system has been adapted to existing biocompatible interconnect methods.

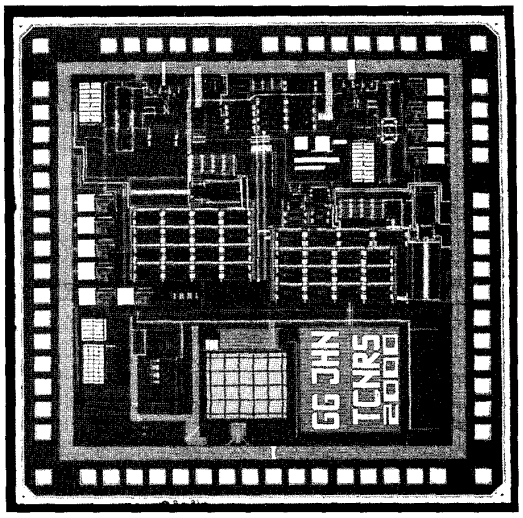


Figure 6. Chip die microphotograph.

We have implemented a simple system consisting of a transceiver and a single sensor chip, but the concept can easily be extended to more general combinations of sensors and actuators, in order to created a complete neural stimulation system with a closed feedback loop.

[1] G. E. Loeb and R. A. Peck, "Cuff electrodes for chronic stimulation and recording of peripheral nerve activity". *J. Neuroscience Methods*, **64** (1996), 95–103.

[2] J. H. Nielsen and T. Lehmann, "An implantable CMOS amplifier for nerve signals", *Proc. ICECS 2001*.

[3] H.-T. Ng and D. J. Allstot, "CMOS current-steering logic for low-voltage mixed-signal integrated circuits", *IEEE Trans. VLSI*, **5** (1997), pp. 301–308.

[4] M. Nardin and K. Najafi "A multichannel neuromuscular microstimulator with bi-directional telemetry", *1995 Conf. on Solid-State Sensors and Actuators*, pp. 59–62.

[5] D. Galbraith, M. Soma and R. White, "Wide-band efficient inductive transdermal power and data link with coupling insensitive gain", *IEEE Trans. BME*, **34** (1987), pp. 265–275.

[6] N. de N. Donaldson and T. A. Perkins. "Analysis of resonant coupled coils in the design of radio frequency transcutaneous links". *Med. and Bio. Eng. and Computing*, **21** (1983), pp. 612–627.

[7] A. Djemouai, M. Sawan and M. Slamani, "An efficient RF power transfer and bidirectional data transmission to implantable electronic devices", *Proc. ISCAS '99*, pp. II-259.

[8] P. R. Troyk and M. A. K. Schwan, "Closed-loop class E transcutaneous power and data link for microimplants" , *Trans. BME*, **39** (1992), pp. 589-599.

[9] Z. Tang et al. "Data transmission from an implantable biotelemeter by load-shift keying using a circuit configuration modulator". *IEEE Trans. BME*, **42** (1995), pp. 524.

A DESIGN METHODOLOGY FOR POWER-EFFICIENT CONTINUOUS-TIME ΣΔ A/D CONVERTERS

Jannik Hammel Nielsen and Erik Bruun

Ørsted•DTU, Technical University of Denmark
DK-2800 Kgs. Lyngby, Denmark
jhn@oersted.dtu.dk

ABSTRACT

In this paper we present a design methodology for optimizing the power consumption of continuous-time (CT) ΣΔ A/D converters. A method for performance prediction for ΣΔ A/D converters is presented. Estimation of analog and digital power consumption is derived and employed to predict the most power efficient configuration of a CT single-loop ΣΔ ADC. Finally, a 10 bit prototype converter is optimized and simulated using a 0.35 μm CMOS technology. The simulation results of the prototype 1.8 V converter show a SNR better than 65 dB and a spurious-free dynamic range of more than 63 dB, consistent with 10 bits performance. Expected power consumption for the prototype is approx. 170 μW.

1. INTRODUCTION

Todays trend toward high portability of everyday electronic devices ranging from telecommunication to biomedical devices, constitutes a major design challenge. The need for high durability of battery-powered devices coupled with ever-increasing performance requirements only enhances the need for low-power solutions. Another factor is the decreasing voltage supply driven by technology scaling. This is due to the fact that low-voltage operation of analog circuits requires high power consumption in order to retain sufficient dynamic range.

An important building block in many devices is the A/D converter (ADC). Of the many possible variations of ADC topologies, the ΣΔ-modulator type remains highly interesting. This topology offers an elegant way to obtain high resolution for low- to moderate input frequencies and yet still maintain modest requirements for the analog building blocks. In the past decade, this ADC topology has shown it's robustness through many implementations using the discrete-time (DT) switched-capacitor (SC) technique. However as supply voltages decrease with technology scaling, the SC technique is becoming less viable.

CT implementations holds promise of lower power consumption and are less dependent on supply scaling than their DT counterparts. The price paid is increased sensitivity towards e.g. clock jitter and other non-idealities.

2. CT ΣΔ-MODULATORS

Employing a CT loop-filter for the ΣΔ ADC realisation holds several advantages. As the sampler is located deep in the loop, implicit anti-aliasing (AA) is native to the CT ΣΔ ADC, often making an explicit AA-prefilter redundant. The necessary gain-bandwidth (GBW) of the integrators is typically a factor of 3 lower than their DT counterparts [2] for a given sampling frequency f_s.

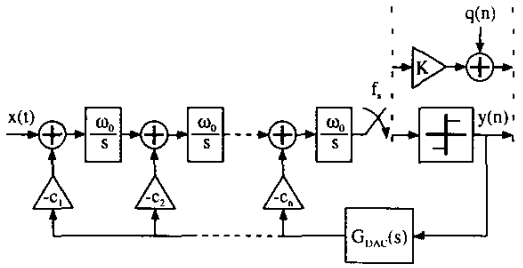
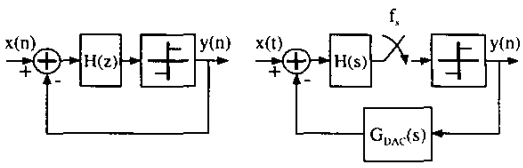


Figure 1: Continuous-Time ΣΔ-modulator.



(a) DT Feedback Filter (b) CT Feedback Filter

Figure 2: Feedback filters in DT and CT modulators.

These properties imply that a significantly lower power consumption can be achieved by using a CT loop-filter over the DT counterpart.

A single-loop CT ΣΔ ADC of order N is shown in fig. 1. This configuration is well suited for low-power as only the first integrator has stringent requirements imposed on it. This is due to the filtering from all other internal nodes to the output. The quantizer is sampled at f_s from which we can define the oversampling ratio: OSR $= f_s/2f_b$, where f_b is the baseband frequency.

2.1. DT to CT Mapping

In order to be able to obtain fast and yet precise simulation of the CT ADC performance, a DT/CT equivalence can be established. This also allows us to use the design methods usually employed for DT ΣΔ ADCs. In fig. 2, a linear model of the ADC is shown from a loopfilter perspective in both a DT and CT version. The

This work was supported by the Danish Medical Research Council

0-7803-7761-3/03/$17.00 ©2003 IEEE

N	
2	$\left(\frac{\omega_0}{f_s}\right)^2 \left\{c_1 = \frac{d_1+d_2}{2},\ c_2 = d_1 + d_2\right\}$
3	$\left(\frac{\omega_0}{f_s}\right)^3 \left\{c_1 = \frac{2d_1-d_2+2d_3}{6},\ c_2 = d_3 - d_1,\ c_3 = d_1 + d_2 + d_3\right\}$
4	$\left(\frac{\omega_0}{f_s}\right)^4 \left\{c_1 = \frac{3(d_4-d_1)+d_2-d_3}{12},\ c_2 = \frac{11(d_1+d_4)-d_2+d_3}{12},\ c_3 = \frac{3(d_4-d_1)+d_3-d_2}{2},\ c_4 = \sum_{n=1}^{4} d_n\right\}$

Table 1: Mapping of DT to CT filter coefficients.

signal transfer function (STF) and noise transfer function (NTF) for fig. 2(a) are respectively:

$$\mathrm{STF}(z) = \frac{KH(z)}{1+KH(z)} \approx 1, \quad \mathrm{NTF}(z) = \frac{1}{1+KH(z)} \quad (1)$$

An often used approach for the design of $\Sigma\Delta$-modulators is to choose a desired NTF (e.g. butterworth etc.) [1]. From the NTF, the feedback filter $H(z)$ can be derived, determining the feedback coefficients. Since the CT modulator in fact constitutes a sampled system, a CT feedback filter equivalent to the DT filter can be obtained by requiring that the outputs are equal at the sampling instant. Referring to fig. 2, we thus have that:

$$H(z) = \mathcal{Z}\left\{H(s)G_{\mathrm{DAC}}(s)\right\} = \widehat{H(z)} \quad (2)$$

where $H(s)G_{\mathrm{DAC}}(s) = \widehat{H(s)}$, constitutes the combined feedback filter. Assuming a zero-order-hold feedback DAC with a NRZ feedback waveform, the s-domain transfer function of the DAC is: $G_{\mathrm{DAC}}(s) = [1 - \exp(-sT)]/s$. For a single-loop modulator of order N, the CT and DT feedback filters can be written respectively:

$$H(s) = \sum_{n=1}^{N} c_n \left(\frac{\omega_0}{s}\right)^n, \quad H(z) = \sum_{n=1}^{N} d_n z^{-1} \left(\frac{z}{z-1}\right)^n \quad (3)$$

Taking the z-transform of each order of $H(s)$ can be shown to be:

$$\mathcal{Z}\left\{H(s)\right\} = \sum_{n=1}^{N} \frac{c_n (\omega_0 T)^n}{n!} \frac{B_n(z)}{(1-z^{-1})^n} \quad (4)$$

where $B_n(z)$ is a polynomial in z of order n. By comparing the coefficients of z in $\widehat{H(z)}$ and $H(z)$, a set of linear equations result from which the mapping between the DT d-coefficients and the CT c-coefficients can be established. In table 1, the mapping for modulators of orders 2 through 4 is given.

2.2. Performance Prediction

The quantizer in the $\Sigma\Delta$ constitutes a highly nonlinear element. As shown in fig. 1, the quantizer can be modelled as a gain factor K, and an additive noise source $q(n)$.

In [1] a quasilinear model of the 1-bit quantizer is proposed. If a signal with a sufficiently high amplitude is applied at the modulator input, the modulator will become unstable. The maximum stable amplitude (MSA) can be shown to be a function of the feedback filter cut-off frequency f_c. Assuming a heavily oversampled input signal, the output power: $\sigma_y^2 = 1 - m_y$, can be attributed

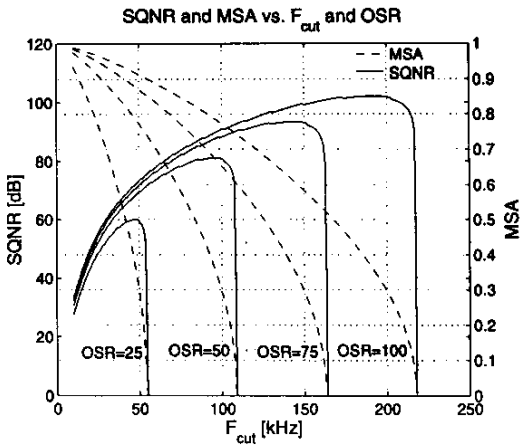


Figure 3: SQNR and MSA in a a 3rd order $\Sigma\Delta$ ADC vs. f_c for increasing OSR.

solely to the quantization noise σ_q^2. Hence we can define a noise amplification factor, $A(K)$ [1]:

$$A(K) = \frac{1-m_y^2}{\sigma_q^2} = \int_0^{\pi} \left|\mathrm{NTF_K}\left(e^{j\omega}\right)\right| d\omega = \sum_{n=1}^{\infty} |\mathrm{ntf_K}(n)|^2 \quad (5)$$

where Parseval's theorem has been used to obtain the time-domain representation. m_y is the mean of the modulator output. Due to the high feedback filter gain, we have that $m_x \approx m_y$.

As the quantization noise is in fact not an independent signal source, but generated by the input signal itself, we can estimate σ_q^2 from m_x and the quantizer input probability density function (PDF) [1]. For higher order modulators (N>2), the quantizer input is well approximated by a gaussian PDF. For a gaussian PDF, we have [4]:

$$\sigma_q^2 = 1 - m_y^2 - \frac{2}{\pi} \exp\left[-2\left(\mathrm{erf}^{-1}(m_y)\right)^2\right] \quad (6)$$

where erf^{-1}, is the inverse error function.

As the lowest amount of quantization noise is attained at the MSA, the minimum value of $A(K)$ yields the MSA. Hence by combining eq. (5-6) we can obtain an estimate of the MSA .

However, as $A(K)$ is an increasing function of the NTF cut-off frequency [1], the MSA is inverserly proportional to f_c and hence a trade-off exists between a high f_c, i.e. high quantization noise supression, and a high MSA, i.e. high signal power.

Due to the intrinsic minimization of the quantization noise, an equilibrium quantizer gain K_{eq}, can be estimated by the K which minimizes $\|\mathrm{ntf_K}\|_2^2/K^2$. The obtained K_{eq} determines both $\mathrm{NTF_K}(z)$ and $A(K)$. From eq. (5), the quantization noise for zero input can be found: $\sigma_q^2 = 1/A(K_{\mathrm{eq}})$. For a base-band frequency of f_b, the signal to quantization-noise ratio (SQNR) for a sinusoid at MSA can be estimated:

$$\mathrm{SQNR} = \frac{\mathrm{MSA}^2}{2A(K_{\mathrm{eq}})f_s} \int_0^{f_b} \left|\frac{1}{1+K_{\mathrm{eq}}H(e^{j\omega})}\right|^2 df \quad (7)$$

In fig. 3 the predicted MSA and SQNR for a third order $\Sigma\Delta$ is shown vs. f_c for different OSRs. The abrupt drop-off of the SQNR

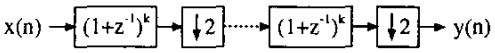


Figure 4: FIR2 combfilter structure.

curves occur as the MSA approaches zero. From fig. 3, it is seen that the maximum SQNR is obtained at a MSA of approx. 0.3. This low value is due to the fact that the noise suppression grows "faster" for increasing f_c than the MSA decreases.

2.3. Decimation filter

A popular choice for decimation filters is the combfilter which can be implemented without using computationally expensive multipliers. The combfilter transfer function is given by [3]:

$$H_{\mathrm{com\,b}}(z) = \left(\frac{1-z^{-D}}{1-z^{-1}}\right)^k = \prod_{n=0}^{\log_2 D-1}\left(1+z^{-2^n}\right)^k \qquad (8)$$

where D is the decimation factor and k is the order of the filter. The second form of eq. (8) can be implemented by a series of FIR filters each decimating by a factor of two as shown in fig. 4.

The wordlength of the registers at the ouput of each FIR stage is equal to $(W_0 + ki)$ bits, where W_0 is the number bits at the filter input. Hence we have a filter configuration where the full sampling frequency is only applied to the first stage, which only has a wordlength of a few bits. Whereas the other stages, with increasing wordlength, are clocked at decimated frequencies giving a power-efficient decimation filter [3]. For a $\Sigma\Delta$-modulator of order N, a decimation filter of order N+1 suffices for out-of-band quantization noise suppression.

3. CT $\Sigma\Delta$-MODULATOR OPTIMIZATION

As was previously seen, the MSA is inversely proportional to the feedback filter cut-off frequency. However the SQNR curves suggests that the maximum SQNR is achieved for low values of the MSA. A very low MSA points towards a very high power consumption in the analog implementation, since in order to obtain the desired dynamic range, very low noise analog blocks must be developed, indicating high biasing currents.

Clearly, a trade-off exists between the analog power consumption, the MSA and the SQNR.

3.1. Power estimation

The core elements of an analog integrator consists of an integrating capacitor and a regulating active element, e.g. an op-amp. The power consumption of the active element is then determined by the GBW necessitated by the integrator specifications, e.g. max. input frequency. Whereas the size of the capacitor is determined by noise limitations hereby setting dynamic power consumption.

We have previously derived an expression for the quantization noise σ_q^2. Referring the quantization noise to a reference voltage V_{ref} and taking the combined filtering of $\mathrm{NTF}(z)$ and $H_{\mathrm{com\,b}}(z)$ into account, we have the resulting quantization noise power:

$$P_{\mathrm{Q}} = \frac{V_{\mathrm{ref}}^2}{A(K_{\mathrm{eq}})f_s}\int_0^{\pi}\left|\frac{H_{\mathrm{com\,b}}\left(e^{j\omega}\right)}{1+K_{\mathrm{eq}}H\left(e^{j\omega}\right)}\right|^2 d\omega \qquad (9)$$

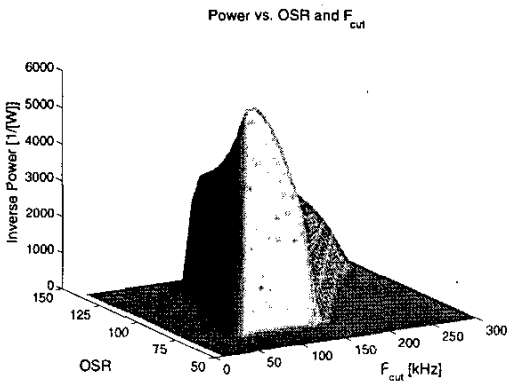


Figure 5: Inverse power consumption of a 3rd order $\Sigma\Delta$ ADC.

The thermal noise accumulated on the integrating capacitor is:

$$P_{\mathrm{Th}} = \frac{\mathrm{NEF}\cdot kT}{C_{\mathrm{int}}} \qquad (10)$$

where NEF is the Noise Excess Factor taking the added circuitry noise into account.

The total amount of noise allowable for a resolution of b bits is given by:

$$P_{\mathrm{N}} = \frac{(\mathrm{MSA}\cdot V_{\mathrm{ref}})^2}{2\cdot 10^{(6.02b+1.76)/10}} \qquad (11)$$

Combining equations (9-11), we can solve for the minimum allowable capacitor size:

$$C_{\mathrm{int}} = \mathrm{NEF}\cdot kT\left[\frac{(\mathrm{MSA}\cdot V_{\mathrm{ref}})^2}{2\cdot 10^{(6.02b+1.76)/10}} - P_{\mathrm{Q}}\right]^{-1} \qquad (12)$$

The static power consumption is determined by the required GBW of the regulating cell. A first order approximation of the necessary bias current I_{B}, can be found:

$$I_{\mathrm{B}} = \frac{\mathrm{GBW}^2 C_l^2}{2\beta} \qquad (13)$$

C_l is the loading capacitance and β is the beta of the driving MOS transistor.

As the output of the first integrator goes unattenuated to the ADC output, the requirements of this block are as strict as for the overall system. The outputs of the other blocks however, are filtered and the requirements for these can hence be significantly loosened. The GBW of the integrators can be lessened, implying a lower I_{B} and more noise can be tolerated on the integrating capacitors, allowing for a smaller capacitors C_{int}.

Let P_c denote the quantizer power per frequency, the analog power is then approximated by:

$$P_{\mathrm{an}} \approx 4V_{\mathrm{ref}}f_s\sum_{i=1}^{N}C_{\mathrm{int,i}} + mV_{\mathrm{DD}}\sum_{i=1}^{N}I_{\mathrm{B,i}} + P_c f_s \qquad (14)$$

m is the number of current branches in the regulating element and the summations indicate the scaling of bias currents and capacitors for the cascade of integrators.

Power Consumption	171 μW
Supply Voltage	1.8 V
Sampling Frequency	1.5 MHz, (OSR=75)
Feedback filter cut-off frequency, f_c	100 kHz
SNR	> 65 dB
SFDR	> 63 dB
Technology	0.35 μm CMOS

Table 2: Test design performance summary.

The digital circuitry when using the FIR2 decimation filter structure is confined to D-flip flops, half- and fulladders. For a given clock frequency we have:

$$P_{\mathrm{dig}} = \alpha \sum_{i=1}^{\log_2 D-1} \frac{f_{\mathrm{clk}}}{2^{i-1}} \left(P_{\mathrm{fa}} + P_{\mathrm{ha}} + P_{\mathrm{dff}}\right) \qquad (15)$$

The frequency dependent power consumption of the digital blocks can be found in the datasheets of the employed technology. α denotes the activity factor. The total power of the modulator is thus given by: $P_{\mathrm{tot}} = P_{\mathrm{an}} + P_{\mathrm{dig}}$. The inverse power is plotted for a third order modulator in fig. 5, where 10 bits of accuracy is required and $V_{\mathrm{ref}} = 200\,\mathrm{mV}$. A clear peak is observed, indicating that a global optimum with regards to power efficiency exists. Hence an optimum f_{clk} and f_c can be extracted for a modulator of order N.

4. TEST CASE

In order to verify the developed design method, a prototype ADC requiring 10 bits performance was simulated. In the table below, the projected power consumption of $\Sigma\Delta$ ADCs of order 2-5 are shown, indicating that a third order $\Sigma\Delta$ is suitable for the specifications.

N	2	3	4	5
P [μW]	308	164	195	257

The prototype $\Sigma\Delta$ ADC was simulated at transistor level using a 0.35 μm CMOS process. The integrators were implemented as $G_m - C$ cells, where the G_m was obtained by placing a MOST in the linear region and the necessary GBW given by the speed requirements for the regulating loop. An example output spectrum with a tone at the MSA of the prototype modulator is shown in fig. 6. The third harmonic is below -63 dB, indicating 10 bits spurious-free performance is obtained. The fifth harmonic is not visible as it is buried beneath the noise floor. The even order harmonics are suppressed due to the fully differential structure. Overlaid the spectrum, the employed decimating comb-filter transfer function is shown. A summary of the prototype $\Sigma\Delta$-modulator performance is given in table 2.

A comparison with an equivalent SC $\Sigma\Delta$ ADC is in it's place. The number of time constants n, necessary for first order settling behaviour to b bits can be found from: $n > (b+1)\ln(2)$, which resolves to $n \approx 8$ for 10 bits performance. Using a first order model of an op-amp, we have that: $GBW = g_m/C_{\mathrm{int}} = 1/\tau$, where the transconductance is: $g_m = \sqrt{2\beta I_B}$. Requiring the output to settle within half a clock-period, we have that: $1/\tau = 16 f_s$. Hence we can approximate the necessary bias current:

$$I_B \approx \frac{(16 f_s C_{\mathrm{int}})^2}{2\beta} \qquad (16)$$

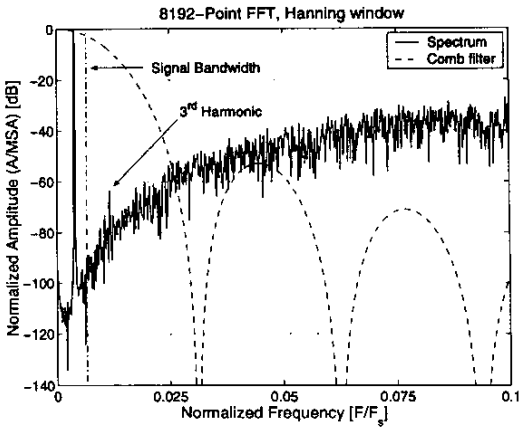


Figure 6: Output spectrum of the test design $\Sigma\Delta$ ADC.

Using the same capacitor sizes and parameters as in the prototype ADC and employing a folded cascode op-amp, the static power consumption of the integrators can be found to be: $P_{\mathrm{stat,DT}} \approx 335\,\mu\mathrm{W}$. Whereas the corresponding CT consumption in the prototype ADC is: $P_{\mathrm{stat,CT}} \approx 111\,\mu\mathrm{W}$. Hence an estimated 2/3 of the static integrator power consumption has been saved by employing a CT structure.

5. CONCLUSION

A design methodology for CT single-loop $\Sigma\Delta$ ADCs has been presented in this paper. The theory necessary for performance prediction was shown and a method for estimation of power consumption. For a given set of specificatons, the theory allows an optimum $\Sigma\Delta$ order, feedback filter cut-off frequency and sampling frequency to be extracted. Finally, a test case was specified and optimized in order to verify the theory. The prototype has performance in good correspondance with the expected results from the optimization routine, verifying the validity of the proposed methodology. An estimate of the power consumption of an equivalent DT cascade of integrators showed that power savings of approx. 2/3 had been obtained by using the CT configuration.

6. REFERENCES

[1] L. Risbo, "Stability Predictions for High-Order $\Sigma - \Delta$ Modulators Based on Quasilinear Modeling". *Proc. ISCAS '94*, May-Jun. 1994, Vol. 5, pp. 361-364.

[2] F. Gerfers, Y. Manoli, "A 1.5V Low-Power Third Order Continuous-Time Lowpass $\Sigma\Delta$ A/D Converter". *Proc. of the Int. Symposium on Low Power Electronics and Design*, Jun. 2000, pp. 219-221.

[3] Aboushady, et. al., "Efficient Polyphase Decomposition of Comb Decimation Filters in $\Sigma\Delta$ Analog-to-Digital Converters". *IEEE Trans. Circuits and Syst. II*, vol. 48, pp. 898-903, Oct., 2001.

[4] S. H. Ardalan and J. J. Paulos, "An analysis of nonlinear behaviour in delta-sigma modulators". *IEEE Trans. Circuits and Syst.*, vol. 34, pp. 593-603, Jun. 1987.

Analog Integrated Circuits and Signal Processing, 36, 153–164, 2003

An Implantable CMOS Amplifier for Nerve Signals

JANNIK HAMMEL NIELSEN[1] AND TORSTEN LEHMANN[2]

[1]*Ørsted•DTU, Technical University of Denmark, DK-2800 Kgs. Lyngby*
[2]*Cochlear Ltd., 14 Mars Road, Lane Cove 2066, NSW, Australia*
E-mail: jhn@oersted.dtu.dk; tlehmann@cochlear.com.au

Received April 30, 2002; Revised January 29, 2003; Accepted April 18, 2003

Abstract. In this paper, a low noise high gain CMOS amplifier for minute nerve signals is presented. The amplifier is constructed in a fully differential topology to maximize noise rejection. By using a mixture of weak- and strong inversion transistors, optimal noise suppression in the amplifier is achieved. A continuous-time current-steering offset-compensation technique is utilized in order to minimize the noise contribution and to minimize dynamic impact on the amplifier input nodes. The method for signal recovery from noisy nerve signals is presented. A prototype amplifier is realized in a standard digital 0.5 μm CMOS single poly, n-well process. The prototype amplifier features a gain of 80 dB over a 10 kHz bandwidth, a CMRR of more than 87 dB and a PSRR greater than 84 dB. The equivalent input referred noise in the bandwidth of interest is 4.8 nV/$\sqrt{\text{Hz}}$. The amplifier power consumption is 275 μW, drawn from a power supply; $V_{\text{DD}} = -V_{\text{SS}} = 1.5$ V.

Key Words: neural sensor, implantable microsystems, FES, ENG

1. Introduction

Thousands of individuals sustain damages to the central nervous system, e.g. spinal chord injury, stroke, etc. which potentially results in paralyzed limbs. However, such damages usually leave the nerves and muscles in the limbs unaffected. In recent years, biomedical research has focused on retrieving nervous information from the natural sensors of the body and using this to activate paralyzed muscles by Functional Electrical Stimulation (FES) [1,2]. Successful trials have been conducted using external devices, correcting dropfoot in early schlyrosis patients [1], and restoring basic hand functions [2], by using natural sensor feedback.

Our involvement comes from developing implantable ASICs for both sensing neural information and for stimulating muscles through FES. An implantable stimulator has been developed and is reported in [3]. Developing implantable devices, reduces the risk for infection as skin continuity is ensured. Also, the need for external sensors which need calibration, may be bulky and are subject to mechanical stress limiting device life-span [1], is eliminated by using the natural sensors of the body.

This paper reports the implementation of the proposed MOSFET amplifier in [4], for amplification of minute nerve signals.

2. Proposed Sensing Implant

The overall sensing system for implantation to be developed is shown in Fig. 1.

The electrical contact with the nerve is obtained using a so-called *cuff electrode* [1,2], shown in Fig. 2. The cuff electrode, which has a length of 10–20 mm, is typically fabricated using silicone as the insulating material and is fitted with conducting electrodes sewn on the inside of the cuff tube for signal pickup. The cuff is placed loosely around the nerve trunk in order not to cause any damage to the nerve. The initial contact resistance directly after implantation rises in the first few months after implantation to about 5 kΩ as tissue regrowth takes place. The neural signal recorded from the cuff is termed the electroneurogram (ENG) and when recorded in a tripolar fashion, provides a differential nerve signal of about ± 10 μV.

Control signals, data transmission and power supply are conveyed through the inductive link, dictating a small supply current, in the order of a few hundred μA.

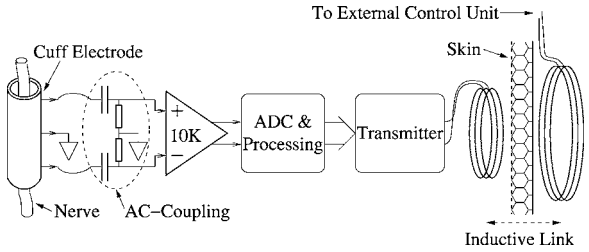


Fig. 1. Proposed nerve signal recovery implant.

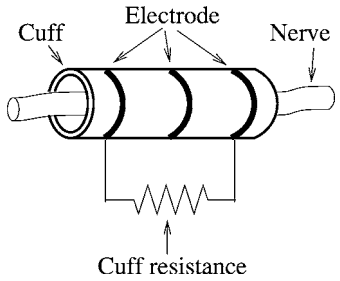


Fig. 2. Silicone cuff and tripolar electrode configuration.

The signal from the cuff is AC-coupled to remove inherent DC offsets. Due to the very small signal amplitude, it is necessary to preamplify the signal prior to any processing. The preamplifier increases the input amplitude from $\pm 10\ \mu$V to ± 100 mV. As we are not interested in the absolute value of the signal, gain variation of 10–15% can be tolerated. The nerve signals reside primarily in the bandwidth $400 < f_n < 400$ Hz [2], giving the necessary amplifier bandwidth.

In order to suppress artifacts which may arise from the stimulus of nearby muscles, it is of importance that the amplifier has a good common mode rejection ratio (CMRR). An excited muscle evokes a so-called electromyogram (EMG), which primarily will couple to the cuff electrodes as a common mode signal. Furthermore, a good power supply rejection ratio (PSRR) is of importance, as the power supply is inductively coupled from an external source.

After amplification the signal is anti-aliased and A/D converted. Some local processing to recover the nerve signal is then performed and the resulting signal is transmitted to the external control system through the inductive link [10].

3. Signal Recovery

The technique used for signal recovery [1,2], samples the nerve signal at $f_s = 10$ kHz. After A/D conversion, the signal is rectified, integrated and downsampled to $f_r = 20$ Hz, giving the signal envelope. We maintain this digital domain signal recovery technique in our proposed system as this allows us to get external access to the raw digitized amplifier output. This is important for the first experimental implants.

The nerve signal can be described as a zero mean stochastic variable V_n, with variance and mean:

$$\sigma_n^2 = \mathrm{E}\{V_n^2\} \quad \text{and} \quad \mu_n = \mathrm{E}\{V_n\} = 0 \tag{1}$$

The rectified signal V_r and its mean are given by:

$$V_r = \sqrt{V_n^2} \quad \text{and} \quad \mu_r = \mathrm{E}\{\sqrt{V_n^2}\} \tag{2}$$

We can thus find the variance of the rectified signal:

$$\begin{aligned}\sigma_r^2 &= \mathrm{E}\{(V_r - \mathrm{E}\{V_r\})^2\} = \mathrm{E}\{(\sqrt{V_n^2} - \mu_r)^2\} \\ &= \mathrm{E}\{V_n^2\} + \mu_r^2 - 2\mu_r \mathrm{E}\{\sqrt{V_n^2}\} \\ &= \sigma_n^2 + \mu_r^2 - 2\mu_r^2 = \sigma_n^2 - \mu_r^2\end{aligned} \tag{3}$$

Assuming a gaussian distribution of V_n, we can find the mean of V_r:

$$\begin{aligned}\mu_r &= \frac{2\sigma_n}{\sqrt{2\pi}} \int_0^\infty x \exp(-1/2 \cdot x^2)\, dx \\ &= \sqrt{\frac{2}{\pi}} \cdot \sigma_n\end{aligned}$$

Inserting μ_r in equation (3) yields the rectified signal variance:

$$\sigma_r^2 = \sigma_n^2 \left(1 - \frac{2}{\pi}\right) \tag{4}$$

The variance of the downsampled signal is reduced by the ratio of sampling frequencies $k = f_s/f_r$. Thus the SNR of the downsampled signal is:

$$\mathrm{SNR} = 10 \log\left[k \frac{\mu_r^2}{\sigma_r^2}\right] \simeq 29\,\mathrm{dB}, \quad f_s = 10\,\mathrm{kHz}$$

Though this SNR may seem low, it is sufficient to determine the state of a human nerve.

4. Amplifier

A diagram of the designed amplifier is shown in Fig. 3. The circuit is devised as a fully differential topology

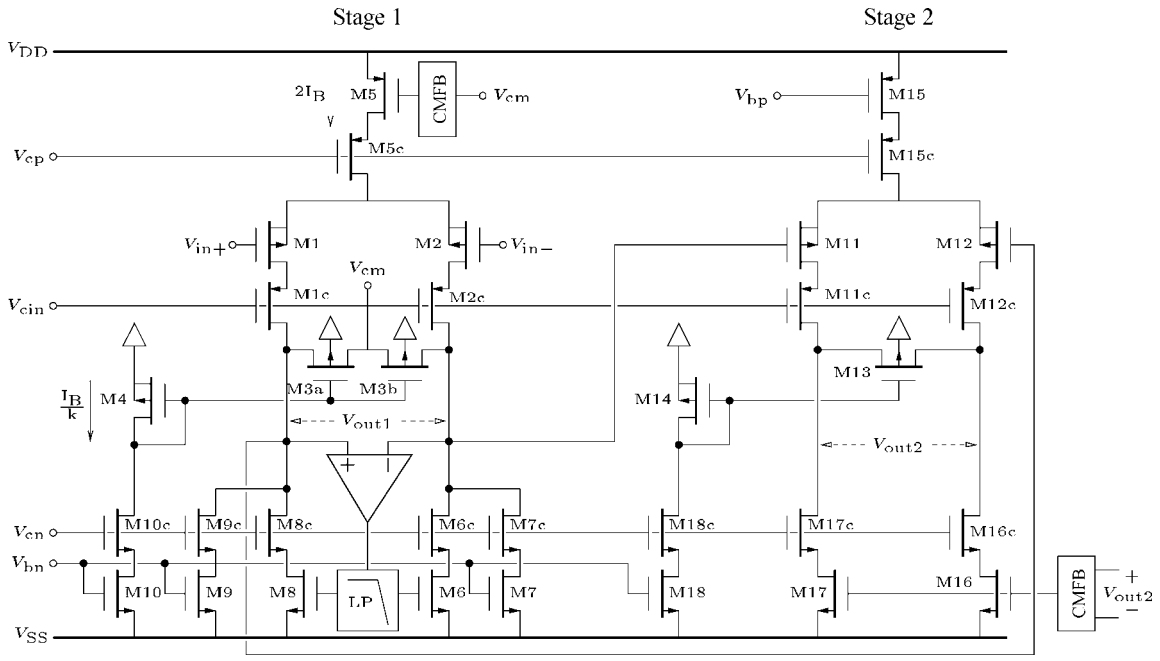


Fig. 3. Amplifier schematic.

in order to maximize CMRR, noise rejection and minimize even order harmonic distortion. Transistors denoted by 'c' are used as cascodes. The amplifier consists of two amplification stages with similar topology, each has a gain of 100.

In the first stage, optimum noise suppression is desirable due to the minute input signal. As the output swing is limited (±1 mV), we can bias the first stage input pair and load in weak inversion hereby maximizing the noise suppression.

Since the second stage needs to support a much higher output swing (±100 mV) and since noise suppression for this stage is not critical, biasing all signal path transistors in strong inversion at a much lower current level is the natural choice.

4.1. Principle of Operation

A simplified schematic of a gain stage showing only the principal transistors is shown in Fig. 4. The differential pair M1 and M2 is loaded by M3, which operates in the linear region. M3 is biased by the diode-connected transistor M4. The DC gain is thus given by:

$$A_V = -\frac{g_{m1}}{2g_{ds3}} \tag{5}$$

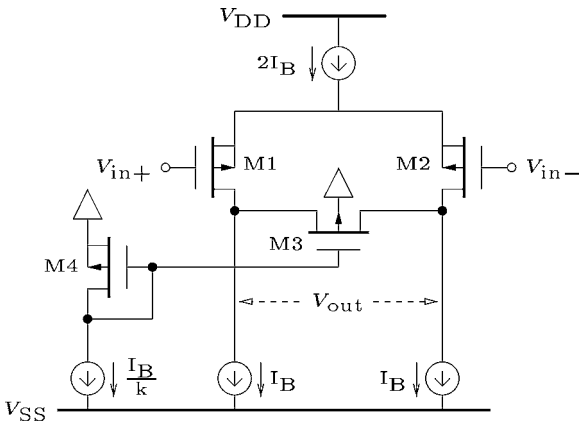


Fig. 4. Simplified schematic.

When proper biasing is applied to this configuration, i.e. all transistors have the same channel inversion level, the gain is independent of the operating mode of the transistors, e.g. strong/weak inversion. Using small signal expressions for the gain parameters obtained from the EKV model [7], we can derive the following expression for the gain with the transistors biased in weak inversion:

$$g_{m1,weak} = 2nK'_P\left(\frac{W}{L}\right)_1 \frac{V_T}{n} e^{\frac{V_{G1}-V_{th0}-nV_{S1}}{V_T}} \tag{6}$$

$$g_{\mathrm{ds3,weak}} = 2nK'_{\mathrm{P}}\left(\frac{W}{L}\right)_3 V_{\mathrm{T}} e^{\frac{V_{\mathrm{G3}}-V_{\mathrm{th0}}-nV_{\mathrm{S3}}}{V_{\mathrm{T}}}} \quad (7)$$

$$\Rightarrow A_{\mathrm{V,weak}} = \frac{1}{2n}\frac{(W/L)_1}{(W/L)_3}\frac{\exp\left(\frac{V_{\mathrm{G1}}-V_{\mathrm{th0}}-nV_{\mathrm{S1}}}{V_{\mathrm{T}}}\right)}{\exp\left(\frac{V_{\mathrm{G3}}-V_{\mathrm{th0}}-nV_{\mathrm{S3}}}{V_{\mathrm{T}}}\right)} \quad (8)$$

where n is the slope factor, V_{th0} is the threshold voltage and V_{T} is the thermal voltage which at the normal body temperature assumes a value: $V_{\mathrm{T}} \simeq 26.7$ mV at 37°C.

Using again the EKV model now for the strong inversion region, the gain expression becomes [7]:

$$g_{\mathrm{m1,strong}} = \frac{K'_{\mathrm{P}}}{n}\left(\frac{W}{L}\right)_1 (V_{\mathrm{G1}} - V_{\mathrm{th0}} - nV_{\mathrm{S1}}) \quad (9)$$

$$g_{\mathrm{ds3,strong}} = K'_{\mathrm{P}}\left(\frac{W}{L}\right)_3 (V_{\mathrm{G3}} - V_{\mathrm{th0}} - nV_{\mathrm{S3}}) \quad (10)$$

$$\Rightarrow A_{\mathrm{V,strong}} = \frac{1}{2n}\frac{(W/L)_1}{(W/L)_3}\frac{V_{\mathrm{G1}} - V_{\mathrm{th0}} - nV_{\mathrm{S1}}}{V_{\mathrm{G3}} - V_{\mathrm{th0}} - nV_{\mathrm{S3}}} \quad (11)$$

From equations (8) and (11) we see that if the same gate-source voltage is applied to load and input pair, both gain expressions simplify to:

$$A_{\mathrm{V}} = \frac{1}{2n}\frac{(W/L)_1}{(W/L)_3} \quad (12)$$

According to equation (12) the gain of the stage can be set by the device aspect ratios. To ensure equal gate-source voltages of M1, M2 and M3 the diode-connected M4 is biased by a scaled current I_{B}/k, and V_{G4} is used to bias M3. M4 is accordingly scaled $(W/L)_4 = 1/k(W/L)_1$, M4 has both source and bulk connected to ground, and thus has the same channel inversion level as the other devices. The source and drain voltages of M3 are set equal to the ground potential by a common mode feedback circuit (not shown).

A prerequisite of equation (12) being valid in the weak inversion region, is that both the input signal being applied and the resulting output voltage stay well within the bound given by: $v_{\mathrm{out}} \ll |V_{\mathrm{GS3}}|$. Oversteering the input will bring the load transistor towards moderate/strong inversion and thus heavily modulate the small signal resistance of M3 r_{ds3}, and induce harmonic distortion.

Gain variation due to device dimension mismatch can be minimized by layout techniques. The gain is however also dependent on the slope factor n, which is given by [7]:

$$\frac{1}{n} = 1 - \frac{\gamma}{2\cdot\sqrt{V_{\mathrm{G}} - V_{\mathrm{th0}} + \left(\frac{\gamma}{2} + \sqrt{\psi_0}\right)^2}} \simeq \frac{1}{1.25} \quad (13)$$

where γ is the body effect factor and ψ_0 is the semiconductor surface potential. It can be shown that n is ultimately dependent on the substrate doping concentration N_{sub}. For variations in N_{sub} of ± 1 order of magnitude, the variation in n was found to be less than 10%, which is within a tolerable range.

4.2. First Stage

Due to the small amplitude of the input signal, it is of prime importance that the noise of the first stage is kept at a minimum. Hence the main power consumption of the entire amplifier is dominated by the first stage. The gain of the stage is given by equation (12) and is chosen to 100, bringing the input signal amplitude from ± 10 μV to ± 1 mV.

4.2.1. Noise Analysis The input referred noise in a strong inversion MOSFET is given by [5]:

$$v_{ni}^2(f) = 4kT\left(\frac{2}{3}\right)\frac{1}{g_m} + \frac{K_{\mathrm{f}}}{WLC_{ox}f} \quad (14)$$

The first term of equation (14) is the inherent thermal noise in the MOSFET and the second term is the flicker (1/f) noise, where K_{f} is a process-dependent constant. The mechanism of flicker noise is still under scrutiny, and research suggests that different mechanisms may be at play for N- and P-type MOSFETs [6]. However, in general, 1/f noise is found to be lower in PMOS transistors than in NMOS transistors. In our case, K_{f} is about an order of magnitude lower in PMOS transistors than in their NMOS counterparts of equal gate area. Thus PMOS transistors will be used in the input stage. In general, the noise in weak inversion transistors is slightly less than in strong inversion [7], however equation (14) can be used as a worst case estimate for calculations. Equation (14) shows that the input referred thermal noise is minimized by maximizing the transconductance g_{m}. Using the EKV-model, the g_{m} of the MOSFET in weak- and strong inversion can be

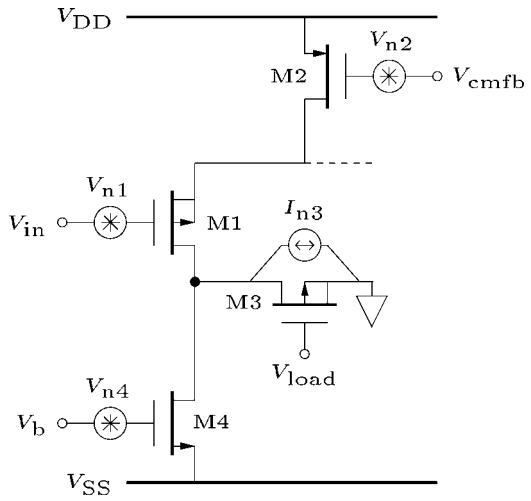


Fig. 5. Half circuit for noise analysis.

shown to be [7]:

$$\text{Weak inversion: } g_{m,weak} = \frac{I_D}{nV_T} \tag{15}$$

$$\text{Strong inversion: } g_{m,strong} = \frac{2nI_D}{V_{eff}} \tag{16}$$

The ratio of the transconductances is: $g_{m,weak}/g_{m,strong} \simeq 5.6$ for a typical effective voltage of 300 mV. Thus, by biasing the input transistors in weak inversion, we can obtain maximum thermal noise suppression for a given drain current I_D.

Due to the minute amplitude of our input signal, the weak inversion region is well suited for our application. So in order to obtain a low input referred noise, most of the amplifier current is drawn by the stage 1 input pair M1 and M2.

A simplified half circuit of a gain stage for calculating the noise contributions of the different transistors, is shown in Fig. 5. As all the noise sources are uncorrelated, their powers add linearly. The noise source corresponding to the top device M2 and the output referred noise of the CMFB circuit, will couple equally to both output nodes and can thus be ignored as the stage is fully differential. The total noise power seen at the output is:

$$v_{no,tot}^2(f) = v_{n1}^2(f)\left(\frac{g_{m1}}{g_{ds3}}\right)^2 + v_{n4}^2(f)\left(\frac{g_{m4}}{g_{ds3}}\right)^2 + \left(\frac{i_{n3}(f)}{g_{ds3}}\right)^2 \tag{17}$$

Referring all noise to the input, i.e. divide by the power gain $(g_{m1}/g_{ds3})^2$, we have the total input referred noise:

$$v_{ni,tot}^2(f) = v_{n1}^2(f) + v_{n4}^2(f)\left(\frac{g_{m4}}{g_{m1}}\right)^2 + \left(\frac{i_{n3}(f)}{g_{m1}}\right)^2 \tag{18}$$

Requiring that the noise contributions from the other devices in the gain stage be negligible compared to the noise inherent to the input device M1, we have:

$$v_{n1}^2(f) \gg v_{n4}^2(f)\left(\frac{g_{m4}}{g_{m1}}\right)^2 + \left(\frac{i_{n3}(f)}{g_{m1}}\right)^2 \tag{19}$$

The thermal noise of a MOSFET biased in the triode region is given by:

$$i_n^2(f) = 4kTg_{ds} \tag{20}$$

Inserting the expressions for the thermal noise from equation (14) and equation (20) in equation (19) then results in:

$$|g_{m1}| \gg |g_{m4}| + \frac{3}{2}|g_{ds3}| \tag{21}$$

The second term of equation (21) is easily fulfilled as the ratio between g_{m1} and g_{ds3} is equal to the gain and hence chosen to 100. The first term suggests that in order to minimize the thermal noise contribution of bias device M4, the ratio of the transconductances should be maximized, i.e. M4 should be biased in strong inversion with a large effective voltage $V_{eff} = V_G - V_S - V_{th}$, and M1 should be biased in weak inversion in order to maximize it's g_m. This result is coherent with our prior decision to bias the input pair in weak inversion to minimize the input pair noise contribution.

As the triode load device has no DC bias current, the flicker noise component for this transistor can be ignored. However for device M4, inserting the flicker noise expression from equation (14) in equation (19) gives us:

$$|g_{m1}| \gg |g_{m4}|\sqrt{\frac{(W/L)_1}{(W/L)_4}\frac{K_{f,N}}{K_{f,P}}} \tag{22}$$

We see that the flicker noise suppression is dependent on the ratio of K_f factors, which in our case is about 10, and thus imposes that the aspect ratio of M1 be set 10 times larger than the aspect ratio of M4 in order to obtain the same relative flicker noise suppression as thermal noise suppression.

By requiring that the thermal noise level of the input device M1, is below the thermal noise of the cuff resistance, we get the following constraint:

$$g_{\mathrm{m1}} > \frac{2n}{R_{\mathrm{cuff}}} \tag{23}$$

Inserting the g_{m} expression for weak inversion in equation (23) gives us a design equation for the required current:

$$I_{\mathrm{D1}} > \frac{2n^2 V_{\mathrm{T}}}{R_{\mathrm{cuff}}} \tag{24}$$

and is in our case set to 37.5 μA for each device in the input pair.

All cascodes are omitted from these considerations, as their noise contributions are negligible. This can be seen from the fact that each cascode transistor can be considered degenerated by an r_{ds} resistance at the source terminal, thus reducing the effective transconductance of the cascode transistor to:

$$g_{\mathrm{m,casc}} = g_{\mathrm{m}}\left(1 - \frac{g_{\mathrm{m}} r_{\mathrm{ds}}}{1 + g_{\mathrm{m}} r_{\mathrm{ds}}}\right) \tag{25}$$

For an example value of $g_{\mathrm{m}} r_{\mathrm{ds}} = 10$, the effective transconductance $g_{\mathrm{m,casc}}$, reduces to $\simeq$9.1% of the nominal value.

4.2.2. Offset Compensation Any threshold offset present in the stage 1 input pair M1, M2, will couple to the output by the amplification factor. The inherent offset ΔV_{t0} in the process used may indeed be orders of magnitude larger than the input nerve signal amplitude, and will thus force the input stage out of weak inversion. Hence some scheme for offset cancellation is needed.

A popular approach for offset cancellation is the *auto-zero* technique [9]. However, as the auto-zero technique is a sampled method, undersampling of wide-band noise is inevitable causing downfolding of thermal noise into the signal band. This will drastically increase the noise floor causing a reduction of the amplifier SNR.

A frequency domain approach such as chopping [8], will remove the output offset to the chop frequency and it's odd order harmonics. This scheme is viable if the gain of a stage is sufficiently low to ensure that the stage does not saturate. In our case, the gain is relatively high and correct biasing levels can not be upheld in the face of the intrinsic offset, making this method unsuitable for our needs.

Instead, a continuous-time current steering scheme for offset cancellation is used. In Fig. 3, the output of stage 1 is amplified, low-pass filtered and used to control current source transistors M6 and M8 in order to match the current offset in M1 and M2. As we are only interested in frequencies $\geq$400 Hz, this scheme is applicable provided the time constant around the control loop is large enough.

The compensation scheme should be well capable of nulling the offset in M1 and M2 and hence we can derive the partition ratio between current source transistors M6, M7 and M8, M9. Let Δv_{th} denote the M1, M2 threshold offset, the resulting current error in the pair is then found to be:

$$\Delta i_{\mathrm{err}} = g_{\mathrm{m1}} \Delta v_{\mathrm{th}} \tag{26}$$

$$= \frac{I_{\mathrm{D}}}{n V_{\mathrm{T}}} \Delta v_{\mathrm{th}} \tag{27}$$

Let Δv_{os} denote the small signal output swing of the offset correction block controlling devices M6 and M8 and α denote the partition ratio between M6 and M7. The resulting current for offset correction is then:

$$\Delta i_{\mathrm{corr}} = g_{\mathrm{m6}} \Delta v_{\mathrm{os}} \tag{28}$$

$$= \frac{2 n I_{\mathrm{D}} \alpha}{V_{\mathrm{eff6}}} \Delta v_{\mathrm{os}} \tag{29}$$

Requiring that $\Delta i_{\mathrm{corr}} > \Delta i_{\mathrm{err}}$ then gives us an expression for the current source partitioning ratio:

$$\alpha > \frac{V_{\mathrm{eff6}}}{2n^2 V_{\mathrm{T}}} \frac{\Delta v_{\mathrm{th}}}{\Delta v_{\mathrm{os}}} \tag{30}$$

For setting the ratio, a worst case estimate of Δv_{th} should be used, e.g. $3\sigma(V_{\mathrm{th0}})$. In practice Δv_{os}, will be some fraction of V_{eff6} as the offset compensation circuit devices M9 and M10 will be replicas of the stage 1 steered current source devices to ensure good matching. In our case, the ratio A, was set to 13%.

A schematic of the offset compensation circuit is shown in Fig. 6. The input differential pair M1 and M2 are connected to the stage 1 differential output. The current output from M1 and M2 is mirrored by a reduction factor 10:1 and converted to a single-ended output voltage. The single-ended output is applied to the low-pass capacitor C_{LP}, which sets the dominant pole in the loop. The M7 and M8 differential pair compares

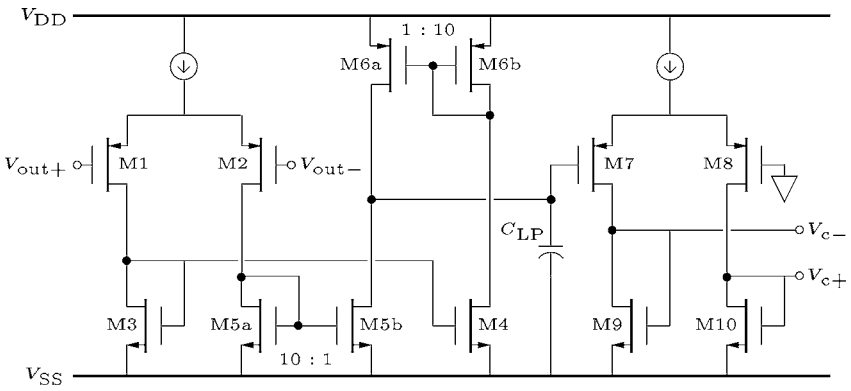


Fig. 6. Offset compensation circuit.

the single-ended voltage to a reference voltage (gnd) and converts the single-ended signal to the differential steering voltages v_{c+} and v_{c-} on the diode-connected devices M9 and M10.

The 1:10 current mirror ratio is intended for scaling of the input pair g_m, so that the gain bandwidth product is given by $GBW = g_m/10 \cdot C_{LP}$. All current mirrors are implemented as cascoded current mirrors to improve accuracy and all current sources are cascoded in order to ensure a high CMRR which is utilized in the single-ended to differential voltage conversion.

Another issue of using a continuous-time offset compensation circuit, is that no charge injection or clock feedthrough will disturb the amplifier input nodes. Indeed, such dynamic effects may induce voltage spikes with amplitudes far in excess of the nerve signal amplitude.

The low-pass filtering of the offset compensation circuit has the effect of introducing a zero at the origin of the transfer function of stage 1, thereby highpass filtering the input signal. The lower 3 dB frequency is then found to be equal to the *GBW* of the offset compensation loop. This high-pass filtering is desirable as it will reduce the amount of flicker noise seen at the output.

Using this offset cancellation scheme will cause some bias current mismatch in the amplifier input stage transistors M1 and M2. The current offset can be modeled by letting a small perturbation ΔV_{GS}, offset the common operating point transconductance g_m, where g_m is the weak inversion transconductance given in equation (6):

$$g_{mi} = g_m \exp\left(\frac{\pm \Delta V_{GS}}{n V_T}\right), \quad i = 1,\ 2 \qquad (31)$$

As $|\Delta V_{GS}| \ll n \cdot V_T$, we can use a 1st order approximation for equation (31):

$$g_{mi} \simeq g_m\left[1 \pm \frac{\Delta V_{GS}}{n V_T}\right], \quad i = 1,\ 2 \qquad (32)$$

The total transconductance of the differential pair can be expressed:

$$g_{m,tot} = 2\frac{g_{m1} g_{m2}}{g_{m1} + g_{m2}} = g_m\left[1 - \left(\frac{\Delta V_{GS}}{n V_T}\right)^2\right] \simeq g_m$$

Hence to a first order approximation, the gain of the stage will not be affected.

4.2.3. Common Mode Feedback As seen from Fig. 3, the common-mode signal (CMS) of stage 1 is readily available in the 'middle' of the load, consisting of transistors M3a and M3b.

Since the signal swing across the load is on the order of ±1 mV the signal will not affect the CMS significantly. Hence, the CMS can be used readily for comparison with a reference and used to control stage 1 current source device M5.

The employed circuit is shown in Fig. 7 and consists of a simple differential pair M1–M2, comparing the CMS to gnd and producing the common-mode feedback voltage on the diode-connected M4.

4.3. Second Stage

The second amplification stage basically has the same topology as the first stage. However, due to the magnitude of the output voltage swing and the resulting gain

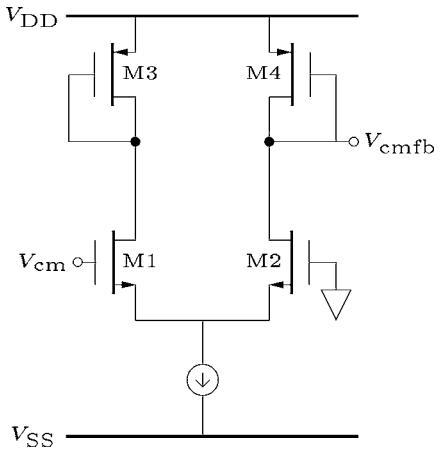


Fig. 7. Common-mode feedback circuit for stage 1.

modulation, weak inversion operation is not applicable here. Hence, all transistors are now operated in the strong inversion region. The gain is again determined by equation (12) and set by the device dimensions to 100, thus bringing the total gain of the amplifier to 80 dB.

The thermal noise requirements for the second stage are much relaxed as the input signal has already gained 40 dB in magnitude. A PMOS input pair is again chosen due to the superior flicker noise performance. The bias current requirements are hence dictated by other specifications and in our case the bandwidth of the second stage. In order to achieve a bandwidth of 10 kHz for the capacitive load driven, the total current for the differential pair was set to 5 μA.

As the transistors are biased in strong inversion, the inherent offset of the stage 2 input devices can be tolerated as this will not bring the transistors out of saturation. However, AC-coupling the two stages was done in the experimental setup to minimize the contribution to offset from the first stage. As seen in the section on experimental results, residual offset at the output of stage 1 may indeed necessitate AC-coupling between the two stages.

A candidate on-chip scheme is shown in Fig. 8. Here, two back-to-back coupled diodes are used as a resistor for the AC-coupling to the input of the second stage. The $I - V$ characteristic of the pair is qualitatively shown to the right. For very low input voltage amplitudes, the equivalent conductance of the pair g_d, will be very small. The diode characteristic is given by:

$$I_D = I_S\left[\exp\left(\frac{V_D}{V_T}\right) - 1\right] \tag{33}$$

Where I_S is the diode-area dependent *scale current*. The diode conductance is then:

$$g_d = \frac{\partial I_D}{\partial V_D} = \frac{I_S}{V_T}\exp\left(\frac{V_D}{V_T}\right) \tag{34}$$

For a 1 μm $\times$ 1 μm diode at 37°C in the technology used for the amplifier prototype, I_S is found to be on the order of 3 fA. Using this value we see that for a voltage swing of 100 mV, the equivalent resistance is in the GΩ range reducing the necessary size of capacitor C_{AC} to the sub pF range making the scheme feasible for integrated solutions.

Common-mode feedback for the second stage is not as simply achieved as for stage 1. As the load device of stage 2: M13, is operated in strong inversion and since the output swing of stage 2 is on the order of $\pm$100 mV, the channel of M13 will change shape with the output voltage and hence the voltage in the 'middle' of the load will be modulated by the input signal.

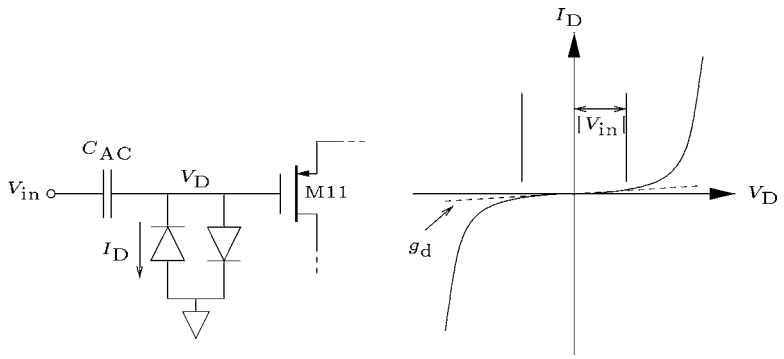


Fig. 8. Diode pair applicable for AC-coupling stage 2.

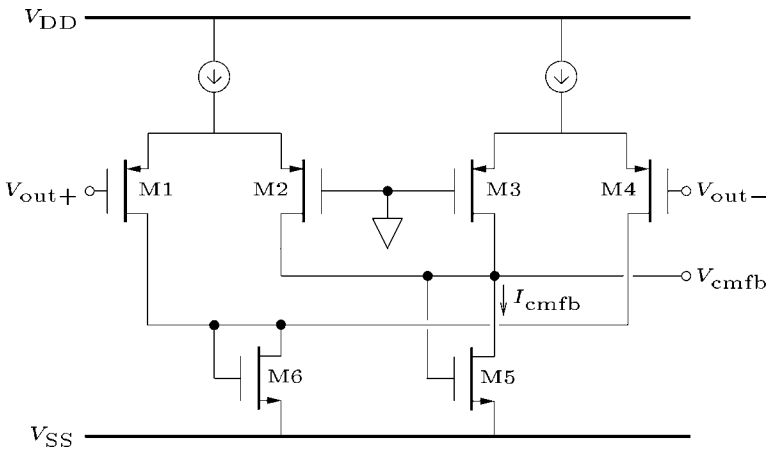


Fig. 9. Common-mode extraction circuit for stage 2.

Instead, the CMS is extracted from the differential output voltages using the difference amplifier shown in Fig. 9. The differential output voltage is compared to a reference (gnd) and an output control voltage v_{cmfb}, is generated. The small-signal output current of the two matched differential pairs M1–M2 and M3–M4 is seen to be:

$$i_{\text{cmfb}} = g_{\text{m1}}(v_{\text{out+}} - v_{\text{out-}}) \tag{35}$$

So an output signal current is only generated when an output voltage imbalance is present.

The resulting output current is fed to the diode-connected M5, generating the control voltage used to control stage 2 current source devices M16 and M17 is Fig. 3.

5. Experimental Results

A test chip with the proposed amplifier has been fabricated in a standard digital 0.5 μm, single poly, N-well CMOS process. A microphotograph of the upper right corner of the chip containing the experimental amplifier is shown in Fig. 10. Unfortunately, the test chip was covered with metal 3 dummy filling, reducing the quality of the microphotograph considerably.

Table 1 summarizes the measured performance of the amplifier.

Figure 11 shows the output power spectrum for an input sine at 1 kHz and amplitude $v_{\text{in}} \simeq 22\ \mu\text{V}_{\text{rms}}$. Due to the fully differential structure of the amplifier, it is seen that the even order harmonics are well below the odd harmonics. The total distortion power is -33 dB below

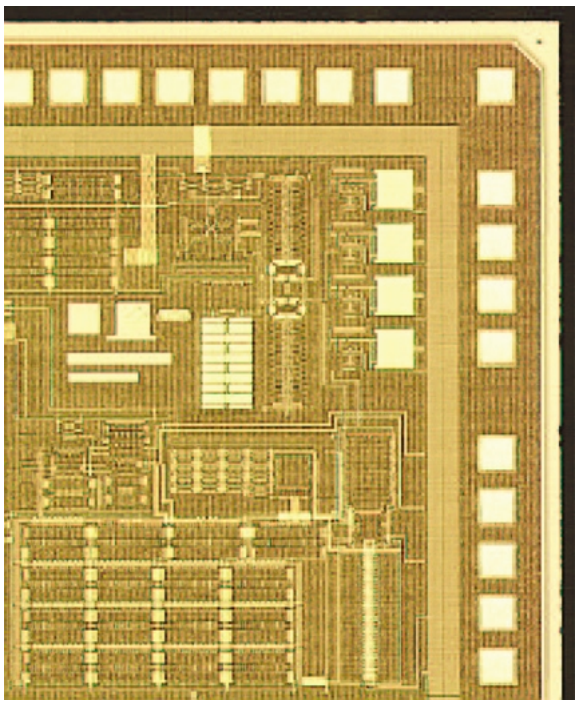


Fig. 10. Chip microphotograph.

the fundamental, equivalent to a THD of 2.2%. As the signal recovery method only yields an SNR of approx. 29 dB, the harmonics will be below the noise floor. Thus the seen performance is acceptable for our application.

The measured output magnitude frequency response is shown in Fig. 12. The high-pass response is due to the current-steering offset compensation scheme. The low-pass response is the band-limiting at the output of

Table 1. Measured amplifier performance.

Supply Voltage (V_{DD}–V_{SS})	3 V
Power Consumption	275 μW
SNR	33.8 dB
THD(@ $f = 1$ kHz)	2.2%
Typ. input offset (stage 1)	<40 μV
CMRR	>87 dB, $f < 100$ kHz
PSRR	>84 dB, $f < 100$ kHz
Gain	80 dB
Equiv. input referred noise	4.8 nV/$\sqrt{\text{Hz}}$

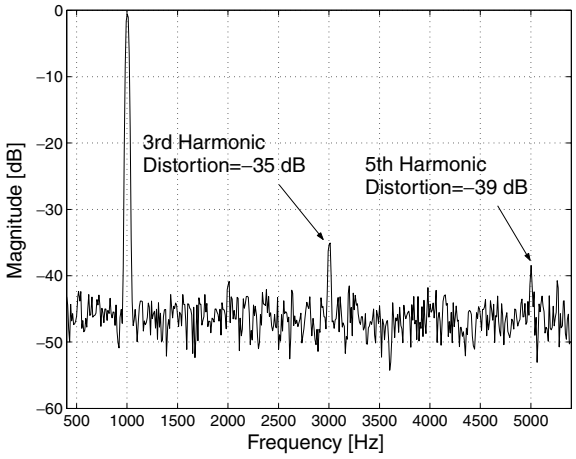


Fig. 11. Measured spectrum at the amplifier output.

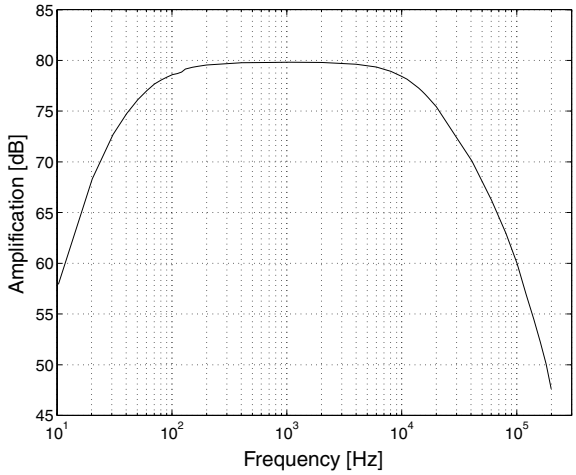


Fig. 12. Amplifier magnitude frequency response.

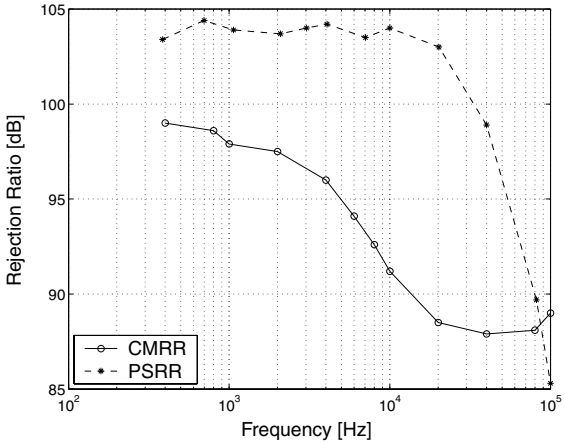


Fig. 13. Measured CMRR and PSRR.

the second amplifier stage. The upper and lower 3 dB frequencies are seen to be approx.:

$$f_{3\text{dB,low}} = 100\,\text{Hz}$$
$$f_{3\text{dB,high}} = 10\,\text{kHz}$$

The CMRR and PSRR frequency responses are shown in Fig. 13. For frequencies $f < 100$ kHz, the amplifier has a CMRR > 87 dB and a PSRR > 84 dB. The CMRR was measured applying a common-mode voltage of 100 mV_{pp}. The circuit draws a total of 91.5 μA from a 3 V supply (excluding output buffers), thus consuming 275 μW. This power consumption is within the bounds set forth for inductively powered devices [10]. The equivalent input referred noise over the signal bandwidth is 4.8 nV/$\sqrt{\text{Hz}}$. Integrating this figure over the ENG bandwidth (400 Hz–4 kHz) gives us the following maximum SNR for a peak-peak ENG of 20 μV:

$$\text{SNR}_{\max} = \frac{20\ \mu\text{V}}{\sqrt{2} \cdot 4.8\ \text{nV} \cdot 60} \simeq 33.8\ \text{dB}$$

Some gain variation was observed in the test chips. These variations are mainly due to offsets in the common mode voltage of stage 1, which in turn modulates the gain as V_{GS} of M3 is altered.

In Fig. 14, the simulated bias currents I_1 and I_2 are shown versus input DC-offset. Alongside the currents, the measured output voltages of stage 1 are shown. Figure 14 shows that offsets <3.5 mV will be compensated by the current steering offset compensation circuit. As we expect a maximum threshold offset of: $3\sigma(V_{\text{th0}}) = 1$ mV for the input pair, a good safety margin is ensured by the offset compensation circuit.

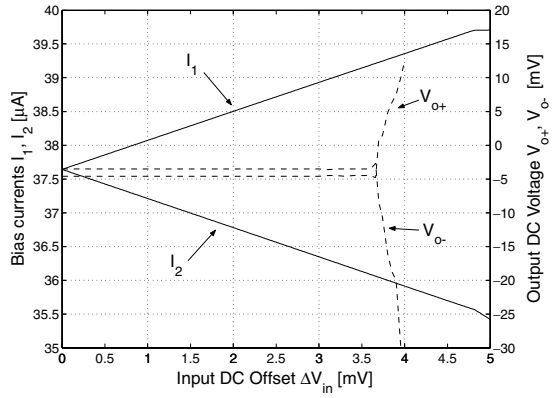


Fig. 14. Offset cancellation by current steering.

Some residual offset will be present due to the device matching in the control loop and is $\simeq$1.1 mV at the output in Fig. 14.

6. Conclusions

In this paper an amplifier for minute nerve signals was presented. The utilization of weak inversion input transistors, with maximized g_m/I_D-ratio to suppress inherent device noise, has been shown to be feasible for the tight power consumption limits inherent to implantable devices. Furthermore, the restrictions on signal amplitude imposed by devices operating in the weak inversion region, combined with it's low noise characteristic, makes this region attractive for signal processing of minute signals.

For offset cancellation, a continuous-time scheme of steering bias currents was used. This was done in order to avoid the subsampling and resulting wideband noise aliasing introduced by switching schemes. A condition of this method is that only AC-information needs to be amplified as DC is filtered out. The continuous-time offset compensation will impose a zero at the origin and hence the flicker noise will be high-pass filtered.

Finally, the characteristics of a test chip with a prototype amplifier was presented. The prototype exhibits performance within the boundaries put forth. The simple scheme of using a single fixed bias transistor as load, does not provide high linearity as seen from Fig. 11. However, as the signal recovery method only provides an SNR of approx. 29 dB, the THD is within acceptable bounds. The achieved noise suppression, allows for an SNR with some margin to the SNR provided by the signal recovery.

Acknowledgment

This work was supported by the Danish Medical Research Council.

References

1. M. K. Haugland, *Natural Sensory Feedback for Closed-loop Control of Paralyzed Muscles*. Ph.D. thesis, University of Aalborg, 1994.
2. A. Lickel, *Restoration of Lateral Hand Grasp in a Tetraplegic Patient Applying Natural Sensory Feedback*. Ph.D. thesis, University of Aalborg, 1998.
3. G. Gudnason, E. Bruun, and M. Haugland, "An implantable mixed analog/digital neural stimulator circuit," in *Proc. IEEE ISCAS*, Orlando, FL, May 30–June 2, 1999, pp. 375–378.
4. K. Papathanasiou and T. Lehmann, "An implantable CMOS nerve signal conditioning system for recording nerve signals with cuff electrodes," in *Proc. IEEE ISCAS*, Geneva, May 28–31, 2000, pp. 281–284.
5. K. G. Lamb, S. J. Sanchez, and W. T. Holman, "A low noise operational amplifier design using subthreshold operation," in *Proc. IEEE Midwest Symp. Circuits Syst.*, Notre Dame, IN, August 9–12, 1998, pp. 35–38.
6. J. Chang, A. A. Abidi, and C. R. Viswanathan, "Flicker noise in CMOS transistors from subthrehold to strong inversion at various temperatures." *IEEE Trans. On Electron Devices* 41, pp. 1965–1971, 1994.
7. C. C. Enz, "The EKV model: A most model dedicated to low-current and low-voltage analogue circuit design and simulation," in *Low-Power HF Microelectronics A Unified Approach*, Machado, G. A. S. (ed.), IEE, 1996, pp. 247–299.
8. A. Bakker, K. Thiele, and J. H. Huijsing, "A CMOS nested-chopper instumentation amplifier with 100-nV offset." *IEEE Journal of Solid-State Circuits* 35(12), pp. 1877–1883, 2000.
9. C. C. Enz and G. C. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: Autozeroing, correlated double sampling, and chopper stabilization," *Proc. of the IEEE* 84(11), pp. 1584–1614, 1996.
10. G. Gudnason, J. H. Nielsen, E. Bruun, and M. Haugland, "A distributed transducer system for functional electrical stimulation," in *Proc. ICECS*, 2001, pp. 397–400.

Jannik Hammel Nielsen was born in Nuuk, Greenland, in 1972. He received the M.Sc. degree in

electrical engineering from the Technical University of Denmark, Lyngby in 1999. He is currently working toward the Ph.D. degree at the Technical University of Denmark.

His main research interests are in low-voltage low-power analog systems, medical electronics and data converters.

Torsten Lehmann was born in Bagsværd, Denmark, in 1967. He received the M.Sc. degree in electrical engineering from the Technical University of Denmark, Lyngby in 1991, and the Ph.D. degree in 1995 for the work "hardware learning in analog VLSI neural networks."

He spent two years as a Research Fellow funded by the European Union at the University of Edinburgh, Scotland, where he worked with biologically inspired neural systems using pulse stream techniques. From 1997 to 2000, he was an Assistant Professor in microelectronics at the Technical University of Denmark, working with low-power low-noise low-voltage analog and mixed analog-digital circuits. He worked briefly as an ASIC design engineer with Microtronics A/S, Denmark. He is currently with Cochlear Ltd., Austrialia, developing cochlea implants. His main research interests are in solid-state circuits and systems (analog and digital), medical electronics and artificial neural networks.

A CMOS CHOPPER AMPLIFIER FOR IMPLANTABLE SENSORS

Jannik Hammel Nielsen and Erik Bruun

Ørsted-DTU, Technical University of Denmark
DK-2800 Kgs. Lyngby, Denmark
jhn@oersted.dtu.dk

Abstract

This paper describes a low-power, low-noise chopper stabilized CMOS instrumentation amplifier for biomedical applications. Low thermal noise is achieved by employing MOSTs biased in the weak inversion region, whereas chopper stabilization is utilized to shift 1/f-noise out of the signal band hereby ensuring overall low noise performance. The resulting equivalent input referred noise is approx. $5nV/\sqrt{Hz}$ *for a chopping frequency of* $20\,kHz$. *The amplifier operates from a modest supply voltage of* $1.8\,V$, *drawing* $136\,\mu A$ *of current thus consuming* $245\,\mu W$ *of power. The gain is* $74\,dB$ *over a* $4\,kHz$ *bandwidth and a CMRR of* $96\,dB$.

1. Introduction

In recent years, research has focused on the implementation of systems-on-chip (SOC) for sensing, recording and processing of various physical signals, e.g. biological signals, silicon microsensor readouts, etc. A common denominator for these systems is the necessity for precision instrumentation amplifiers prior to any further processing as many of the sensors employed, only provide very weak signal strengths.

As many of these systems are intended for battery operation or inductively coupled power supplies [2], low power consumption of the circuitry employed is of prime importance. For very weak signal amplification, this constraint poses a severe challenge as the inherent electronic circuit noise is inversely proportional to the power consumption. Hence high amplification using few components in the initial processing is crucial.

In this paper, we propose a chopped micropower instrumentation amplifier suitable for amplification of nerve signals recorded by cuff-electrodes. The typical amplitude of these nerve signals are in the range of $\pm 10\,\mu V$ lying primarily in the frequency band $400\,Hz \leq f_{\text{sig}} \leq 4\,kHz$ [3].

A block diagram of the amplifier is shown in fig. 1. The amplifier consists of two stages where the first stage has a gain of $40\,dB$ and the second stage has a gain of $34\,dB$. The input signal is modulated by chopping with a square signal at frequency f_{chop}, and demodulated at the output of stage 1.

2. Noise considerations

Because of the very weak input signal, noise suppression is of utmost importance in the first stage of the amplifier,

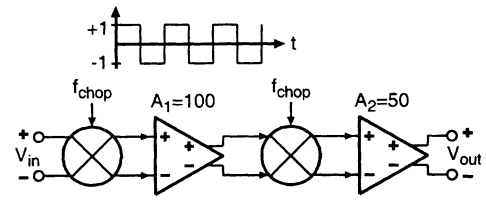


Figure 1: Amplifier block diagram.

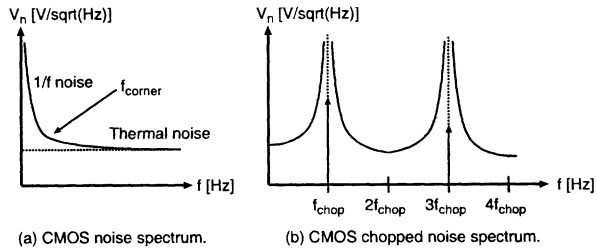


Figure 2: CMOS noise spectra.

whereas for the second stage it is less critical as the signal at that point has gained $40\,dB$ in magnitude. The input referred noise for a MOST is given by:

$$v_{\text{ni}}^2(f) = 4kT\left(\frac{2}{3}\right)\frac{1}{g_{\text{m}}} + \frac{K_{\text{f}}}{WLC_{\text{ox}}f} \tag{1}$$

The first term of eq. (1) is the thermal noise and the second term is the flicker (1/f) noise. The noise spectrum is qualitatively shown in fig. 2a.

According to eq. (1) the thermal noise can be minimized by maximizing g_{m}. From the EKV model we have the transconductances for MOSTs working in strong and weak inversion respectively [1]:

$$\frac{g_{\text{m,weak}}}{I_{\text{DS}}} = \frac{1}{nV_{\text{T}}} \tag{2}$$

$$\frac{g_{\text{m,strong}}}{I_{\text{DS}}} = \frac{2}{V_{\text{eff}}} \tag{3}$$

V_{T} is the thermal voltage and $V_{\text{eff}} = V_{\text{G}} - V_{\text{th}} - nV_{\text{S}}$ is the effective voltage, where V_{th} is the threshold voltage and n is the slope factor. All terminal voltages are referred to the bulk. We see that for a given current I_{DS}, the g_{m} is maximized in weak inversion and hence the thermal noise is minimized.

The 1/f-noise reaches the level of the wideband thermal noise at the corner frequency f_{corner}. Unfortunately,

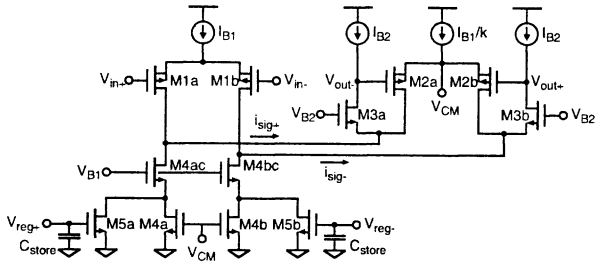


Figure 3: Stage 1 schematic.

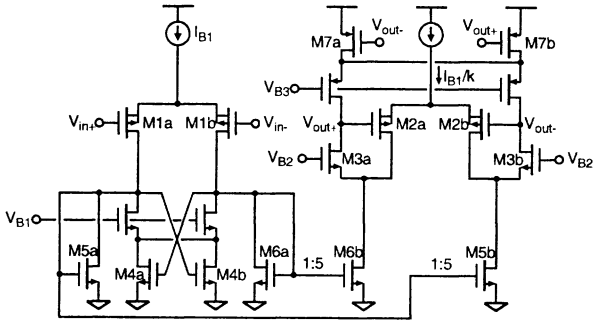


Figure 4: Stage 2 schematic.

f_{corner} can be located at several kHz for MOSTs, dominating the noise spectrum at low frequencies. A method for reducing the effect of 1/f-noise is to employ chopping in the amplifier. As seen from fig. 1, the V_{in} signal is chopped twice, prior and after amplification hence shifting the signal frequency during amplification and then shifting it back to the baseband. The 1/f-noise however, is only chopped once, leaving it at the odd multiples of the chopping frequency as illustrated in fig. 2b hence effectively shifting it out of the baseband.

3. Amplifier

3.1. Stage 1

A schematic of the first amplifier stage is shown in fig. 3. The M1 input pair provides a differential signal current $i_{\text{sig}} = i_{\text{sig}+} - i_{\text{sig}-}$, which is folded to the load branch by the M3 folded cascode pair. The high impedance of the output, forces the signal current to flow through the transimpedance load consisting of the M2 pair. Hence we have the gain of the stage given by:

$$A_1 = \frac{g_{\text{m1}}}{g_{\text{m2}}} \tag{4}$$

The g_{m} for a MOST biased in weak inversion can be written using the EKV model [1]:

$$g_{\text{m}} = 2V_{\text{T}}K'\left(\frac{W}{L}\right)\exp\left(\frac{V_{\text{G}} - V_{\text{th}} - nV_{\text{S}}}{nV_{\text{T}}}\right) \tag{5}$$

Inserting eq. (5) in eq. (4) the gain expression becomes:

$$A_1 = \frac{\left(\frac{W}{L}\right)_1 \exp\left(\frac{V_{\text{G1}} - V_{\text{th}} - nV_{\text{S1}}}{nV_{\text{T}}}\right)}{\left(\frac{W}{L}\right)_2 \exp\left(\frac{V_{\text{G2}} - V_{\text{th}} - nV_{\text{S2}}}{nV_{\text{T}}}\right)} = \frac{\left(\frac{W}{L}\right)_1}{\left(\frac{W}{L}\right)_2} \tag{6}$$

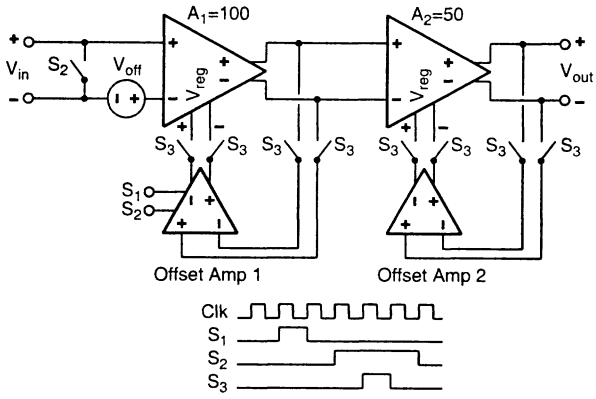


Figure 5: Offset cancellation scheme.

A prerequisite of eq. (6) being valid is that the gate-source voltage of the two pairs is the same. This can be accomplished by ensuring that the current density is the same in both pairs when employing MOSTs of equal length. This implies that alongside the dimension scaling, the bias current of the pairs should be scaled accordingly so that the bias current scale factor k of fig. 3 is given by: $k = A_1$.

The necessary bias current for suppressing the thermal noise sufficiently below signal levels, was set to $I_{\text{B1}} = 100\,\mu A$. For a gain of 100, the scaled bias current in the load pair was hence set to $1\,\mu A$. The relatively high bias current in the input pair combined with the much smaller output pair, gives the stage a $3\,dB$ bandwidth of several MHz making the stage very suitable for chopping as there will be no significant phase shift at the chopping frequency.

Common-mode stabilization of the stage is accomplished by connecting the gates of M4a and M4b with the source potential of the M2 pair and biasing the M4 pair in the triode region. For small output amplitudes, this node will not carry any signal and hence represent a level shifted common-mode signal.

The extra port V_{reg}, which controls the M5 pair is used for offset compensation as explained in section 4.

3.2. Stage 2

A schematic illustrating the principle of operation of the second amplifier stage is shown in fig. 4. Only principal MOSTs are shown in order not to complicate the schematic unnecessarily. This stage has a similar structure as stage 1, however some differences remain.

Weak inversion operation of the principal MOSTs in the second stage is not applicable as the higher signal swing would force them out of the desired operating range. Hence all MOSTs are biased in strong inversion in order to acommodate the higher signal swing. The gain expression for the second stage is equal to that of the first: $A_2 = A_1 = g_{\text{m1}}/g_{\text{m2}}$. In strong inversion we have from the EKV-model [1]:

$$g_{\text{m}} = \frac{K'}{n}\left(\frac{W}{L}\right)(V_{\text{G}} - V_{\text{th}} - V_{\text{S}}) \tag{7}$$

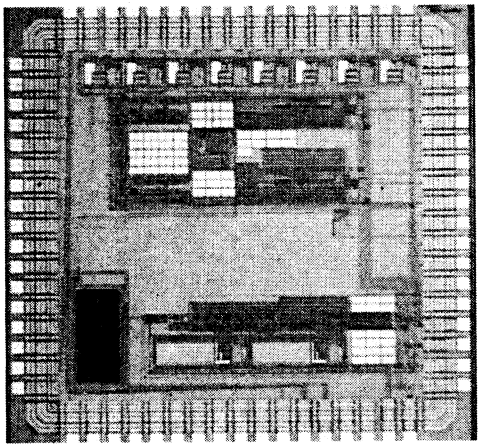


Figure 6: Chip die microphotograph.

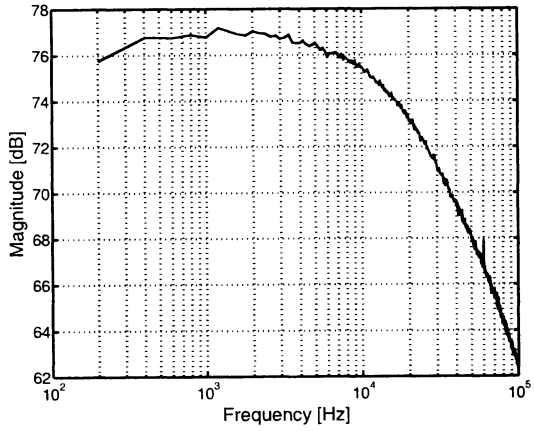


Figure 7: Amplifier magnitude response.

Assuming again equal current density in input- and output pairs, we then have:

$$A_2 = \frac{\left(\frac{W}{L}\right)_1}{\left(\frac{W}{L}\right)_2} \tag{8}$$

In the first stage a high biasing current I_{B1}, was necessary to ensure sufficient thermal noise suppression. For the second stage, the signal has gained $40\,dB$ in magnitude and a high biasing current is no longer needed. In order to achieve sufficient gain without biasing the load pair at extremely low current levels, part of the amplification is done in the current domain. The M5 and M6 current mirror pairs each multiply the signal current by a factor 5. By setting $I_{B1} = 10\,\mu A$ and $k = 10$, we thus achieve a gain of $A_2 = 50$.

Common-mode stabilization is achieved by cross-coupling the M7 pair, connecting their drains and cascoding to ensure sufficient differential output impedance. In order to implement the M5 and M6 current mirrors without mirroring the bulk of the I_{B1} bias current, the M4 pair is designed to take up most of the DC biasing current, whereas the signal current flows in the mirrors.

3.3. Chopper

The chopper circuit is the conventional 4-switch fully differential type, where the switches are implemented as NMOSTs.

The input chopper is noise critical as it modulates the weak sensor readout. Thus the size of the input chopper MOSTs is determined by the acceptable amount of thermal noise from the resistive channel. The output chopper switches can be minimum size due to the stage 1 gain.

4. Offset compensation

As ideal matching of MOSTs cannot be achieved, some offset voltage will always be present in the signal path. As this offset voltage V_{off}, in our case may have a magnitude of several millivolts, saturation of the amplifier is inevitable due to the high gain. To reduce the offset an extra amplifier input port V_{reg}, has been included as shown in fig. 3. By controlling V_{reg}, the bias current of each of the input branches of an amplifier stage can be controlled and hence by monitoring the amplifier stage output, the offset can be cancelled. This offset regulation port is also implemented in the second amplifier stage however not shown for simplicity in fig. 4.

In fig. 5, the scheme for offset cancellation is shown. A discrete-time scheme is employed in order to avoid unnecessarily large time-constants which imply large capacitor sizes. To measure the offset voltage, the input of the instrumentation amplifier is shorted during $S2$. The resulting output voltage of each stage is then fed to an auxilliary OTA (Offset Amp 1 and 2 respectively) during $S3$, which controls the offset nulling port V_{reg}, of the stage. The resulting regulation voltage is then stored on C_{store} as shown in fig. 3. When $S3$ opens, the instrumentation amplifier is offset regulated. The stage 1 auxilliary OTA is furthermore itself offset compensated using the auto-zero technique. The $S1$ signal is used for internal sampling of the Offset Amp 1 offset.

5. Measurements

The circuit has been implemented in a standard $0.35\,\mu m$ CMOS process. A chip die microphotograph is shown in fig. 6.

The measured magnitude response is shown in fig. 7. The gain is seen to be slightly above the specified $74\,dB$ across the nerve signal bandwidth.

The chopped output noise PSD is shown in fig. 8 and is approx. $32\,\mu V/\sqrt{Hz}$ over the nerve signal bandwidth corresponding to $5\,nV/\sqrt{Hz}$ at the input. The chopping frequency was set to $20\,kHz$.

The auto-zero offset compensation technique employed for the stage 1 offset amplifier presented in section 4 appeared to have some problems, leaving us with a residual stage 1 output offset of approx. $60\,mV$ in the test chips. Therefore the auto-zero action for offset amp 1 was disabled, the $S3$ switches of stage 1 were shorted and a large capacitor ($100\,nF$) was added externally to the offset amp 1 output to slow down the regulating loop.

However, the discrete-time regulation of the amplifier

	This work	[4]	[5]	[6]
$V_{DD} - V_{SS}$ $[V]$	1.8	5	1.8	3
Power consumption $[\mu W]$	245	1300	775	275
Input referred noise $[nV/\sqrt{Hz}]$	5	3.3	56	5
Gain $[dB]$	76	40	51	80
CMRR $[dB]$	96	82	100	87
THD $[\%]$	<1	-	-	<2.2
Technology $[\mu m]$	0.35 CMOS	0.8 BiCMOS	0.35 CMOS	0.5 CMOS

Table 1: Preliminary measured amplifier performance and comparison with prior work.

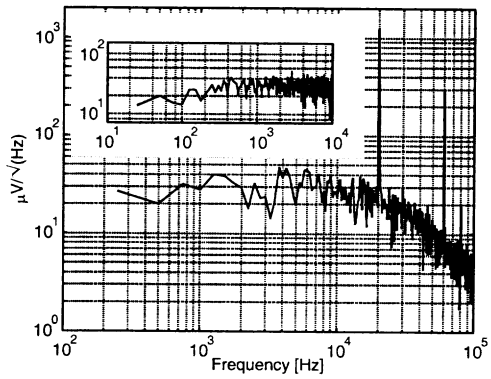


Figure 8: Output noise PSD.

V_{reg} port proved to be operational for the stage 2 offset amp which was *not* auto-zeroed. Fig. 9 shows the time sequence of the stage 2 output offset for a calibration period of 170 s. The offset nulling action is clearly seen proving the validity of the concept.

The current consumption excluding biasing and output buffers is $136\,\mu A$ drawn from a $1.8\,V$ supply, giving us a power consumption of $245\,uW$. However as the chopping signal signal was applied externally, a slightly higher power consumption is expected for a solution with an on-chip oscillator.

The measured performance of the circuit is summarized in table 1 together with results of recent similar work.

6. Conclusion

A high-gain, low-power chopper-stabilized instrumentation amplifier has been presented. The circuit was implemented in a standard $0.35\,\mu m$ CMOS process. In order to minimize circuit thermal noise, weak inversion MOSTs are employed in the input stage of the amplifier. To eliminate 1/f noise, which dominate the noise spectrum at low frequencies, chopper modulation is used to shift the 1/f noise out of the signal band.

Offset reduction is accomplished by including an offset nulling port in each amplifier stage controlled by a local discrete-time feedback loop. However, this scheme was only verified for the second stage as the first stage had some problems.

Comparison with prior art shows that this work has

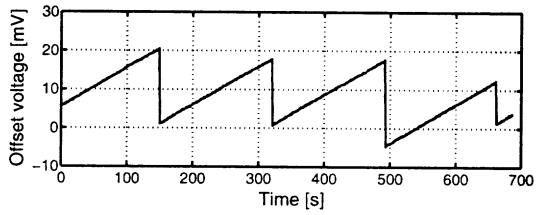


Figure 9: Discrete-time offset nulling of stage 2 offset.

comparable performance at a relatively low power consumption.

7. Acknowledgement

This work was supported by the Danish Medical Research Council.

8. References

[1] C. Enz, et. al., "An Analytical MOS Transistor Model Valid in All Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications", *Analog Integrated Circuits and Signal Processing*, vol. 8, pp. 83-114, 1995.

[2] G. Gudnason, et. al., "A Distributed Transducer System for Functional Electrical Stimulation", *Proc. IEEE Int. Conf. Electronics, Circuits and Systems*, vol. 1 pp. 397-400, 2001.

[3] J. J. Struijk, et. al., "Cuff electrodes for long-term recording of natural sensory information", *IEEE Eng. in Medicine and Biology*, vol. 18(3), pp. 91-98, 1999.

[4] R. Rieger, et. al., "Design of a Low Noise Preamplifier for Nerve Cuff Recording Systems", *IEEE J. Solid-State Circuits*, vol. 8, pp. 1373-1379, 2003.

[5] Y. Hu and M. Sawan, "CMOS Front-end Amplifier Dedicated to Monitor Very Low Amplitude Signal from Implantable Sensors", *Analog Integrated Circuits and Signal Processing*, vol. 33(1), pp. 29-41, 2002.

[6] J. Nielsen and T. Lehmann, "An Implantable CMOS Amplifier for Nerve Signals", *Analog Integrated Circuits and Signal Processing*, vol. 36(1-2), pp. 153-164, 2003.

A 200MSPS, 14b, 97mW DAC in 0.18µm CMOS

Qiuting Huang, Pier Andrea Francese, Chiara Martelli, ETH Zurich
Jannik Nielsen, DTU Lyngby

High-speed, high-resolution D/A converters are an essential element for direct frequency synthesizers, arbitrary wave generators, video displays and communications transmitters. In recent years significant progresses in DAC techniques have enabled resolutions of more than 12b to be achieved for more than 100MSPS output rates [1~4]. Beyond sophisticated layout [2], the main technique to achieve 12b-plus resolution is calibration. Doing so at startup using off-chip calibration engines has the advantage that the DAC circuit itself is simpler, so that higher frequency [3] or lower power [4] can be achieved with moderate chip area. The drawbacks of such an approach are the cost of the off-chip circuitry and vulnerability to time-varying errors such as temperature-dependent offset and (signal-dependent) variation of load impedance. Background calibration based on the concept of floating current source [1] removes both static and slow varying errors. Such a concept is further explored in this contribution under low voltage constraints. A full analog solution is proposed for the background calibration loop in order to improve area, power and performance of prior art.

The overall 14b DAC shown in Fig.20.3.1 is segmented into a 5b unary MSB array, a 4b unary ULSB array and a 5b binary LLSB array of current sources. Each of the 32 floating currents of the MSB array is calibrated in turn in the background. Instead of trimming the voltage developed on a common resistor by each MSB current using an ADC-DSP-DAC loop [1], the architecture in Fig.20.3.1 compares each MSB with a reference current directly.

The current comparator has a regulated-cascode input node (A) to source the MSB current accurately and folds it up to a better point (B) for high output impedance. From node B any difference between the MSB and the reference current flows into the storage capacitor C_{storei} connected to it during the *i*th calibration phase to modify the corresponding trimming voltage V_{Si} until it reaches the correct value to compensate the mismatch. Each calibration phase consists of 25 sampling clock cycles. Each of the 32 MSB sources is calibrated at 4μs intervals.

During the guard time between two consecutive calibration phases when MSB cells are being switched, the parasitic capacitance at node B may discharge due to the temporary current imbalance. To prevent subsequent charge-sharing from causing glitches on each V_{Si}, the latter is sampled by a common tracking capacitor C_{hold} during the phase before the particular MSB is to be calibrated. This pre-sampled voltage V_{Hi} is then used to hold the voltage on node B through a buffer during the guard time to make the transition seamless. Source followers are used to shift V_{Si} to a lower level, so that the gate of each trimming transistor can be biased close to V_{SS}. This helps each stacked transistor in the floating source to have enough V_{DS} to stay in saturation. Since more than five transistors are stacked under a 1.8V supply, careful biasing and transistor sizing are required. To reduce the uncertainty of total current consumption and the trimmable range of each floating current source, the ambiguity of the NMOS-to-PMOS current mirror in [1] needs to be removed. This is achieved by a common PMOS-to-NMOS mirror, which pre-distorts the required bias for each MSB.

With 0.5mA in each MSB source, the total DAC signal current is 16mA. To ensure load-independent accuracy and provide output voltage swing, the DAC currents are first summed

by a differential regulated-cascode structure (Fig.20.3.2) then folded down towards the load impedance. The price paid for this load-independent output stage under low voltage constraint is the doubling of current due to the folding operation of the current buffer. Each differential branch of the current buffer is biased at 18mA to provide some margin over the signal, using a cascode current source from V_{DD}. To maintain signal-independent accuracy the effective input impedance of the summing node must be much lower than those of the signal and bias current sources. The main regulation amplifiers are therefore simple cascoded inverters to sustain high gain at high frequency. Cross-coupled switches are used before the output impedance to provide a return-to-zero (RZ) option to mask switching transients. To match the DAC output to test instruments two 25Ω resistors are placed on chip at both output nodes. When a 50Ω instrument is connected via a transformer the maximum output voltage is 400mV.

Implemented in a 0.18µm 1P6M CMOS process, the DAC chip consumes 97mW from 1.8V. Fig.20.3.3 shows the measured static nonlinearities at 200MHz update rate. The maximum INL of 0.65 LSB and DNL of 0.55 LSB are both determined by the matching accuracy of the ULSB array that causes a step error every 512 codes. The fully analog calibration achieves much higher accuracy for the MSB array. Dynamic measurements are performed at 200 MSPS for different signal frequencies and shown in Fig.20.3.4 for both NRZ and RZ mode. As expected the NRZ produces a better maximum SFDR of 85dB at 249kHz but suffers more from the distortion associated with switching transients when signal frequencies are higher. The RZ mode has the best all-round performance. Constant at ~76dB until 2MHz, its SFDR stays no less than 70dB up to 20MHz and more than 60dB up to 90MHz. Third harmonic is the limiting factor in all SFDR measurements. Fig.20.3.5 shows the measured output

spectrum with the 10kHz signal removed by AC coupling. The average noise PSD is –160dBm/Hz, as intended by design. The spurs due to the calibration sequence can also be seen to be less than –97dBm (–93dBFS). Fig.20.3.6 summarizes the overall performance. The chip micrograph is shown in Fig.20.3.7. The core area is only 1mm^2.

In a broadband transmitter such as that for a VDSL modem, the signal replicas at clock multiples must be removed by a reconstruction filter to restrict out-of-band spurious transmission. An adequate oversampling ratio facilitates the realization of such a filter. For the 12MHz signal bandwidth of VDSL the present DAC offers 8x oversampling, 12~13b linearity in-band and 10b-plus beyond. Together with the low noise and low spurs described above the overall performance is well suited for VDSL and other higher speed applications.

Acknowledgment: Prof. Z. Hong, Dr. Z. Zhu and Dr. C. Hammerschmied for substantial contributions during various stages of this project.

References:

[1] A. Bugeja, B.S. Song, "A Self-Trimming 14b 100MSample/s CMOS DAC", ISSCC Dig. of Tech. Papers, pp.44-45, 2000

[2] A. Van den Bosch et al, "A 12b 500MSample/s Current-Steering CMOS D/A Converter", ISSCC Dig. of Tech. Papers, pp.366-367, 2001

[3] W. Schofield, D. Mercer, L. St. Onge, "A 16b 400MS/s DAC with < -80dBc IMD to 300MHz and < -160dBm Noise Power Spectral Density", ISSCC Dig. of Tech. Papers, pp.126-127, 2003

[4] Y. Cong, R. Geiger, "A 1.5V 14b 100MS/s Self-Calibrated DAC", ISSCC Dig. of Tech. Papers, pp.128-129, 2003

List of Figure Captions:

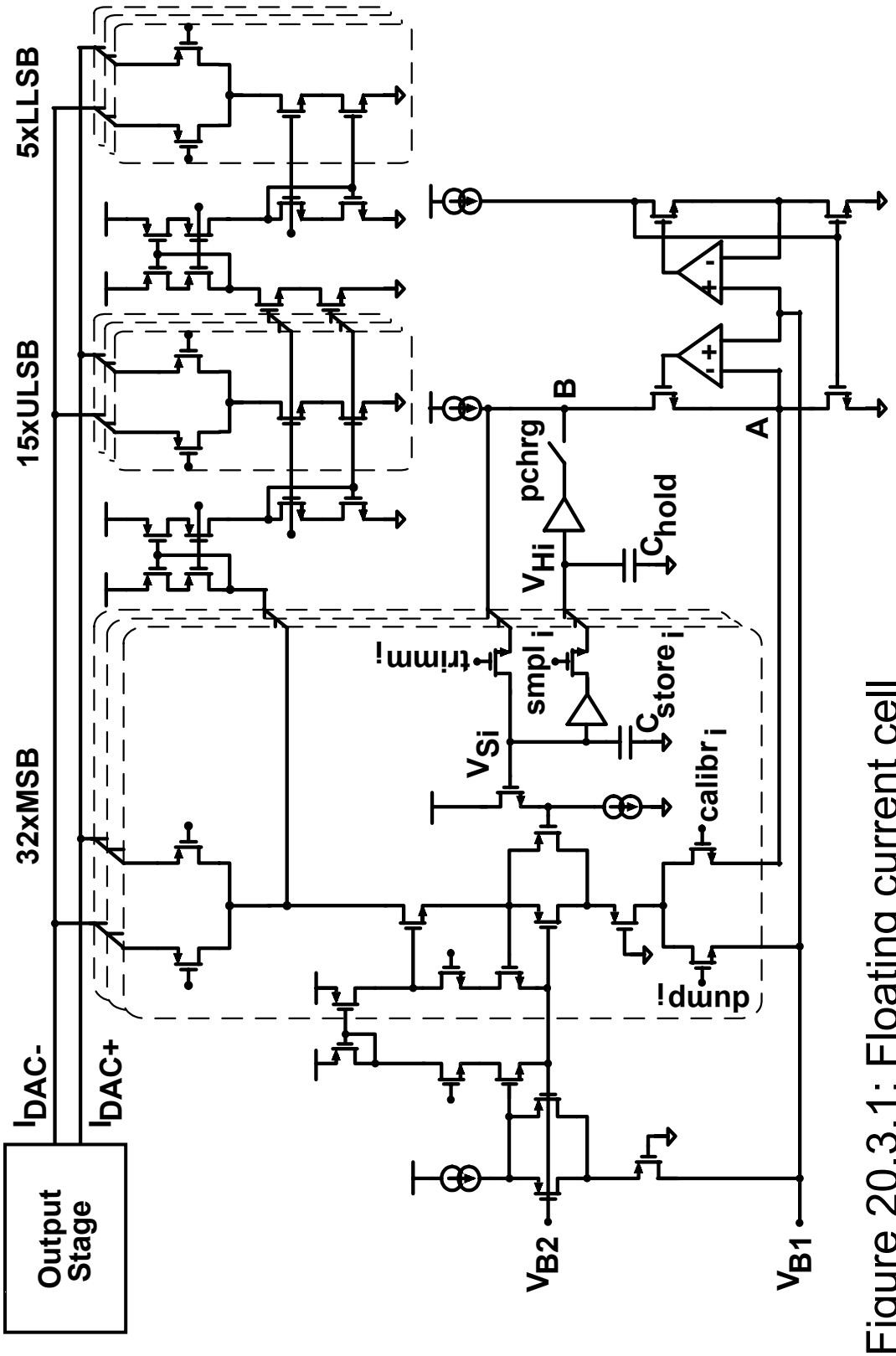


Figure 20.3.1: Floating current cell

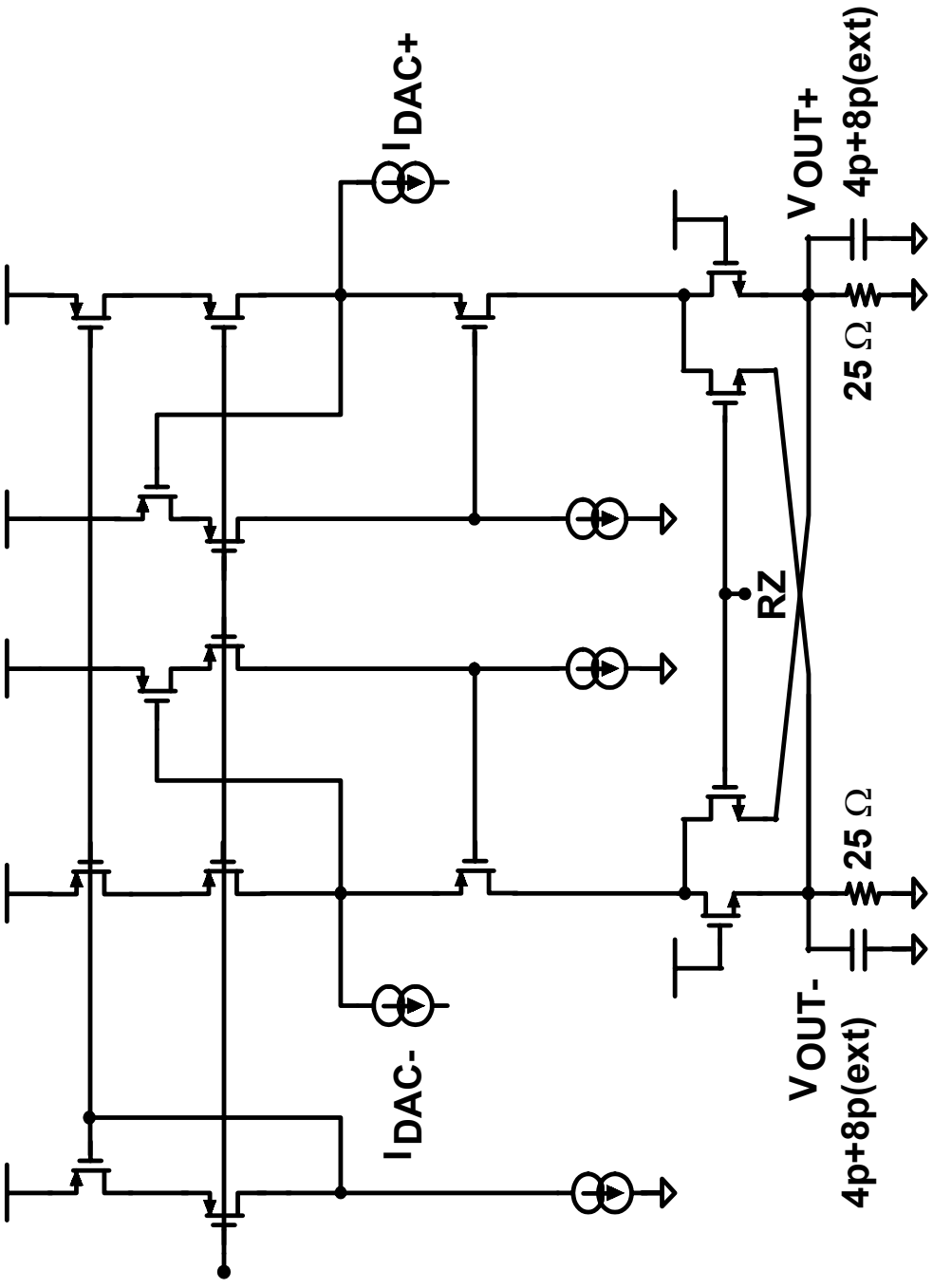


Figure 20.3.2: Output stage

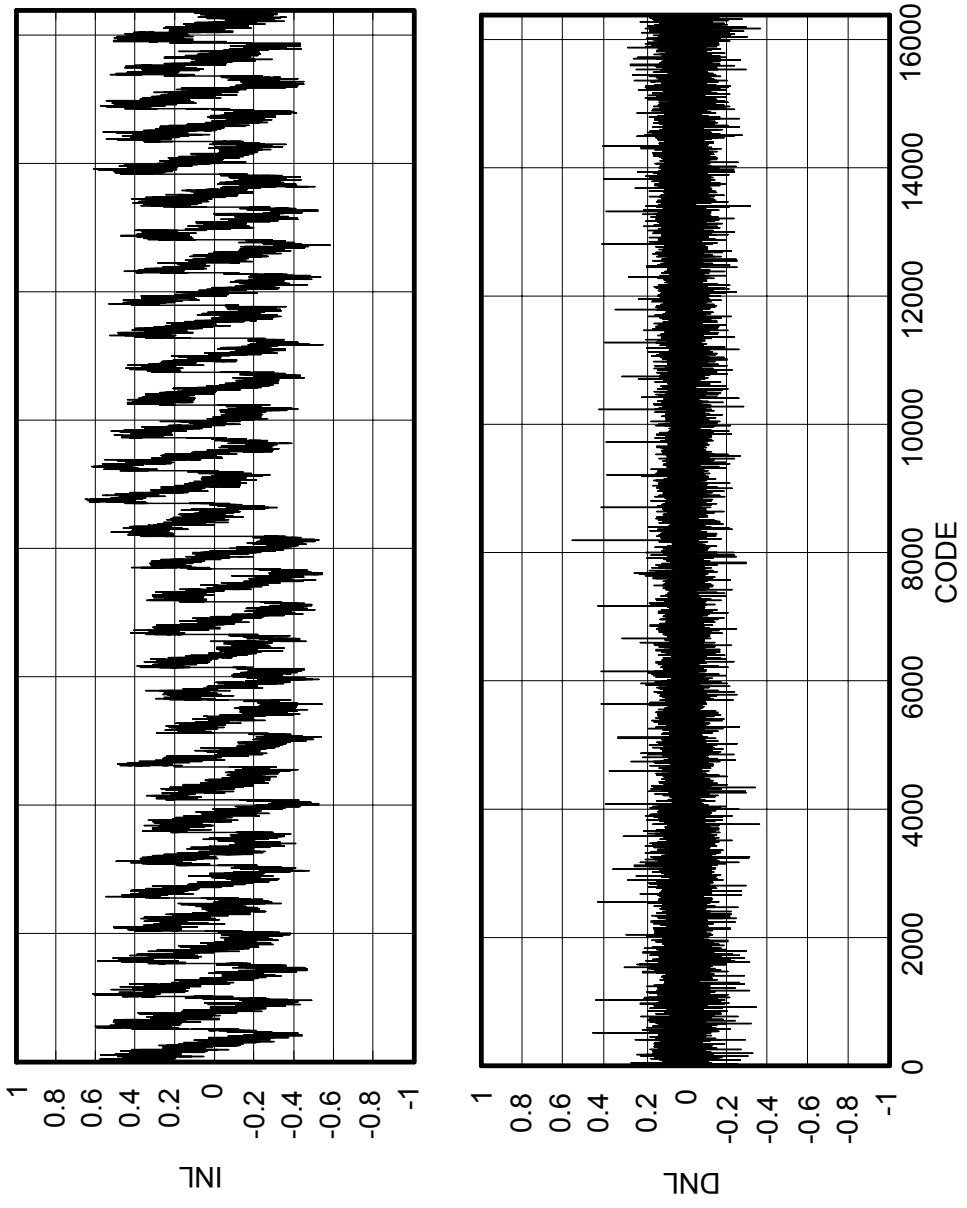


Figure 20.3.3: INL and DNL

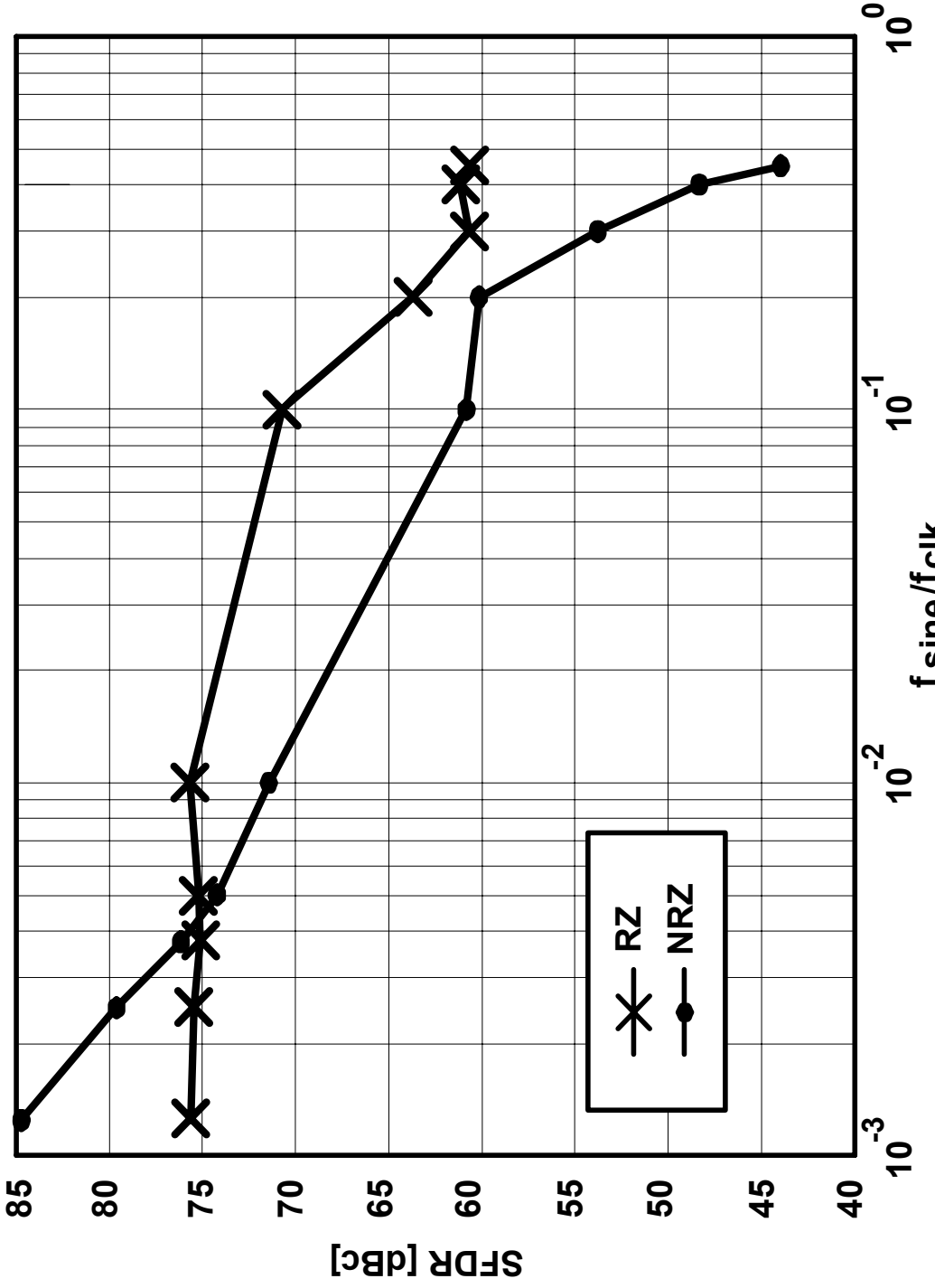


Figure 20.3.4: SFDR RZ/NRZ

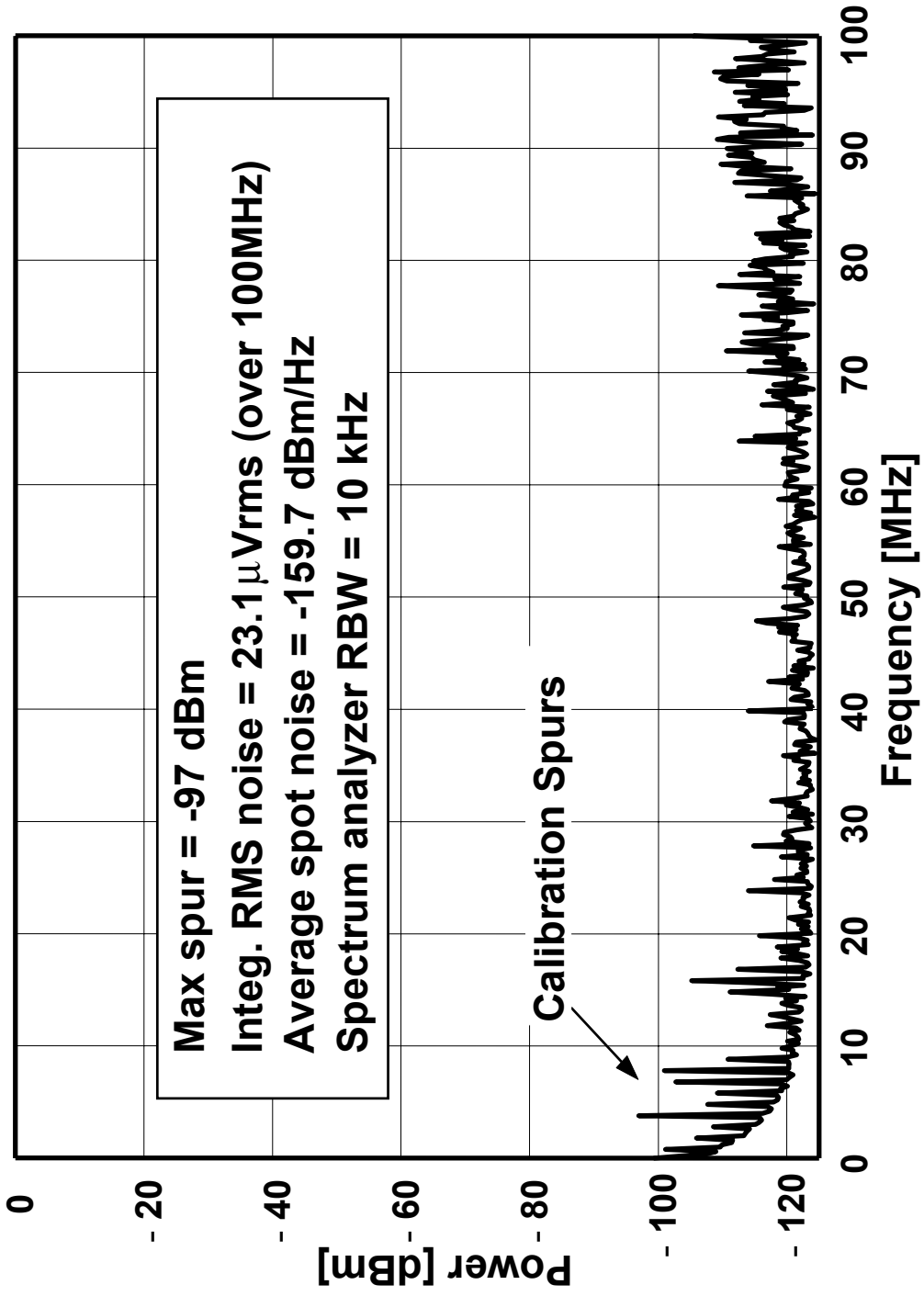


Figure 20.3.5: Noise spectrum

Max SFDR (NRZ)	85 dB
Max SFDR (RZ)	76 dB
Average spot noise	-160 dBm/Hz
Output current	16 mApp
Power supply	1.8 V
Power consumption	97 mW
Process	0.18μm 1P6M CMOS
Core area	1 mm^2

Figure 20.3.6: Performance summary

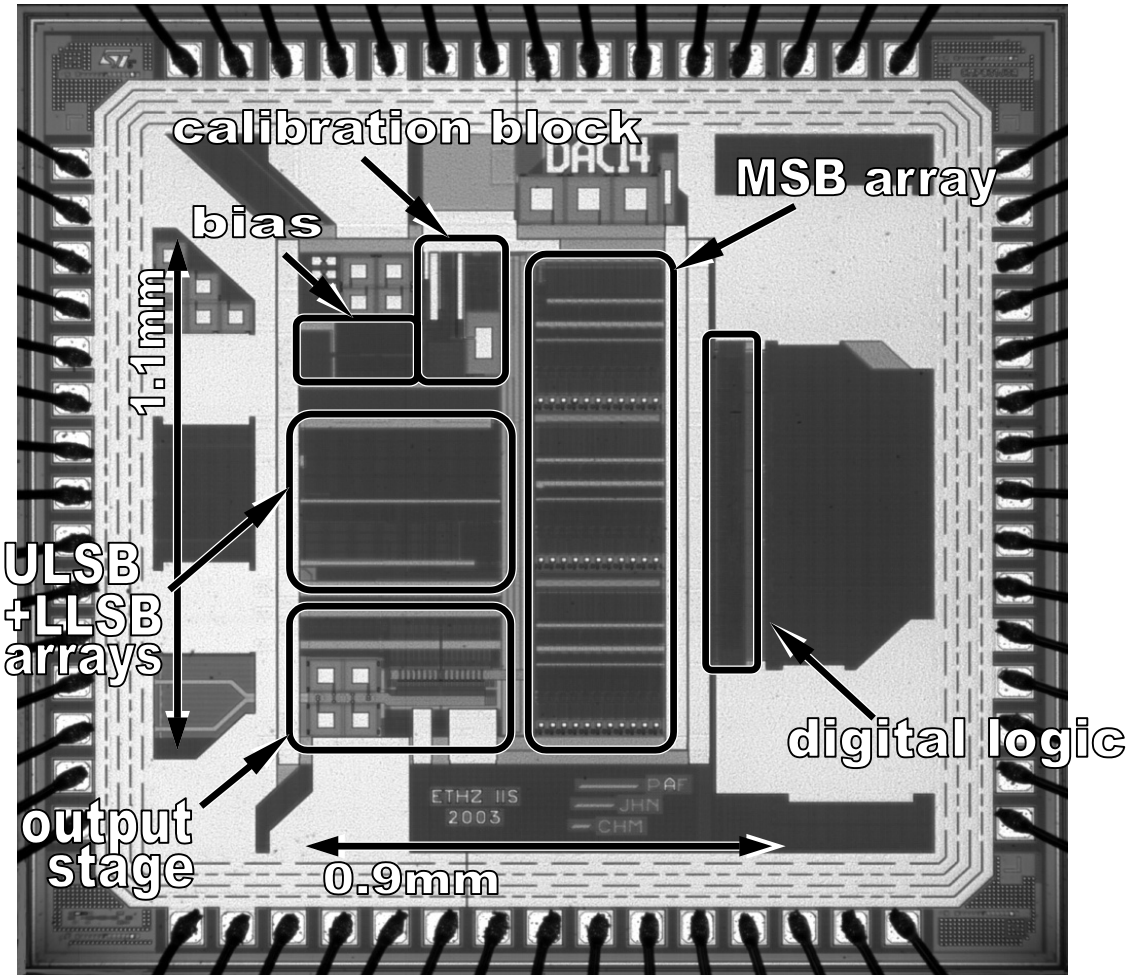


Figure 20.3.7: Chip micrograph

A LOW-POWER 10-BIT CONTINUOUS-TIME CMOS $\Sigma\Delta$ A/D CONVERTER

Jannik Hammel Nielsen and Erik Bruun

Ørsted•DTU, Technical University of Denmark
DK-2800 Kgs. Lyngby, Denmark
jhn@oersted.dtu.dk

ABSTRACT

This paper presents the design of a third order low-pass $\Sigma\Delta$ analog-to-digital converter (ADC) employing a continuous-time (CT) loop filter. The loop filter is implemented using $G_{\mathrm{m}} - C$ integrators, where the transconductors are implemented using CMOS transistors only. System level as well as transistor level design issues for power efficiency is discussed. A prototype $\Sigma\Delta$ ADC intended for weak biological signals restricted to bandwidths below $4\,\mathrm{kHz}$ has been manufactured in a standard $0.35\,\mu\mathrm{m}$ CMOS technology. The ADC has a measured resolution of 10 bits and a dynamic range (DR) of 67 dB at a sampling rate of $f_{\mathrm{s}} = 1.4\,\mathrm{MHz}$, while drawing a bias current of $60\,\mu\mathrm{A}$ from a modest supply voltage of $1.8\,\mathrm{V}$, thus consuming $108\,\mu\mathrm{W}$ of power.

1. INTRODUCTION

The last few decades has seen a rapidly growing interest in the implementation of systems-on-chip for sensing, recording and processing of various physical signals, ranging from micromachined silicon sensor readouts to biological signals. A typical configuration for these systems consists of an analog frontend for preprocessing the often weak signal input followed by an A/D converter leaving the signal for further digital processing. As many of these systems are designed for portability, low power consumption of the system components is crucial.

In this paper we propose a continuous-time $\Sigma\Delta$ A/D converter suitable for quantizing preamplified nerve signals recorded by cuff-electrodes. The main energy of human nerve signals lie in the frequency band $400\,\mathrm{Hz} < f_{\mathrm{sig}} < 4\,\mathrm{kHz}$ and has a typical maximum amplitude of $\pm 10\,\mu\mathrm{V}$ [1]. A preamplifier which brings the nerve signal to $\pm 100\,\mathrm{mV}$ has previously been implemented and is reported in [2].

CT loop filters holds several advantages over their discrete-time counterparts. As the sampler is located deep inside the modulator, implicit anti-aliasing is provided by the CT loop filter. Furthermore, the gain-bandwidth-product (GBW) of the active element in the CT loop filter can be significantly lowered compared to traditional switch-capacitor implementations for a given sampling frequency implying power savings [3]. The drawbacks of using a CT loop filter is increased sensitivity to clock jitter and loop delay.

For low power implementation of $\Sigma\Delta$-modulators, the single-loop architecture as shown in fig.1 is preferred, as only the first integrator in the loop has strict requirements imposed on it. The requirements for the remaining integrators can be lessened as all internal nodes are filtered to the output.

This work was supported by the Danish Medical Research Council

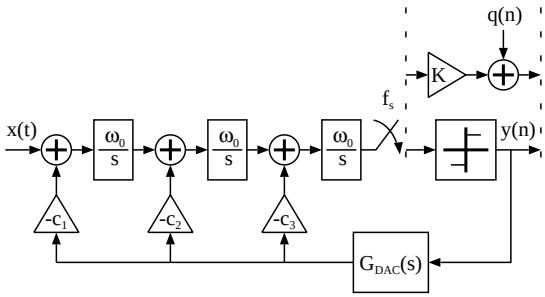


Fig. 1. Continuous-time 3^{rd}order $\Sigma\Delta$- modulator block diagram.

2. CT $\Sigma\Delta$-MODULATOR SYSTEM DESIGN

As shown in fig.1, the quantizer is modelled by a gain factor K and an additive white noise source $q(n)$. The feedback DAC is a NRZ zero-order hold waveform and the s-domain transfer function of the DAC is given by: $G_{\mathrm{DAC}}(s) = [1 - \exp(-sT)]/s$. The overall loop filter function can be written:

$$H(s) = \sum_{n=1}^{3} c_{\mathrm{n}} \left(\frac{\omega_0}{s}\right)^n G_{\mathrm{DAC}}(s) \tag{1}$$

The signal transfer function (STF) and noise transfer function (NTF) from a loop-filter perspective are respectively:

$$\mathrm{STF}(s) = \frac{KH(s)}{1 + KH(s)} \approx 1, \quad \mathrm{NTF}(s) = \frac{1}{1 + KH(s)} \tag{2}$$

By choosing an appropriate NTF, e.g. a butterworth highpass filter, the feedback coefficients can be derived [3].

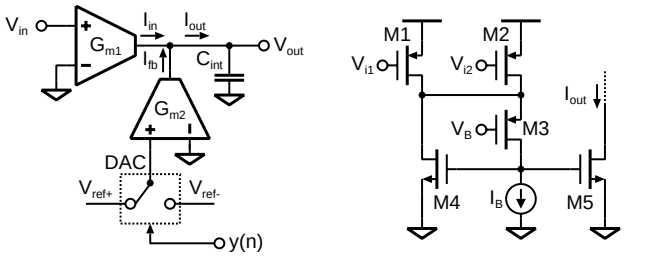
$\Sigma\Delta$-modulators are characterized by having a maximum stable input amplitude (MSA), which will cause the modulator to become unstable when exceeded. The MSA is inversely proportional to the loop filter order and cutoff frequency ω_0 [3], i.e. the more aggressively we high-pass filter the quantization noise, the less signal power can be applied while maintaining stable operation [3].

For a given quantization noise power σ_{q}^2, the inband quantization noise power for a channel $[f_{\mathrm{low}};\ f_{\mathrm{high}}]$ can be found:

$$P_{\mathrm{Q}} = \frac{\sigma_{\mathrm{q}}^2}{f_{\mathrm{s}}} \int_{f_{\mathrm{low}}}^{f_{\mathrm{high}}} \left| \frac{1}{1 + KH\left(e^{j2\pi f}\right)} \right|^2 df \tag{3}$$

The resulting maximum signal-to-quantization-noise-ratio (SQNR) at the MSA is then:

$$\mathrm{SQNR} = 10 \log \left(\frac{\mathrm{MSA}^2}{P_{\mathrm{Q}}} \right) \tag{4}$$



(a) Integrator block diagram (b) Transconductor schematic

Fig. 2. Continuous-time integrator with feedback and corresponding single-ended dual input transconductor MOST schematic.

The optimum loop filter order, sampling frequency and loop filter cut-off frequency ω_0, with regards to minimizing the power consumption for a required SQNR can be found by using the design methodology described in [3]. For 10 bits performance for our given bandwidth, a third order loop filter with a cut-off frequency of $80\,\mathrm{kHz}$ was found to be optimal as a second order filter would require a high sampling frequency, whereas a fourth order filter would degrade stability too much.

3. IMPLEMENTATION

3.1. Transconductor Core

The integrators were implemented as $G_\mathrm{m} - C$ structures. A principle block diagram of an integrator is shown in fig. 2(a), where the integrator unity-gain frequency is given by $\omega_0 = G_{\mathrm{m}1}/C_{\mathrm{int}}$. The integrator is shown in a single-ended version for simplicity whereas the implemented version is fully differential. The integrator employs two transconductors for the signal input and the modulator DAC feedback respectively. The feedback DAC function is implemented by letting the ADC output $y(n)$, control a switch selecting either a positive or negative voltage reference $V_{\mathrm{ref+/-}}$. The feedback factor is determined by $c = G_{\mathrm{m}2}/G_{\mathrm{m}1}$.

The two transconductors employed in an integrator were implemented in a single MOST-only structure by using an extended version of the transconductor introduced in [4], shown in a single-ended version in fig. 2(b). The transconductors $G_{\mathrm{m}1}$ and $G_{\mathrm{m}2}$ of fig. 2(a) are realized by biasing M1 and M2 in the triode region thus realizing a two-input transconductor. The drain voltage of M1 and M2 is set by the bias voltage V_B and held constant through the negative feedback loop comprising M3 and M4. M3 is biased in weak inversion to maximize the $g_\mathrm{m}/I_\mathrm{D}$-ratio whereas all other MOSTs are biased in strong inversion for good matching. Hence, the source node of M3 constitutes a virtual ground node well suited for current summing. The signal currents produced by M1 and M2 are thus sensed by M3 resulting in a voltage change at the gate of M3 modulating the M3 current i_DS. The summed signal current can be accessed through M4 which mirrors the M3 i_DS.

Using the EKV-model, the drain current for a PMOST biased in the triode region is given by [5]:

$$I_\mathrm{D} = K_\mathrm{P}^{'}\,(W/L)\left[V_\mathrm{G} - V_{\mathrm{th,p}} + \frac{n}{2}\,(V_\mathrm{D} - V_\mathrm{S})\right](V_\mathrm{D} - V_\mathrm{S}) \tag{5}$$

where $V_{\mathrm{th,p}}$ is the threshold voltage, n is the slope factor and all terminal voltages are referred to the bulk. Eq. (5) is valid for $V_\mathrm{G} - V_{\mathrm{th,p}} \leq n\,(V_\mathrm{D} - V_\mathrm{S})$. From eq. (5) it is seen that the drain current varies linearly with the gate voltage if the drain-source voltage is held constant. The transconductance of M1 and M2 constituting $G_{\mathrm{m}1}$ and $G_{\mathrm{m}2}$ in the overall dual input transconductor is equal to the triode region small-signal transconductance:

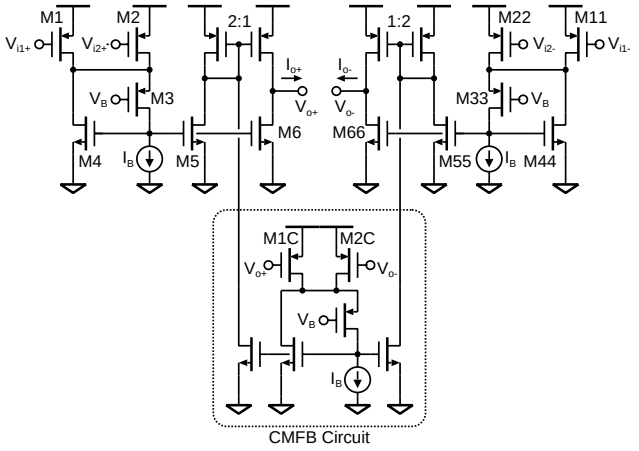


Fig. 3. Fully differential dual-input transconductor including CMFB circuit.

$$G_{\mathrm{m,single}} = g_\mathrm{m} = \frac{\partial I_\mathrm{D}}{\partial V_\mathrm{G}} = K_\mathrm{P}^{'}\,(W/L)\,(V_\mathrm{D} - V_\mathrm{S}) \tag{6}$$

From eq. (6) we see that the transconductance is tunable by regulating the drain-source voltage enabling us to tune the loop filter cut-off frequency ω_0 to the desired frequency.

The necessary bias current of the transconductor core is determined by the speed requirements of the negative feedback loop. An approximate expression for the loop GBW is given by: $\mathrm{GBW} = g_{\mathrm{m}4}/(2c_{\mathrm{gs}4} + c_{\mathrm{gd}3} + c_{\mathrm{IB}})$. Voltage excursions on the drain of M1 and M2 will result in modulation of the transconductor output current given by: $\Delta i_{\mathrm{ds}1,2} = \Delta v_{\mathrm{ds}1,2}\,(g_{\mathrm{ds}1} + g_{\mathrm{ds}2})$. The amplitude of the voltage excursions is given by regulated impedance seen at the source node of M3:

$$Z_\mathrm{x}(\omega) = \frac{\omega}{g_{\mathrm{m}3}\mathrm{GBW}} \tag{7}$$

From eq.(7), we see that the impedance can be lowered by increasing the loop GBW and $g_{\mathrm{m}3}$, hence giving us the required bias current necessary for a given input frequency.

3.2. Differential Transconductor

A fully differential version of the dual-input transconductor can be easily realized by duplicating the circuit of fig. 2 as shown in fig. 3. Common mode stabilization is achieved by including a common-mode-feedback (CMFB) circuit which is constructed by replicating the transconductor core as seen in fig. 3. Output common-mode voltage deviation is sensed by M1C and M2C which injects current through the 2:1 mirrors to correct the output voltage. The aspect ratio of M1C and M2C are set equal to half the aspect ratio sum of M1 and M2. So we have $W_{1\mathrm{C}} = W_{2\mathrm{C}} = (W_1 + W_2)/2$, hereby ensuring that the correct common-mode level is set for the succeeding integrator.

The 2:1 mirrors employed by the CMFB circuit will also half the small signal current, hence the differential transconductor has

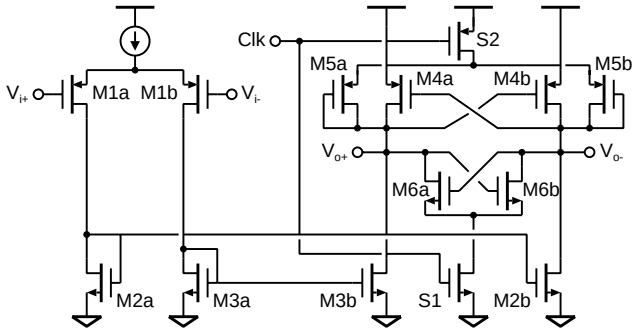


Fig. 4. Comparator schematic.

transconductance equal to half the single-ended version:

$$G_{\mathrm{m,diff}} = 1/2 G_{\mathrm{m,single}} = 1/2 g_{\mathrm{m1}} \tag{8}$$

Due to the filtering of the internal nodes in the loop filter, the requirements for the 2nd and 3rd transconductors can be relaxed. A summary of the transconductor performances is given below:

	$I_{\mathrm{B}}\,[\mu\mathrm{A}]$	Reg. GBW [MHz]	DC-gain A_0[dB]
G_1	3	10	61
G_2	1,25	4.5	61
G_3	0.75	2.7	48

3.3. Comparator

As the quantizer in the ADC modulator consists of a single bit, a comparator suffices. A schematic of the realized comparator is shown in fig. 4. It consists of a preamplifier stage and a track and latch output stage. During the low state of the clock, the M6 positive feedback pair is switched off and the M5 pair is switched on. The loop gain around the M4 and M5 pairs is in this case <1 and the comparator operates in track mode with a gain given by:

$$A_{\mathrm{V}} = \frac{v_{\mathrm{o}}}{v_{\mathrm{i}}} = g_{\mathrm{m1}} \frac{g_{\mathrm{m4}} + g_{\mathrm{m5}}}{g_{\mathrm{m5}}^2 - g_{\mathrm{m4}}^2} \approx \frac{g_{\mathrm{m1}}}{g_{\mathrm{m5}}} \tag{9}$$

As the clock goes high, switch S2 turns off and the loop gain around the M4 pair is now >1. Switch S1 turns on enabling the M6 positive feedback pair and the comparator output is latched. A digital set-reset latch succeeds each of the comparator outputs in order to restore the output to full digital levels.

3.4. Capacitor dimensioning

A figure of merit for transconductors is the noise-excess-factor (NEF) which is defined as the ratio of transconductor output thermal noise conductance to the transconductance:

$$\mathrm{NEF} = \frac{\mathrm{G_{N,out}}}{G_{\mathrm{m}}} \tag{10}$$

For the employed transconductor, the NEF can be shown to be [4]:

$$\mathrm{NEF} = \frac{g_{\mathrm{ds1}}}{g_{\mathrm{m1}}} \left(1 + \frac{g_{\mathrm{ds1}}}{g_{\mathrm{m3}}} \right) + 2\frac{g_{\mathrm{m4}}}{g_{\mathrm{m1}}} + \frac{g_{\mathrm{m,IB}}}{g_{\mathrm{m1}}} \tag{11}$$

The thermal noise accumulated on the integrating capacitor is:

$$P_{\mathrm{Th}} = \frac{\mathrm{NEF} \cdot kT}{C_{\mathrm{int}}} \tag{12}$$

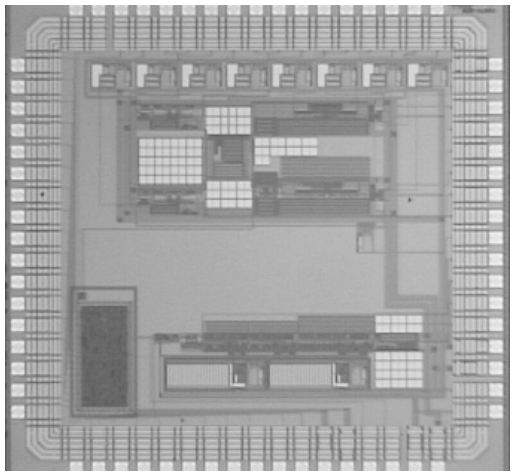


Fig. 5. Chip die microphotograph.

Referred to a given reference voltage V_{ref}, the total amount of noise allowable for a resolution of b bits is given by:

$$P_{\mathrm{N}} = \frac{(\mathrm{MSA} \cdot V_{\mathrm{ref}})^2}{2 \cdot 10^{(6.02b+1.76)/10}} \tag{13}$$

We have previously derived an expression for the inband quantization noise eq. (3). Combining equations (12-13) and eq. (3), we can solve for the minimum allowable capacitor size:

$$C_{\mathrm{int}} = \mathrm{NEF} \cdot kT \left[\frac{(\mathrm{MSA} \cdot V_{\mathrm{ref}})^2}{2 \cdot 10^{(6.02b+1.76)/10}} - P_{\mathrm{Q}} \right]^{-1} \tag{14}$$

The noise filtering of the intenal nodes allows for smaller capacitors in the 2nd and 3rd integrators. The implemented capacitor sizes are given below:

	C_1	C_2	C_3
Cap. size [pF]	7.5	1.75	1.25

4. MEASUREMENTS

The circuit was implemented in a standard $0.35\,\mu\mathrm{m}$ CMOS process. A chip microphotograph is shown in fig. 5, where the $\Sigma\Delta$-modulator is contained in the lower half of the chip.

An example output spectrum is shown in fig. 6, for an input tone at $1.3\,\mathrm{kHz}$ with an amplitude of $0.8 V_{\mathrm{ref}}$. It is seen that the spurious performance is dominated by the second harmonic.

Fig. 7 shows the measured signal-to-noise-plus-distortion ratio (SNDR) for two different tone frequencies versus the normalized input signal $V_{\mathrm{in}}/V_{\mathrm{ref}}$ for one of the test chips. Though the harmonic performance is unexpectedly dominated by the second harmonic we still achieve a max. SNDR of approx. $62\,\mathrm{dB}$at the low tone frequency at $340\,\mathrm{Hz}$, which deteriorates to approx. $56\,\mathrm{dB}$ for the higher tone frequency at $1300\,\mathrm{Hz}$. The average max. SNDR over the ten test chips for an input tone of $1864\,\mathrm{Hz}$ was measured at $58.4\,\mathrm{dB}$.

The measured average current drawn in the ten test chips is $60\,\mu$Afrom a $1.8\,\mathrm{V}$supply giving an average power consumption of $P = 108\,\mu\mathrm{W}$, excluding bias circuitry and input buffers.

The measured performance of the circuit is summarized in table 1 together with results of recent similar work.

	This work	[6]	[7]	[8]	[9]
Power Dissipation [μW]	108	135	950	340	40
Signal Bandwidth [kHz]	3.6	25	25	8	16
Sampling Frequency [MHz]	1.4	2.4	5	1.024	1.538
Peak SNDR [dB]	62	70	85	69	62
$V_{\rm DD} - V_{\rm SS}$ [V]	1.8	1.5	1	1.95	0.9
Dynamic Range [dB]	67	80	88	73	77
Technology	$0.35\,\mu$m CMOS	$0.5\,\mu$m CMOS	$0.35\,\mu$m CMOS	$1.2\,\mu$m CMOS	$0.5\,\mu$m CMOS

Table 1. Measured performance summary of the third order $\Sigma\Delta$-modulator and comparison with prior work.

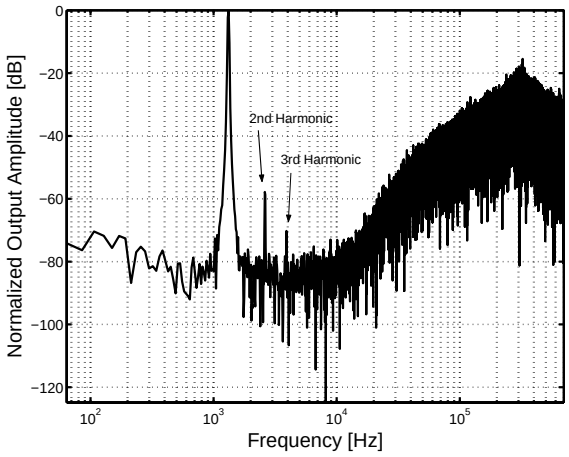


Fig. 6. Measured output spectrum.

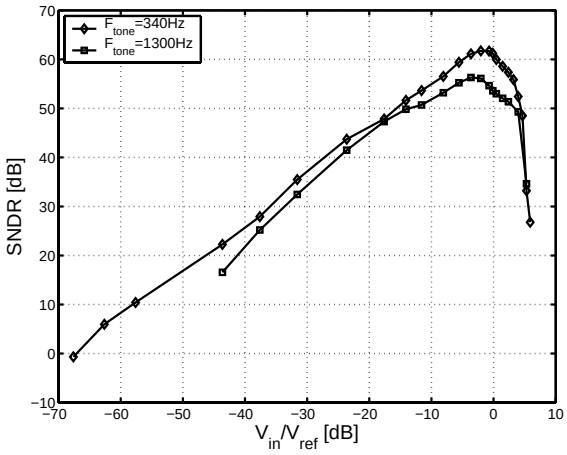


Fig. 7. SNDR for two different input tone frequencies.

5. CONCLUSION

A low-power CT ADC has been implemented using $G_{\rm m} - C$ integrators. The transconductors were implemented using MOST only, whereas the integrating capacitors were implemented as poly-poly capacitors. However as all capacitors are grounded, poly-well capacitors are feasible for a MOST only implementation of the ADC thus making it suitable for implementation in pure digital CMOS standard processes. Comparison with prior work shows that CT loop-filters are a good alternative to switch-capacitor DT loop-filters for low-power $\Sigma\Delta$ ADCs. The transconductor employed is reported operational at a supply of 1.2 V in a $0.18\,\mu$m technology [4], implying that sub $75\,\mu$A operation of the ADC could be achieved.

6. REFERENCES

[1] J. Struijk et al., "Cuff electrodes for long-term recording of natural sensory information," *IEEE Eng. in Medicine and Biology*, vol. 18, pp. 91–98, 1999.

[2] J. H. Nielsen and T. Lehmann, "An implantable CMOS amplifier for nerve signals," *Analog Int. Circuits and Signal Proc.*, vol. 36, pp. 153–164, 2003.

[3] J. H. Nielsen and E. Bruun, "A design methodology for power-efficient continuous-time sigma-delta a/d converters," in *Proc. Int. Symp. on Circuits and Systems*. IEEE, 2003, vol. 1, pp. 1069–1072, IEEE.

[4] U. Yodprasit and C. Enz, "A 1.5V 75-dB dynamic range third-order Gm-C filter integrated in a standard 0.18-μm standard digital CMOS process," in *Proc. European Solid-State Circuits Conference*, 2002, pp. 647–650.

[5] C. Enz et al., "An analytical CMOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications," *Analog Int. Circuits and Signal Proc.*, vol. 8, pp. 83–114, 1995.

[6] F. Gerfers et al., "A 1.5V 12-bit power-efficient continuous-time third-order $\Sigma\Delta$ modulator," *IEEE Journal of Solid-State Circ.*, vol. 38, pp. 1343–1352, 2003.

[7] M. Dessouky and A. Kaiser, "Very low-voltage digital-audio $\Sigma\Delta$ modulator with 88-dB dynamic range using local switch bootstrapping," *IEEE Journal of Solid-State Circ.*, vol. 36, pp. 349–355, 2001.

[8] S. Au and B. H. Leung, "A 1.95-V, 0.34-mW, 12-b sigma-delta modulator stabilized by local feedback loops," *IEEE Journal of Solid-State Circ.*, vol. 32, pp. 321–328, 1997.

[9] V. Peluso et al., "A 900-mV low power $\Sigma\Delta$ A/D converter with 77-dB dynamic range," *IEEE Journal of Solid-State Circ.*, vol. 33, pp. 1887–1897, 1998.