UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL AND APPLIED SCIENCES

Electronics and Computer Science

Design Methods to Mitigate the Effects of Variation in Analogue and Mixed-Signal Circuits

by

Robert M. Rudolf

Thesis for the degree of Doctor of Philosophy

November 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL AND APPLIED SCIENCES
Electronics and Computer Science

Doctor of Philosophy

DESIGN METHODS TO MITIGATE THE EFFECTS OF VARIATION IN ANALOGUE AND MIXED-SIGNAL CIRCUITS

by Robert M. Rudolf

The continued scaling of CMOS process features enables ever-faster and denser circuits, which comes at the cost of increased device parameter variation. The performance of analogue and mixed-signal circuits in particular degrades in such a high variation environment, which poses an extraordinary challenge in the design and fabrication of such circuits.

This thesis develops a set of tools and methodologies for a post-fabrication calibration system called the Configurable Analogue Transistor (CAT). The principle of the CAT technique is to replace certain transistors in a circuit with calibration devices, which allow adjustment of circuit performance after fabrication to compensate the effects of device parameter variation. Building on initial research on the CAT, this thesis develops a methodology to identify the most suitable calibration devices in their circuit and determine their optimal sizes. Furthermore, the applicability of CAT is extended beyond parameter variation to also include direct compensation of temperature.

A complementary technique to post-fabrication calibration is robust design, where a circuit is designed to be inherently robust against variation in device parameters. In this thesis, a novel closed-loop pick-off circuit for force-balanced MEMS accelerometers is presented. It is comparable in performance to other state-of-the-art techniques, but provides vastly improved robustness against parameter variation and a more intuitive design process.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Robert M. Rudolf, declare that the thesis entitled *Design Methods to Mitigate the Effects of Variation in Analogue and Mixed-Signal Circuits* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as: [1], [2], [3] and [4]

Signed:.....

Date:.....

Acknowledgements

I would like to express my gratitude towards my supervisor, Dr Peter Wilson, for supporting my research in every way possible to him. In addition to encouragement and practical matters, Peter supported this work through countless coffee-fuelled discussions.

Furthermore, I would like to thank Dr Reuben Wilcock for the original work on the Configurable Analogue Transistor, the many hours of inspiring discussions and plentiful advice I received from him. Over the years, Reuben has become a good friend and I owe much of my professional development to him.

Dr Li Ke (李科) has taught me practical analogue IC design and was always there to answer my questions, no matter how stupid they were. I am grateful for his patience and knowledge which made him an outstanding teacher.

Finally, I would like to thank my mother who was always there for me during the last years and who continues to support me in every way possible to her. I am also indebted to my late father, who sparked my interest in Engineering and who would have been proud to see this work completed. I would not have been able to finish my research without their support and therefore dedicate this work to my dear parents.

Für Papa

Chapter 1

Introduction

The continued development of modern microelectronics depends on scaling of CMOS processes to satisfy the ever increasing demand for faster and lower-power devices with more functionality. For decades, every new generation of process technology has enabled us to put more processing power on smaller chips while at the same time reducing power consumption. As much as both analogue and digital circuits have benefited from advanced processes [5], the height of physical obstacles that need to be overcome has increased likewise. Device variability in particular is making the economic design of analogue circuits on highly integrated system-on-chips (SoC) extremely challenging. [6]

Ever since the electronic revolution started with vacuum-tube based radio receivers in the early 20^{th} century, electronic components' inherent tolerances, variability, ageing and susceptibility to the environment have challenged the manufacturers of mass-produced electronic devices. Practices such as trimming of components in the factory, labour-intensive as they may be, can be employed to manufacture and ship working units, but when these units then fail in a customer's hands because a critical component value has drifted out of specification, the reputation of the entire brand is in danger of being tainted. Therefore, designers and manufacturers have always sought to improve the reliability of their devices, be it only good enough to survive the warranty period.

This improvement has come in a number of forms. For instance, tolerances and temperature coefficients of active and passive components have been consistently improved while at the same time maintaining or lowering cost, such that 1% metal film resistors are the norm for prototyping in many labs today, where only 25 years ago 10% or worse carbon resistors would have been used and metal-film resistors were reserved exclusively for high-accuracy applications. On the other hand, it is often ingenious circuit design that is required to overcome limitations in components and manufacturing, either because the cost of better components or techniques would be prohibitive, or simply because they are not available. One famous example is the Hewlett-Packard 200A Audio Oscillator, which used a cheap, off-the-shelf incandescent light bulb as a PTC resistor to stabilise the gain

of the Wien bridge oscillator. This allowed the HP 200A to not only provide a more stable, less distorted signal than many competing products but, more importantly, it did so at a price point far below that of comparable products. Thus, a simple yet ingenious way of compensating for component variability through design laid the foundations for one of the largest electronics business still in existence today.

Figure 1.1: Hewlett-Packard 200A Audio Oscillator *by Colin CC BY-SA* [7]

In modern electronics, the performance of discrete components is nowhere near as significant a concern as it was in 1939 when William Hewlett filed his patent for a stabilised oscillator [8]. Standard passive and active components are generally good enough to be used in any application that has modest requirements on precision, while the design-in of lower-specification components is normally only considered for high-volume consumer products where cost optimisation is vital. Conversely, components with almost arbitrarily high specifications can be readily purchased, assuming one is willing to accept the consequences on the bill-of-materials cost. Instead, the real challenge today and the focus of this work is on the variability of transistors and other components on integrated circuits of increasingly smaller feature sizes that are at the heart of our connected lives.

1.1 CMOS Device Variability

As feature sizes of integrated circuits decrease to the sub-micron region, circuit designers are faced with a variety of challenges arising from such small feature sizes [9]. One key contributor to these challenges is increased variability in nano-scale CMOS circuits.

Variability is caused by a range of physical and manufacturing processes and can manifest itself in a number of ways. Depending on the focus of discussion, variability can be

categorised in several ways, which all have their individual aspects. For the purpose of this discussion, variability is categorised according to its origins during the manufacturing process and during use, which is suggested in [10].

Intrinsic variability covers random processes that occur within each device. Most of these processes are linked to atomic-scale uncertainty in the manufacturing process. Examples of such atomistic effects are dopant distribution, film thickness and line-edge roughness. Random dopant distribution occurs when the active area of a device is very small, reducing the number of dopant atoms in the channel to only a few 100 or below. On these scales, a difference of a few dopant atoms can result in significant parameter variation between devices. Similarly, the equivalent gate oxide thickness is often only a few atoms with an interface roughness of one or two atoms, resulting in significant random variation of oxide thickness. Line-edge roughness (LER) is introduced by the photolithographic process when feature sizes are comparable to the wavelength used in the exposure, resulting in random variation in feature outlines. These processes are the ones that are most commonly associated with variability and give rise to device parameter variation and contribute to mismatch variation, which are usually modelled in IC design kits [11, 12].

Extrinsic variability, on the other hand is caused by variations during the manufacturing process, differences in the chemical processes between wafers or dopant gradients on a wafer. These effects can occur anywhere from chip scale, over chip-to-chip, wafer-to-wafer to batch-to-batch scale. On the larger scales, all devices within a chip are affected equally, which may result in overall parameter changes, but not necessarily in mismatch. For example, an array of matched current sources may have significantly different output currents on a wafer-to-wafer comparison, while the matching of the sources within each chip is not significantly affected. In IC design kits, this is generally modelled as process variation. However, extrinsic variability can affect matching between devices when it is significant on the wafer- or chip-scale. In this case, the placement of a device on the chip or the physical layout of a device and its surroundings can affect its relative parameters. One often-cited example is chemical-mechanical polishing, where the resultant layer thickness is dependent on the feature density in an area. Another example are dopant gradients over the scale of the chip, where device mismatch increases with separation. Such placement-induced variation is the reason for certain design techniques in analogue circuits, such as common-centroid layout to negate parameter gradients or dummy devices to ensure a homogeneous feature pattern around important devices.

Intrinsic and extrinsic variability cover effects that occur during the fabrication of a chip. There are additional processes that describe the change in parameters of a device after fabrication. These can be further divided into wear-out effects and use-induced effects. Wear-out covers many non-reversible temporal processes related to ageing and device stress. These can result in a drift of circuit parameters over time and ultimately lead to

failure of the circuit. The prime example for this is Negative-Bias Temperature Instability (NTBI) [13], where device performance continually degrades under the presence of negative gate bias and elevated temperatures. In contrast, use-induced variability generally covers reversible temporal effects, such as parameter dependence on temperature. Since the power consumption of a chip, and therefore its die temperature, may vary depending on usage, this type of variation can depend on the particular state of a circuit or the operation it is currently performing.

Since digital circuits are at the forefront of process scaling, the effects of variability on digital circuits and their mitigation are the subject of significant research. Regardless of the process that leads to parameter variation, there are two main areas in which digital circuits are generally affected. First, the delay time of any digital circuit depends on the parameters of its constituent transistors. If these transistor parameters are affected by variability, so is the delay time of gates and higher-level circuit blocks. Should the variability then cause the delay time on the critical path of a particular circuit to exceed the design margin, the circuit will no longer operate correctly. Generally, the consequence is then that the circuit has to be operated at a lower clock speed to satisfy timing requirements again. If the original clock speed must be maintained, it is possible to reduce the relative effects of timing variability by increasing the operating voltage, which reduces the initial delay by increasing the device current at the expense of increased power consumption of the chip.

The second main effect of variability on digital circuits is a reduced noise margin. Since the output voltage levels and detectable input logic voltages of a digital gate depend on the parameters of its transistors, these properties are subject to the same variation as the transistor parameters. This in turn means that under the right conditions, the margins between the output levels of one gate and on input levels of another gate can be reduced, leading to reduced noise immunity or even complete device failure. The noise margin can be increased by increasing the supply voltage, which again results in an increased power consumption of the circuit. [14]

In addition to variability, there is a wide spectrum of other challenges that arise from CMOS device scaling. Indeed, for high-performance digital circuits, variability is only one among several key challenges, which have been the subject of sustained research to enable the continued improvement of circuit performance. The most important group of these are generally categorised as short-channel effects and a brief overview of some of these mechanisms is given below.

In classical scaling, reducing the channel length of MOS devices also necessitates a reduction in the supply voltage to maintain a reasonable electric field strength across the channel. Since the supply voltage should be significantly greater than the threshold voltage, a reduction in threshold voltage is also required. However, lowering the threshold

voltage increases the sub-threshold drain-source current, which results in a higher off-state device current. For short channels, the threshold voltage further reduces with drain voltage, in which case the effect is referred to as drain-induced barrier lowering (DIBL). Additionally, gate oxide thickness is also scaled with the supply voltage, which leads to increased gate leakage (tunnelling) currents and thus to an even higher power consumption. There are also several mechanisms reducing carrier mobility and ultimately leading to reduced transconductance and output impedance. Another short-channel effect is hot electrons, which describes the effect when electrons are accelerated by the strong fields in short channel devices such that they are trapped in the gate oxide, which also leads to a degraded performance.

The solutions for these issues are generally found in new materials or device structures. For instance, carrier mobility can be improved again by strained silicon or materials with higher inherent mobility. High-k metal gates can be employed to improve threshold voltage behaviour. Modified channel structures, such as thin silicon channels or multi-gate transistors further improve sub-threshold behaviour, leading to a reduction in sub-threshold conduction. Such technologies are vital to maintain scaling of technology nodes and are already commonplace in high-performance digital ICs, such as processors. However, as will be discussed in the following section, the main obstacle for effectively implementing analogue circuits on small-scale process nodes is variability, which is even more problematic for analogue circuits than it is for digital circuits.

1.2 Analogue Circuits on Nano-Scale CMOS

Today's highly integrated and compact consumer electronics would not be possible without mixed-signal devices and system-on-chips (SoC), where analogue circuits are integrated on the same chip as digital circuits. These high-performance digital components are often bound to small-scale process nodes in order to allow operation at the desired high frequencies. However, this also means that the analogue components must be built on the same process and therefore subjects the analogue devices to the problematic levels variability that are inherent to these processes. Although the physical and device-level mechanisms of variation are the same in for analogue and digital circuits, their effects on analogue circuits are more subtle, which leads to a wide range of techniques to handle them.

There are several different ways in which the different kinds of variability affect the performance of analogue circuits. When the effects of extrinsic variability do not cause mismatch on the chip-scale, they generally affect overall circuit performance or operating points. For example, current mirrors that distribute reference currents in a circuit are all affected equally by chip-wide parameter variation, which results in a uniform change in reference currents. This change in reference currents may then affect other parameters,

such as the gain or bandwidth of an amplifier. Random effects generally cause mismatch between devices within a circuit, which results in asymmetries or imbalances where a circuit was designed symmetrically or balanced. One common example of this is the input offset voltage of a differential amplifier, which is due to a mismatch in threshold voltages of its two transistors. A mismatch in transconductance of the differential pair will also cause an asymmetric distribution of bias current, further contributing to the offset voltage.

Depending on the type of variability, there are different ways how analogue circuits can deal with them. The effects of extrinsic variability on the chip scale can normally be reduced through certain design techniques. An example is the addition of dummy devices in the layout, which lead to more homogeneous patterns around the active devices and thus can help reduce variability. Another example is symmetric or common-centroid layout, where the layout of a device is modified such that it is less affected by process gradients. Design techniques can also be used to reduce the effects of intrinsic variability, albeit to a lesser extent. This is usually achieved by making devices physically larger to reduce the impact of atomistic and lithographic effects. However, this is not always effective and can lead to other limitations. First, there are certain properties of a device's size and 3D shape over which the designer does not have any control, for example gate oxide thickness. As stated earlier, nano-scale CMOS processes feature oxide thickness in the order of a nanometre, or a few atoms. This means that a variation in oxide thickness of only one or two atoms between transistors can result in significant parameter mismatch, but the oxide thickness cannot be increased by the circuit designer to reduce this variation. But even increasing device width and length comes with its own problems: If the device size is increased to reduce variation, the gate-source capacitance of the device is similarly increased. For fast analogue signal processing circuits, this increase in capacitance necessitates higher bias currents to maintain the speed of the circuit. Increasing the device size is therefore a trade-off between matching accuracy, speed and power consumption [15].

However, there are further techniques available to help the designer of analogue circuits reduce the impacts of variability. One of the earliest approaches is trimming, where individual component values are adjusted after fabrication. In the case of integrated circuits, however, this technique requires specialised tools and it generally expensive, which is why it is only applied to extremely high accuracy circuits. More commonly, dedicated calibration elements are added to a circuit to allow its adjustment. The advantages over trimming are that no special tools or processes are required during fabrication and that calibration can in principle be carried out at any point after fabrication. Calibration covers a wide spectrum of techniques, ranging from single transistor solutions to reconfigurable systems. A further method to overcome the effects of variation is the capability of a circuit to actively compensated for it. A classic example for this are auto-zero or

chopper-stabilised amplifiers, which use an internal feedback loop to periodically measure the offset voltage and then subtract it from the input voltage, thus cancelling the offset and any associated drift. [16] This example also highlights an important point for any kind of calibration or circuit-level solution: If the system can be calibrated, adjusted or re-configured during run-time, it may not be possible to continue normal operation during that period. In the case of chopper amplifiers where the offset cancellation takes place continuously at a relatively high frequency, the usable signal bandwidth is reduced according to Nyquist's sampling theorem. In other cases where more complex tests are executed for calibration, normal system operation may have to be suspended for a noticeable amount of time.

Apart from calibration, there is a whole different approach to designing circuits for variation: Inherently robust design. In this case, the circuit is designed to minimise the effects of parameter variation on its performance. Increased robustness can be achieved on a number of different levels, from the initial circuit topology, over device sizing, to the physical layout of the circuit on a chip. Chapter 2 will discuss a variety of calibration and robust design techniques.

1.3 Motivation, Organisation and Contributions

The preceding sections of this chapter served to establish that variability is one of the main factors that hamper development of high-performance analogue and mixed-signal circuits on small-scale silicon processes. Devising means to enable the economic production of such circuits has been the focus of significant research in a wide range of fields ranging from semiconductor devices and processes over calibration techniques to variation-tolerant circuits and systems. The motivation of this work is to contribute to these combined efforts, which is achieved by further developing one particular calibration technique known as the Configurable Analogue Transistor (CAT) and by introducing a novel variation-tolerant method for interfacing force-balanced MEMS sensors called Electromechanical PLL (EM-PLL).

The following list outlines the organisation of the remaining chapters of this thesis and briefly summarises any contributions made.

- Chapter 2 provides a more in-depth review of state-of-the-art calibration techniques and basic robust design techniques to complement the general observations made in the preceding sections of this chapter. Since a large portion of this thesis is based on the Configurable Analogue Transistor (CAT), it is reviewed in more detail.
- Chapter 3 introduces three contributions to the existing CAT technique. The first fills a gap in the existing design flow for CAT, the need for an automated algorithm for Critical Device Identification (CDI). CDI is the process of determining which

transistors in a given circuit are most suitable to adjust circuit performance. The proposed algorithm uses Monte Carlo simulation to explore the parameter space and determine which transistors of a circuit are most suitable to be replaced by CAT. The second development shows how the existing algorithm for optimising the size of CATs in current sources can be extended to optimise the size of CATs in a wider range of configurations. The third advancement extends the applicability of the CAT to online calibration for circuits in extreme environments. It is shown through simulation how the CAT can conceptually be used for temperature compensation.

- In Chapter 4, the CAT is applied to a 14-bit digital-to-analogue converter (DAC) to improve linearity. This chapter first reviews the challenges in the design of DACs. It then introduces the basic operation and building blocks of a segmented current-steering DAC and shows how the CAT can be incorporated to adjust the transfer function to improve linearity. Finally, measurements from the fabricated chip are used to prove the feasibility of calibration through CATs for the first time in a complex circuit.

- Chapter 5 shows that a viable alternative to calibration is inherently robust circuit design. This is exemplified by the introduction of a novel interface circuit for differential force-balanced MEMS accelerometers which is inherently more robust against parameter variation than current high-performance solutions. This is in contrast to the calibration-based CAT technique of the previous chapters to illustrate that novel design approaches can also be used to tackle variability issues without the need for calibration, or in systems where calibration is more difficult to apply.

- Finally, Chapter 6 summarises the contributions of Chapters 3, 4 and 5 and outlines further research paths based on this work.

A number of publications have been made based on the work in this thesis, all of which are listed below:

- Rudolf, R.; Taatizadeh, P.; Wilcock, R.; Wilson, P., "Automated critical device identification for configurable analogue transistors," Design, Automation & Test in Europe Conference & Exhibition (DATE), 2012 [1]

- Rudolf, R.; Wilcock, R.; Wilson, P.R., "Reliability improvement and online calibration of ICs using configurable analogue transistors," Aerospace Conference, 2012 IEEE [2]

- Wilson, Peter R., Rudolf, Robert, Li, Ke, Wilcock, Reuben, Brown, Andrew D. and Harris, Nick, "Fully differential electro-mechanical phase locked loop sensor circuit," Sensors and Actuators A: Physical, Volume 194, 1 May 2013 [3]

- Chapter 46, "Achieving Invariability in Analog Circuits Operating in Extreme Environments" by Peter Wilson, Robert Rudolf and Reuben Wilcock in "Extreme Environment Electronics", edited by John D. Cressler and H. Alan Mantooth [4].

Chapter 2

State of the Art

2.1 Introduction

In this chapter, state of the art techniques to mitigate the impacts of device variation particular to analogue circuits are reviewed. Since this is an area of intensive research and constant innovation, it would be impossible to provide an entirely comprehensive list. Instead, the techniques described in this chapter are included because they are a good example of a commonly used principle, because they are particularly interesting, or because they are directly related to this work.

One key point of this review is focussed on calibration techniques, which allow a circuit's parameters to be adjusted after fabrication to ensure they are within specifications. Calibration techniques lie on a spectrum from purely analogue over digital to software solutions. Analogue calibration techniques are characterised by their ability to adjust the properties of a device such that its fundamental characteristics can be altered. Examples of this category of techniques are substrate biasing, floating gates and some forms of conventional trimming. Digital techniques are characterised in that the fundamental device properties remain unchanged, but circuit characteristics are altered by enabling or disabling a certain number of calibration devices, re-routing signal flow or reconfiguring certain parts of the circuit. Calibration can also take place in software, in which case a system performs a built-in test against a reference and the measured deviations are then used as the basis of a software calibration table.

In contrast to calibration, inherently robust design is another approach to tackle the issue of parameter variation. Design based on calibration accepts the fact that the circuit performance will be affected by parameter variation and introduces calibration schemes to correct them. In contrast, robust design attempts to create a circuit whose parameters remain within specifications even if devices are subject to variation. Robustness in this context is generally considered as a measure for how stable circuit performance

remains when the circuit is subject to parameter variation, and is hence often used synonymous with variation tolerance [17]. The basic concepts of robust design, such as common-centroid transistor layout, are commonly found in most circuits as standard design practice, but these can be extended to make a circuit more resilient against specific device variations. Additionally, there is significant work in the area of model-based robust design, which uses computer algorithms to optimise circuits for robustness.

These two threads, calibration and robust design, are reflected in the contributions made in this thesis. Chapters 3 and 4 are concerned with developing and proving a particular calibration technique, whereas Chapter 5 features an example of how an inherently robust design can be a viable alternative to a very parameter-sensitive design.

Every calibration and inherently robust design technique lies on a continuum between being specific to a particular type of circuit or being generally applicable. These relationships are illustrated in Figure 2.1, which shows the techniques discussed in this chapter. For instance, Switching Sequence Post Adjustment (SSPA) and Calibration DACs are calibration techniques that are only applicable to segmented current-steering Digital-to-Analogue Converters, while conventional trimming is practically universally applicable. Many techniques are somewhere between truly generic and circuit-specific. In most of these cases the fundamental technique or methodology may be quite generic, but it is only applied to certain types of circuits. An example for this are digital calibration devices, which can in principle be widely used, but are most commonly found in mixed-signal circuits, but less frequently in purely analogue circuits.

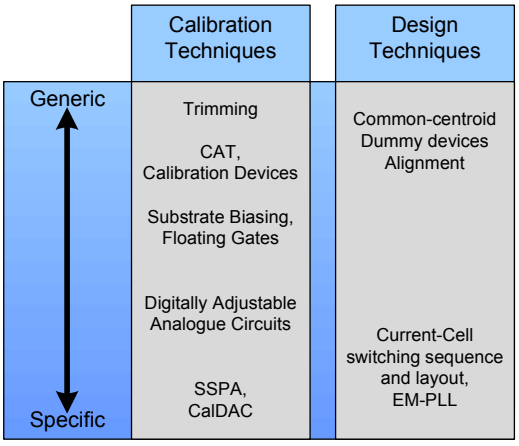


Figure 2.1: Categorisation of design and calibration techniques in this chapter.

The structure of the following sections of this chapter is illustrated in Figure 2.2. First, a selection of important generic calibration techniques are discussed in Section 2.2. Section 2.3 then discusses generic robust design techniques and gives a brief introduction to model-based robust design. Then, Section 2.4 covers current design and calibration techniques specific to current-steering digital-to-analogue converters. Finally, Section

2.5 discusses the CAT technique on which a large part of this work is based, and relates it to other calibration techniques described in this chapter.

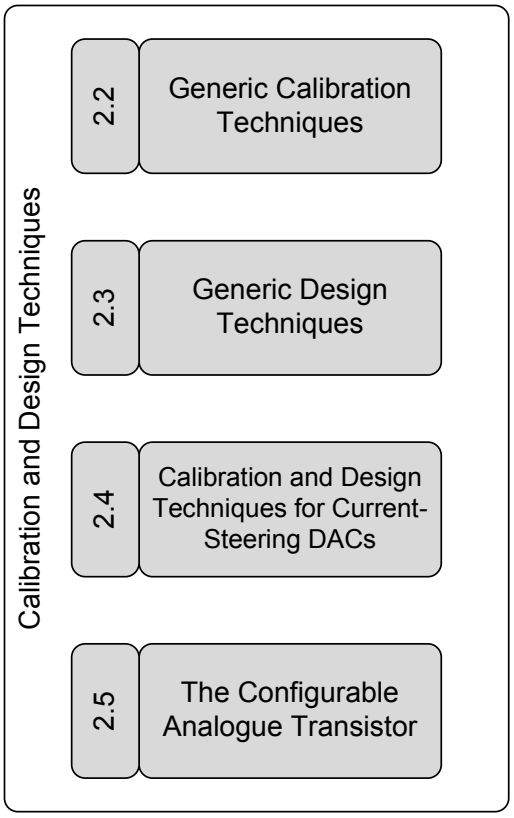


Figure 2.2: Structure of this chapter.

2.2 Calibration Techniques

2.2.1 Introduction

In this section, a number of generally applicable calibration techniques are presented. The defining factor of all of these techniques is that they are applicable to a wide range of different circuits in principle. This generality of use is due to their operation on a physical or device level, rather than the circuit level. For instance, the calibration techniques described in the following subsections alter fundamental device properties, such as resistance, threshold voltage or the channel aspect ratio. These device-level parameters are independent of the circuit in which they are used and can therefore be employed in any circuit where alteration of these parameters result in a changes in circuit performance. Section 2.3 will then continue with an overview of robust circuit design techniques.

2.2.2 Trimming

Component trimming is the earliest and perhaps the most fundamental calibration technique. When applied to a circuit, individual components are manipulated to change their values and thus adjust the operating parameters of a circuit. For integrated circuits, a historically common trimming method is laser trimming, where a laser is used to alter the physical structure of the fabricated chip. Although other devices could be trimmed in principle, laser trimming is most commonly applied to resistors. Early trimming techniques were applied to thick-film (hybrid modules [18]) or thin-film (integrated circuits [19]) resistors on a chip, whose area could be trimmed much like discrete laser-trimmed resistors. However, this arrangement was unfavourable for two reasons. First, a process with resistive material as the uppermost layer is required and second, the area of such trimming resistors is usually large compared to other devices on the chip which is impractical for high-density integrated circuits.

Improved techniques feature a resistor that is not itself trimmed, but rather employs a number of taps, all but one of which are cut with a laser to select a specific resistor value, much a like a potentiometer with discrete wiper settings. This principle is described in [20] and is shown in Figure 2.3 Although metal fuses relax the process requirements, the area requirements of discrete resistors are not reduced significantly.

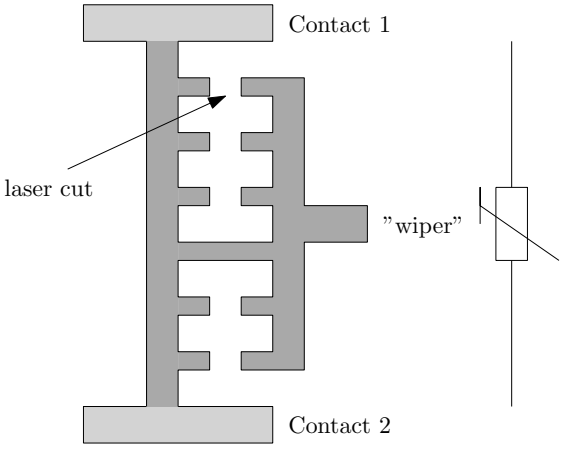


Figure 2.3: Transistor trimming by burning metal links.

Laser trimming has two key drawbacks in particular: First, it requires special machinery to perform the trimming after fabrication, while most other techniques are purely electronic. Second, it is irreversible, or at the very least the resistance can only be trimmed in one direction. The latter, in addition to the fact that physical access to the chip is required to conduct trimming, means that laser trimming is a one-time process in practice that cannot be used to counter the effects of device ageing or to recalibrate the circuit under special environmental conditions. Furthermore, laser trimming is still a comparatively complex and expensive process, which is why it has been largely replaced by other trimming techniques whenever possible.

A more economic and less process-dependent approach to discrete resistor trimming has been used in precision integrated circuits since the 1970's and is commonly referred to as "Zener" zapping, which saw its first widely commercial application in the OP-07 operational amplifier [21]. Zener zapping features a series string of resistors in parallel with Zener diodes, which under normal conditions are reverse biased and will therefore not affect the operation of the resistors. This is illustrated in Figure 2.4, where the resistance is measured between the V+ and V- terminals and the X_n terminals are used to access the internal nodes for trimming. In order to select a particular resistor value, a high current is passed through one or more Zener diodes, which results in local heating sufficient to fuse the diode's metal contacts into the silicon, creating a short in parallel with the corresponding resistor. Hence, the total resistance of the string can be trimmed. A significant advantage of Zener zapping over laser trimming or metal fuse-based processes is that Zener diodes are readily available in bipolar processes as base-emitter junctions and that therefore no special requirements are placed on the fabrication process. The trimming terminals are normally not accessible in the packaged chip and trimming is normally done before packaging.

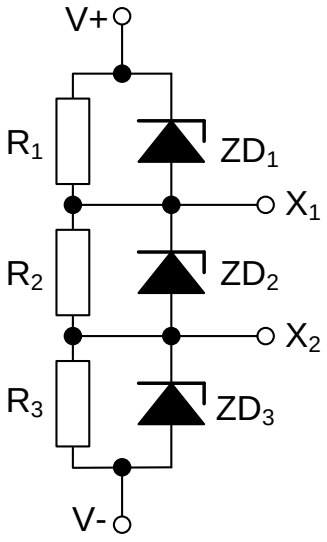
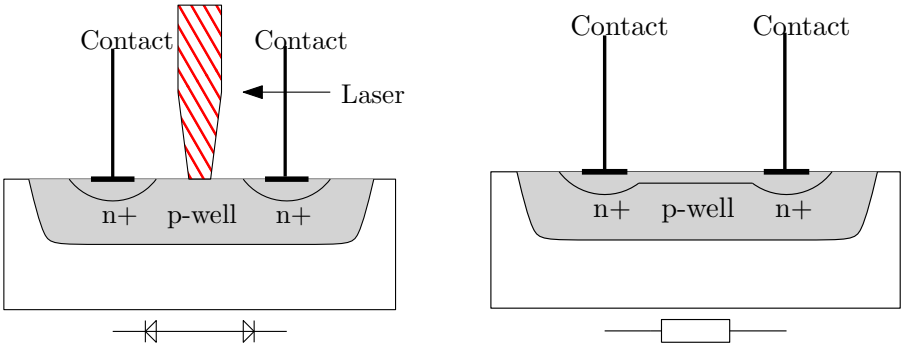


Figure 2.4: Principle of "Zener zapping" resistor trimming.

As analogue circuits began to scale down further, the overhead of on-chip resistors was no longer economically viable and trimming resistors disappeared from all but the most sophisticated precision integrated circuits. However, recently interest emerged again in the field of laser trimming and trimming processes compatible with small-scale CMOS processes were developed. One such device is the laser-diffused resistor described in [22], which is shown in Figure 2.5. It comprises of a standard CMOS device without a gate and works as follows: Under normal circumstances, the n+ diffusions and the p substrate form two back-to-back pn-junctions, which do not pass any current. If a laser beam is focussed on the gap between the two n+ diffusions, the n+ diffusions start to extend



(a) Untreated device. (b) Laser-diffused resistor.

Figure 2.5: Laser-trimmed diffusion resistor

into the gap, thereby forming an electrically conducting link between the electrodes. The duration and intensity of the laser process controls the amount of n+ material that diffuses and thereby controls the resistance between the electrodes.

Another interesting method based on a somewhat similar device is presented in [23], where doped polysilicon is used as the resistor base material. However, instead of laser trimming, the resistance is changed by pulsing a large current through the material, which again results in heating of the material. However, instead of material diffusion, this device utilises dopant redistribution during the cooling process to alter resistance. Importantly, this effect is reversible, which allows resistance to be trimmed multiple times. This particular method has been further developed on SiGe processes [24]

While conventional resistance trimming is no longer relevant to modern highly-integrated devices, it still has a legitimate place in discrete analogue circuits such as precision operational amplifiers. In small-scale CMOS devices, trimming methods based on standard devices provide a viable path for calibration, provided that trimming a resistance can be used to calibrate a circuit parameter.

2.2.3 Substrate Biasing

Contrary to most forms of trimming, which are based on physical manipulation of devices, most other analogue calibration techniques are purely electronic. One example is substrate biasing, which is derived from a method first applied to digital circuits, where the substrate of the entire chip or a part thereof is biased to control gate propagation time delay and leakage power. For analogue circuits, this method can be scaled down to be applied to individual transistors, where it is commonly referred to as Dynamic Threshold MOS (DTMOS), which was introduced in [25].

In a MOS transistor, the threshold voltage is subject to the body effect, which relates the threshold voltage (V_{th}) to the source-bulk voltage (V_{SB}):

$$V_{th} = V_{th0} + \gamma(\sqrt{V_{SB} + 2\phi} - \sqrt{2\phi}) \tag{2.1}$$

where V_{th0} is the threshold voltage at zero bias, γ is the body-effect coefficient and ϕ is the surface potential. Conventionally, the source is tied to the bulk, thereby fixing the threshold voltage. If a bias voltage is applied to the body of a transistor, however, the threshold voltage can be actively adjusted.

Ignoring channel-width modulation, the drain current (I_D) of a MOS transistor in the active region is defined by:

$$I_D = \frac{1}{2}\mu C_{ox}\frac{W}{L}(V_{GS} - V_{th})^2 \tag{2.2}$$

where μ is the carrier mobility, C_{ox} is the gate oxide capacitance per unit area, V_{GS} is the gate-source voltage, and W and L are the channel dimensions. Substituting V_{th} from equation 2.1 into the drain current equation results in:

$$I_D = \frac{1}{2}\mu C_{ox}\frac{W}{L}(V_{GS} - V_{th0} - \gamma(\sqrt{V_{SB} + 2\phi} - \sqrt{2\phi}))^2 \tag{2.3}$$

which shows that for given operating conditions, the drain current can be adjusted by varying the body voltage.

In order to be able to apply body bias voltage to an individual transistor, the CMOS process must support isolation of individual transistors, which necessitates either triple-well or silicon-on-insulator processes. Whilst some specialist processes, such as radio-frequency analogue analogue processes or radiation-hardened processes, may support biasing of individual transistors, this feature is not typically found in modern mainstream general-purpose CMOS processes.

The adjustment of the threshold voltage in this way has two main implications for circuit design. Firstly, as can be seen from equation 2.1, the effective threshold voltage of a transistor can be lowered by applying a constant bias voltage, which allows circuits to operate at lower supply voltages. For example, in [26] an operational amplifier with a nominal single supply voltage of 0.8V is presented. Secondly, by dynamically deriving the body bias voltage from circuit parameters, this technique can also be employed to increase robustness of the circuit against device parameter or bias condition variation. For example, in [27] it was shown that with a relatively simple negative feedback loop, which dynamically controls the body biasing of an NMOS transistor as a common-source amplifier for constant drain current, circuit parameter deviation across the process corners was improved from 77% to 12%. This technique was also applied in the previously mentioned operational amplifier [26], where the bias of the input stage is controlled by a DTMOS transistor.

Due to the aforementioned process requirements, this technique is not applicable to standard analogue CMOS processes. Low-voltage circuits where a reduction of threshold voltage is necessary and therefore already contain DTMOS transistors may benefit greatly from the possibility to add robustness against parameter variation. Since body bias is set by an analogue voltage, a separate DAC would be required to allow for digital calibration, which means that this technique is not well suited for post-silicon calibration of individual transistors. Instead, the body bias voltage is usually derived from an internal node to stabilise the circuit bias [27, 28]. In contrast, digital circuits may be physically clustered and the bias voltage for each cluster generated by a dedicated DAC to optimize a certain parameter, such as propagation delay [29].

2.2.4 Floating Gates

This technique is somewhat related to substrate biasing (Section 2.2.3) in that it affects the input voltage characteristics of a transistor. The main differences are the location at which the device is adjusted and that floating gates provide inherent non-volatile storage of the calibration value. The principle of floating gates was first applied in digital memories such as Electrically Erasable PROMs (EEPROMs). However, the deliberate alteration of MOSFET threshold voltage that is used to store digital data can likewise be applied to trim analogue circuits [30].

A MOSFET containing a floating gate, as depicted in Figure 2.6 is structurally very similar to a conventional MOSEFT, except that is has an additional gate (floating gate, FG) added between the original gate (control gate, CG) and the substrate. The purpose of this additional gate is to store a certain amount of charge that effectively alters the threshold voltage of the device [31].

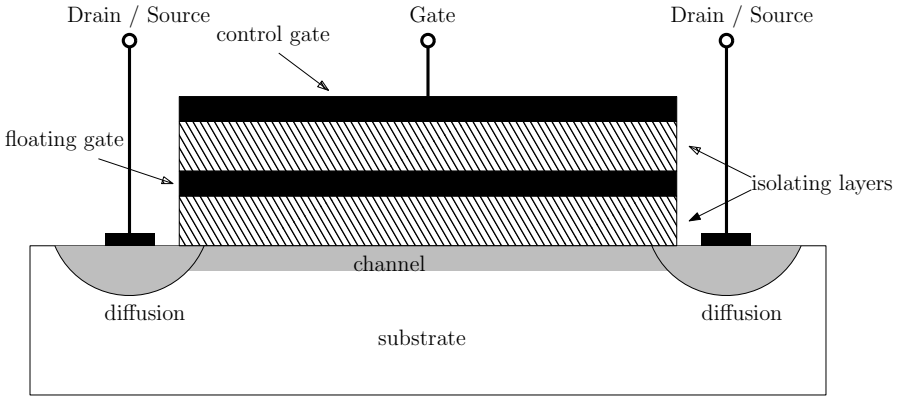


Figure 2.6: Schematic diagram of a floating-gate device.

Since the floating gate is electrically isolated from the control gate and the substrate, charges are added and removed by means of various tunnelling mechanisms. To facilitate this, these devices usually contain a support transistor or other means to enable and

disable carrier injection either through a dedicated terminal or the terminals of the main transistor. The former method is usually preferred since it can be implemented efficiently by sharing the floating gate with a second transistor which is exclusively used for the purpose of charge tunnelling [32], otherwise the transistor would need to be equipped with multiplexers to switch between normal and programming operation.

While the additional floating gate layer is not seen as a drawback in memory chips, it is usually not available on analogue CMOS processes. However, at the expense of greater silicon area, floating gate calibration can be implemented in standard CMOS processes with no special layers. In fact, this method may be preferred to a dedicated floating gate layer, since only transistors that are to be equipped with calibration facility feature a floating gate, while all other devices are standard single-gate CMOS devices.

The principle of the single-layer floating gate transistor is shown in Figure 2.7 and was first introduced for EEPROMs in [33], but later developed to be applied to analogue circuits [34]. The device depicted is an n-channel MOSFET with the n diffusions in the substrate as source and drain. This device's gate is connected to the gate of an adjacent p-channel device and acts as the isolated floating gate. The n-channel device is controlled through the n-well of the p-channel device, whose potential is changed by applying a control ("gate") voltage to the diffusions and implants of the p-channel device. Since with positive gate voltages the junction from the n-well to the p-substrate is reverse biased, this p-channel device is electrically isolated from the n-channel device and controls the n-channel device through the floating gate.

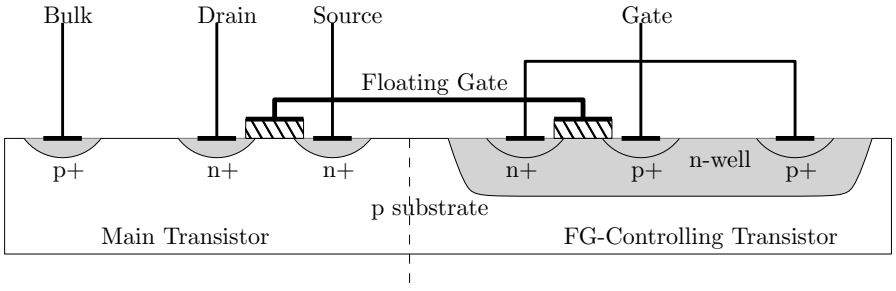


Figure 2.7: Schematic diagram of a floating-gate device on a conventional CMOS process without special layers.

Due to the way the charge on the floating gate affects the effective gate voltage, this technique is commonly used to cancel or compensate offset voltages, for example in differential pairs [35]. As the charge stored on the floating gate is practically non-volatile, floating gates can in principle be used for one-time calibrations, although adaptive adjustment of the charge level during operation is a more commonly described scenario [36]. Conversely, floating gates can also be used to deliberately add offset voltages, which, for example, allows the creation of adjustable current sources where the gate voltage of the current source transistor is controlled by a floating gate.

2.2.5 Calibration Devices

Perhaps the most fundamental way to achieve digital control or calibration of an analogue circuit is to insert calibration devices that complement a main device. These calibration devices can be turned on or off in a digital manner, resulting in digitally adjustable characteristics of the combination of main and calibration devices. This principle is not only used for calibration, but also for programmable gain amplifiers [37] and similar applications. A large number of circuits featuring calibration devices found in the literature are data converters. Therefore, several techniques that employ calibration devices are discussed with other DAC specific techniques in Section 2.4.

An independent application of calibration devices is shown in [38], which features a digitally controlled length MOS transistor. A number of calibration transistors with binary weighted lengths are connected in series to the main transistor. Individual transistors can either be bypassed by connecting their gates to VDD or inserted in the circuit by connecting their gates to the main transistor gate, as depicted in Figure 2.8 [38]. The series transistors can be enabled by a digital control word to effectively give a variable transistor length.

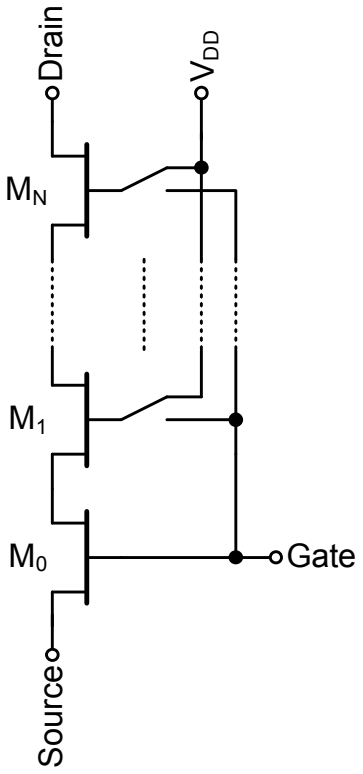


Figure 2.8: MOS transistor with programmable length.

The intended application of this programmable length transistor is calibration and matching of current sources in tunable neuromorphic cells, but the authors present some results geared towards more generic applications. This approach has not been proposed to counter device variability, but it would in principle be applicable for calibration. The binary sizing of the calibration devices results a wider tuning range with the same number of devices when compared to unity-sized calibration devices, e.g. [39]. A drawback of this technique is that the calibration device and its control switch are arranged in series. This can be problematic since disabled devices are still in the signal path and their presence may affect the overall circuit characteristics, e.g. through their finite on-resistance.

2.2.6 Digitally Adjustable Analogue Circuits

The calibration methods discussed thus far operate mostly on the device level, with little or no support circuitry. Even in the more complex cases presented in Section 2.2.5, the support circuitry is only required for calibration of the overall circuit but not to operate the individual calibration devices. In practice, even simple calibration devices will be used with some kind of system to support adjustment of circuit parameters. Such a system where digital devices are controlled with the overall circuit performance in mind can be generalised as a digitally adjustable analogue circuit. In this section, the general aspects and certain details of such digitally controlled circuits will be discussed.

An analogue circuit is equipped with two nodes, a "monitor" and a "control", as shown in Figure 2.9 [40]. The monitor allows the measurement of a certain parameter, while the control adjusts a certain variable. An example monitor parameters is gain, which can be determined by measuring the output voltage of an amplifier for a certain input voltage, while an example control may be the bias voltage by tuning a programmable current source. When the circuit is to be calibrated, either after fabrication, at power-up or periodically during operation, the control variable is varied according to an algorithm until the monitored parameter is within specifications or otherwise optimized. The value at the monitor node does not need to be measured in all cases, as it is often sufficient to simply compare it to a reference value and adjust the control accordingly. Indeed, successive approximation algorithms that rely only on comparison instead of measurement are commonly found in such architectures.

In the ideal case, monitoring and adjusting the control can occur during normal operation of the circuit. However, more commonly normal operation of the circuit may need to be interrupted for several reasons. For example, to adjust gain of a circuit an input signal of known amplitude must be present, which necessitates disconnection of the input signal and connection of a test signal for the duration of calibration. Another case is when the voltage or current at the monitor node would be too low during normal operation or the circuit needs to be reconfigured in a specific way to bring the monitoring node to a certain, easily measurable condition.

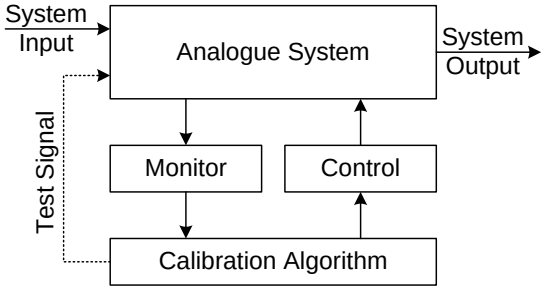


Figure 2.9: Principle of a digitally adjustable analogue circuit.

Depending on the complexity of the calibration system, the calibration cycle itself requires a means to sense the monitor node, such as a comparator or ADC and a DAC or circuit component that serves the purpose of a control. Storage of the calibration word can either be within the calibration circuit itself as volatile memory, fuses or floating gates or in the memory of an on-chip microprocessor system, which may also include the calibration logic, if present. In any case, after the completion of the calibration cycle, normal operation of the circuit is not affected, except for effects introduced by the presence of monitors and controls, which may lead to increased parasitics or altered device output characteristics. A direct example for this are series-connected transistors discussed in Section 2.2.5.

One general property is that the accuracy of the calibration depends on the DAC structure of the control. If the resolution of this DAC is too low or individual bits deviate too greatly, the quality of calibration may degrade or it may not even be possible to achieve satisfactory calibration. On the other hand, introducing too many calibration bits or conservatively designing the DAC from large-area devices may also be infeasible due to the parasitics introduced to the supported circuit.

An interesting solution to this problem is the use of DAC structures with sub-binary bit weights [41]. Whereas bits in a conventional DAC have weights 2^N, the bits in a sub-binary radix DAC have weights r^N, where the radix $r < 2$. This introduces an overlap in the value that is covered by individual bits, i.e. $r^N < \sum_{n=0}^{N-1} r^n$, which adds code redundancy at the expense of number of bits required for a given maximum value. The radix can be chosen to make the DAC tolerant to an expected maximum parameter variation. Note that this technique does not rely on any DAC architecture and can be applied to any calibration technique that employs binary weighted elements.

A more general view of the system depicted in Figure 2.9 is that of built-in self-test (BIST) and calibration facilities. This term is more commonly applied to digital signal processing systems. In addition to the standard ADC, DAC and signal processor, systems equipped with BIST facilities also contain a set of analogue multiplexers that allow to form a test loop on the analogue side, as depicted in Figure 2.10 [42].

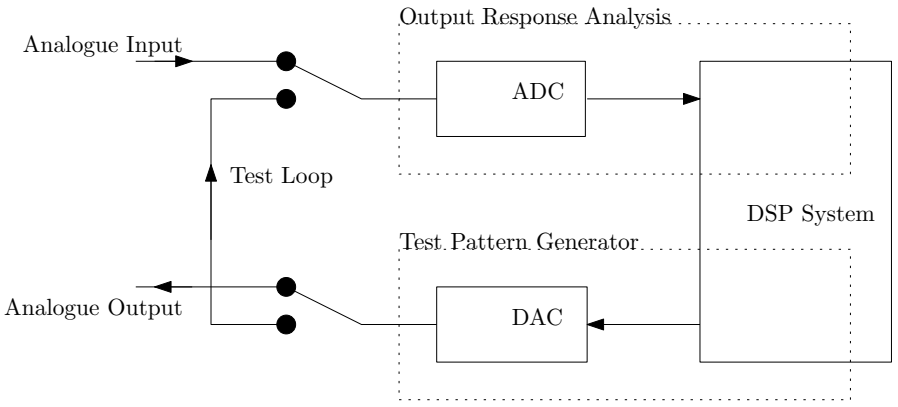


Figure 2.10: DSP system with test loop facility.

To calibrate the system, the test loop is closed, which disconnects the external signals [42]. The DAC is used as a test pattern generator, while the ADC analyses the output response. By applying appropriate test signals and algorithms, the DSP can determine the nonlinearities and errors in the DAC and ADC, which allows a correction pattern to be developed. This information about the characteristics of the system can then be used in normal operation to correct the aforementioned systematic errors. The nature and extent of errors that can be corrected is mainly dependent on hardware properties. While, for example, converter offset can easily be corrected, missing ADC or DAC codes are problematic. Employing $\Sigma\Delta$-modulators, either as a replacement for the system's ADCs and DACs or as dedicated converters for the purpose of BIST, alleviates this particular problem and allows more flexible calibration by adjusting the converter parameters [43, 44].

2.3 Variation-Tolerant Design

2.3.1 Introduction

In contrast to the calibration techniques discussed in section 2.2, robust design is another approach to mitigate the effects of device parameter variation. The conceptual difference is as follows: Calibration accepts that device parameter variation results in circuit performance variation, but provides a way to tune the circuit to bring circuit performance back to acceptable values. Robust design, on the other hand, aims to minimise the effects of device parameter variation on the circuit in the first place. This can be achieved through certain layout strategies, circuit topologies or by ensuring that the circuit is designed with enough margin for error in device parameters.

Section 2.3.2 will examine a number fundamental layout strategies that are commonly used to reduce parameter variation and mismatch in devices. Furthermore, it will also

provide an example of how different circuit topologies can be employed to improve variation tolerance. The common theme across these circuit and layout techniques is that they are largely manual, in that the designer will choose the appropriate circuit topology or layout scheme that is most suitable for a particular application and its requirements. In contrast, Section 2.3.3 will give an overview of computer-aided techniques to optimise the robustness of a circuit.

2.3.2 Robust Design Techniques

The performance of integrated circuits in general not only depends on the circuit topology and device parameters, but to a large part also on the final layout of the chip. Especially for analogue circuits the choice of layout scheme and the particular layout of devices can have a great effect on the performance of the fabricated circuit. Likewise, layout also contributes significantly to how production processes contribute to device parameter variation and how the circuit is affected by variation. The following layout techniques are generally considered fundamental good practice for analogue circuit design, and have been developed since the beginnings of integrated circuit design [45]. Depending on the susceptibility of the circuit on parameter variation, applying some of these layout techniques alone may be sufficient to reduce variation in circuit performance to a level which does no longer require calibration or other compensation.

As discussed in Section 1.2, there are numerous systematic processes which lead to parameter variation and device mismatch on a chip. On the most abstract level, the main sources of such variation are mainly due to physical separation or different orientation and asymmetries in surrounding features of devices which need to have closely matched characteristics. Therefore, the layout techniques of this section are concerned with device and feature placement that minimises these effects. All of these schemes are generally applicable to active and passive devices alike.

If two devices need to be matched, they should have the same orientation. This is because the lithographic and chemical processes used during fabrication can have slightly different properties in different axes. Thus, when two devices are placed in the same orientation, they are less likely to be affected by these differences. [46] Another cause of mismatch are gradients that lead to different properties in two physically separate devices. These can be gradients in underlying physical properties, such as doping, but in the majority of cases temperature gradients across the chip are more significant [47]. The general solution to reduce effects of such gradients in common-centroid design, which is illustrated in Figure 2.11. Conceptually, it means that the two devices are divided into smaller devices, which are then arranged symmetrically and equidistant around a common point, or centroid. This layout scheme will effectively cancel linear gradients in any axis. However, the metal interconnect may become quite complex and asymmetric. This can re-introduce mismatch because the metallisation layers can slightly affect properties

of the devices underneath. A special case of common-centroid layout for devices with a common source or drain terminal that minimises required metal interconnect are interdigitated transistors, illustrated in Figure 2.12. In this case, the two transistors are a differential pair with a common source connection, which allows sharing of the source regions in the layout, as indicated. The example shown is a "true" common centroid layout because it is symmetric in both axes and the centroids of both devices coincide, thus yielding optimal response to parameter gradients. However, this arrangement does in fact not provide perfect matching, since the two devices have different drain and source areas. One variation, shown in Figure 2.13 restores good device matching, but it is no longer a perfect common centroid since the centroids of the two devices are at different locations. Furthermore, this layout requires slightly more complex metal interconnects since the drain region in the centre is not common to both devices and can therefore not be shared. Common-centroid device design therefore requires a trade-off between inherent device matching, matching under the influence of gradients and wiring complexity [46].

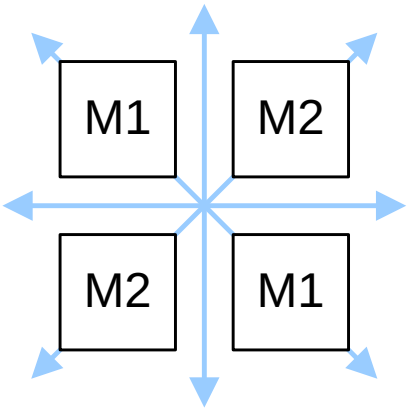


Figure 2.11: Concept of common-centroid layout, arrows indicate axes of symmetry.

In order to avoid the abrupt changes at the end of a strip of interdigitated transistors or a common-centroid array, it is also common practice to insert dummy devices around the edges. These inactive devices have the same physical structure as the active transistors and ensure that all active devices are in a homogeneously structured environment, reducing possible mismatch introduced by density-dependant chemical, mechanical or photolithographic processes. [48]

Another layout technique which is most commonly applied to capacitors is the use of unit-sized devices. This concept is illustrated in Figure 2.14. It can be applied when the absolute values of two or more devices are not variation sensitive, but their relative ratios are. For example, in a charge amplifier, the gain is determined only by the ratio of the feedback capacitors, but not their absolute values. Normally, such different capacitor

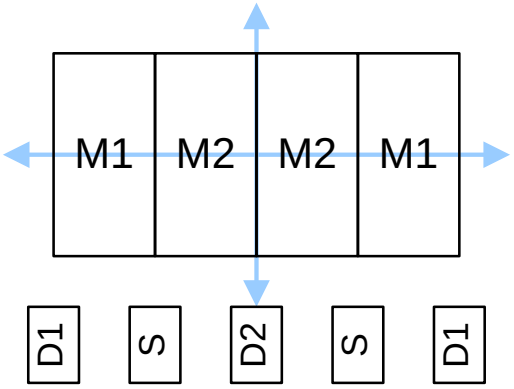


Figure 2.12: Perfect common-centroid arrangement of source-coupled transistor pair.

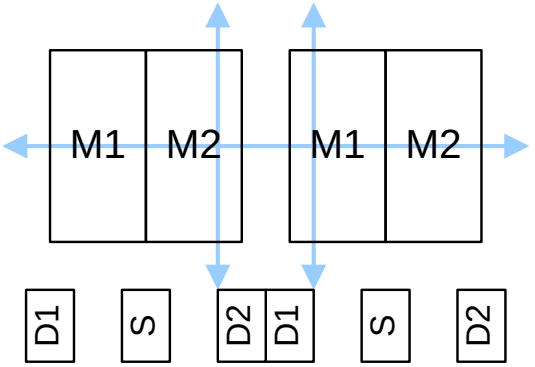
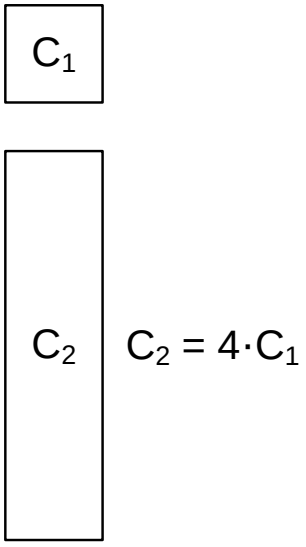


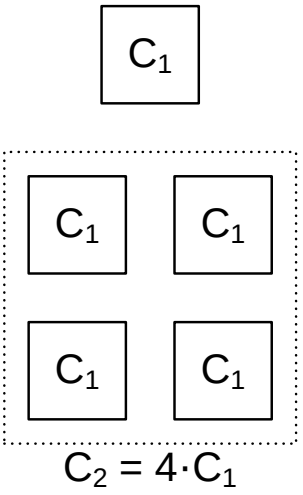
Figure 2.13: Imperfect common-centroid arrangement of source-coupled transistor pair.

values are achieved by having two capacitors with different sizes and possibly different aspect ratios, which is illustrated in Figure 2.14a. Once again imperfections in the manufacturing process can lead to mismatch, which is likely to depend on the physical size and shape of the capacitors. This leads not only to variation in the absolute capacitor value, but also the ratio between the to capacitors. Using unit-sized devices is illustrated in Figure 2.14b. Instead of using capacitors of different sizes, both capacitors are made up of smaller elements of the same size and shape which serve as the unit-sized elements. This way, the unit-sized elements are still subject to variation, but all of them are affected equally. Therefore, the capacitance ratios are defined mainly by the ratio of the number of elements and influence from variation is reduced. The same principle can be applied to other devices where matching of ratios is important, including resistors. [49]

Like the robust layout techniques described thus far in this section, there are also circuit design techniques that result in more robust circuits. Some of them are elementary and used almost without notice, such as various transistor bias stabilisation techniques [50]. Others result in useful "building blocks" that are used whenever a particular function



(a) Capacitors of different sizes.



(b) Capacitors made of unit-sized devices.

Figure 2.14: Application of unit-sized capacitors.

is required [51]. Other still are entirely specific to a particular circuit [52]. Again, the number of such techniques would be too great to give a meaningful comprehensive overview. Instead, the evolution of the bandgap reference will serve as an example for a circuit where temperature stability is achieved through ingenious circuit design.

Among the first semiconductor components suitable for reference voltage generation were Zener diodes. Zener diodes are characterised by a well-defined and relatively abrupt reverse breakdown. When the Zener diode is operated in its breakdown region, the breakdown voltage can be used as a relatively stable reference voltage. Like most semiconductor parameters, the breakdown voltage is subject to temperature drift. In addition to stabilising the operating temperature [53], it is also possible to reduce the temperature coefficient through circuit design. One way to achieve this is illustrated in Figure 2.15 where a regular diode is placed anti-serial with the Zener diode. For Zener diodes with a breakdown voltage greater than approximately 5V, the temperature coefficient of the breakdown voltage is positive. For a regular junction diode, the temperature coefficient of the forward voltage is negative. Thus, if the two are combined as shown in Figure 2.15 and the breakdown and forward voltages added together, the overall temperature coefficient is significantly improved. This scheme only works within a narrow envelope of breakdown voltages, forward currents and temperatures where the two temperature coefficients are approximately equal and opposite. Although discrete temperature compensated Zener diodes are available, this scheme is most commonly found in monolithic integrated circuits, a prime example being the classic LM723 voltage regulator.

Whilst discrete temperature compensated Zener diodes can achieve temperature coefficients in the region of a few ppm/K (e.g. Microsemi 1N4569A [54]), the breakdown

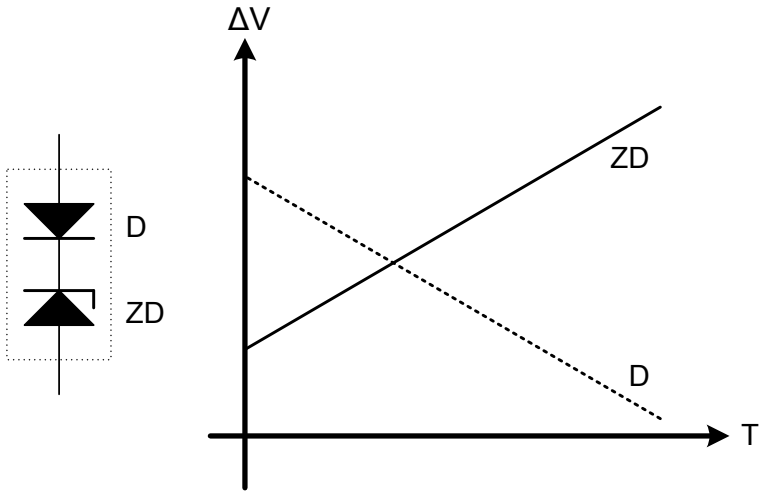


Figure 2.15: Temperature-compensated Zener diode.

voltage where compensation is possible is fixed at around 6V with poor initial accuracy. This, amongst other factors, makes them impractical for use in low-voltage integrated circuits. There, a different reference circuit is commonly used instead: The bandgap reference. The first commercial product featuring an integrated bandgap reference was the LM109 regulator developed by Bob Widlar [55]. Figure 2.16 illustrates the basic bandgap reference circuit used by Widlar. It relies on the fact that two identical transistors have different base-emitter voltages if they operate at different current densities and, crucially, that this difference in voltage has a positive temperature coefficient. Like in the case of the temperature-compensated Zener diode, this voltage can be added to the forward voltage of a junction diode to minimise the overall temperature coefficient. In the circuit of Figure 2.16, transistor Q1 operates at a higher current density than Q2 and thus, the difference in V_{BE} is present at R4, which is amplified to R2 and added to the V_{BE} of Q3, at whose collector is a temperature-compensated reference voltage. The bandgap reference has several advantages over temperature-compensated Zener diodes, most notably the lower output voltage of 1.2V, greater initial accuracy and a hyperbolic temperature coefficient. The latter means that at its designed operating temperature, a bandgap reference has a temperature coefficient of practically zero. Away from that nominal temperature, modern bandgap references with additional compensation can achieve accuracies of below 3ppm/K over their entire operating temperature range without internal temperature control [56]. Naturally, this original design has since been refined several times, resulting in variable bandgap output voltages, enhanced performance, higher-order compensation and FET-only references [57].

Although this example of the development of early bandgap reference circuits is mostly of historical interest, it illustrates how basic physical properties can be exploited solely on a circuit level to achieve higher robustness against parameter variation, in this case temperature.

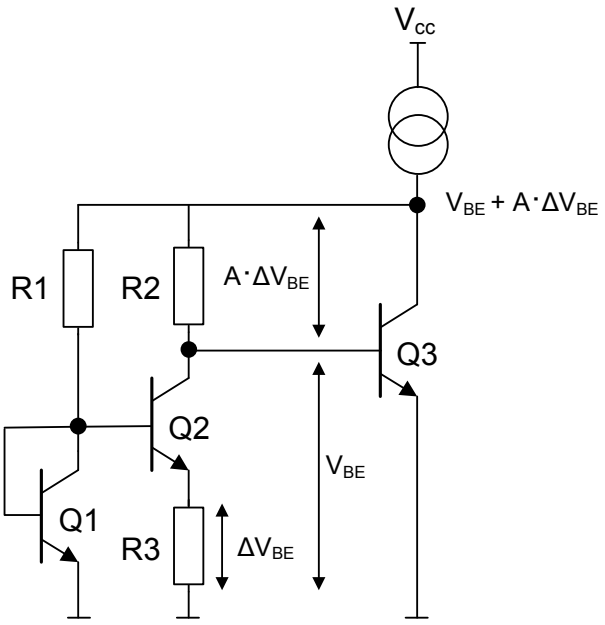
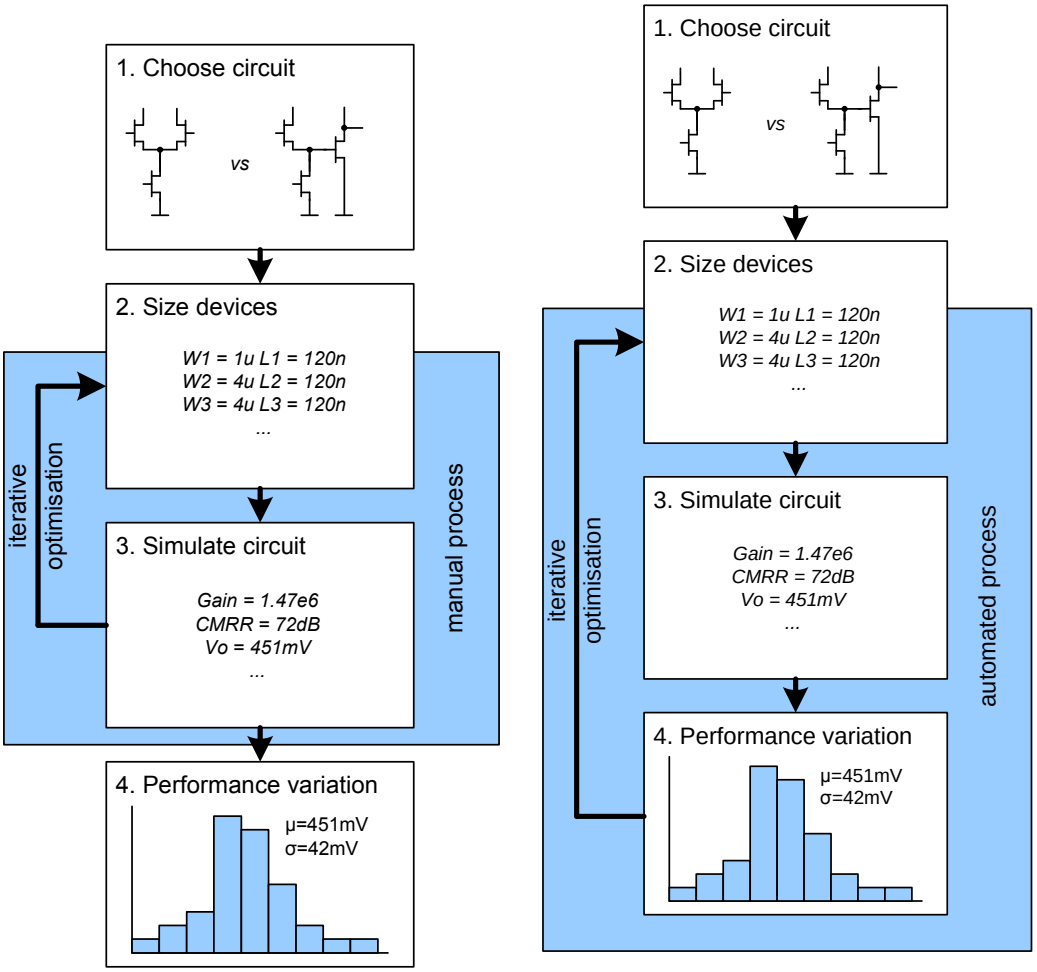


Figure 2.16: Concept of early bipolar bandgap reference [55].

2.3.3 Model-Based Robust Design

Section 2.3.2 introduced the concept of robust design, where a circuit is designed such that the effects of parameter variations on performance are reduced. The examples given are based on optimised circuit topologies or layout features which are inherently more robust. A complementary aspect which has not been discussed thus far is a choice of circuit parameters that lead to a more robust circuit. The importance of this can be illustrated with a hypothetical design example, which is illustrated in Figure 2.17a. First, the designer chooses the appropriate circuit topology and then proceeds to size the devices, based on simplified equations or, more likely, iteratively with the help of a simulator. The goal of sizing is to optimise the desired circuit performances, for example gain or bandwidth of an operational amplifier. However, if no attention is paid to device variability at this stage, it is easy to size a circuit that is extremely susceptible to parameter variation. For example, this can occur when the the operating point of a device is placed close to the boundary to a different operating region, or when an amplifier is designed with very low gain or phase margins. Even when the best practice layout considerations are observed, the fabricated circuit is still likely to have a low production yield, decreased operating temperature range or suffer otherwise from high sensitivity to parameter variation.

It is thus necessary for the designer to consider the impact of the device sizes and other design choices on the robustness of the circuit. A simplistic way to achieve this is to place constraints on the acceptable variation of the circuit performances and ensure they are met when sizing the circuit. The response of circuit performance to parameter



(a) Design flow not considering variation (b) Design flow considering variation

Figure 2.17: Design flow with and without concern for variation.

variation is typically modelled through Monte Carlo (MC) simulation. Conventionally, a set of simulations (e.g. transient, AC, noise, ...) is run with nominal device models to determine the performances of the circuit. In a Monte Carlo simulation, these simulations are repeated a large number of times and in each iteration, the parameters of each device are subject to random variation. The result is a distribution of circuit performances and the designer can then determine if this expected variation of performances is acceptable or not. The numbers for the random variation on device parameters used the the Monte Carlo simulation are typically included in the IC design kit and are based on several measurements of test wafers for a given process.

Using Monte Carlo simulation in the manual design example above allows the designer to verify that a given circuit with given device sizes meets the given robustness criteria. However, it is a rather unwieldy tool in manual iterative design and is therefore likely only used at the end of the design and sizing process as a means of verification instead of continuously driving the design towards a robust solution. This may lead to a final

circuit which may not be optimal in terms of achievable robustness. Using a computer-aided sizing approach can better incorporate robustness as a design goal and at the very least remove the tedious and error-prone manual iterative sizing process.

2.3.3.1 Computer-Aided Design Flow

Creating IC design flows with various degrees of automation is an area of active ongoing research and has been since the early 1980's. An extensive survey of such techniques along with fundamental concepts is presented in [58]. Since the aim of this section is to introduce the concept of computer-aided robust design, it will focus on the underlying concepts and ideas and their application, rather than discussing the literally hundreds of tools.

The simplified IC design flow shown in Figure 2.17a can be readily extended to include automatic optimal device sizing. As before, the designer still chooses the appropriate circuit topology and an initial "guess" for the device sizes. However, instead of manually attempting to optimise the device sizes, this process is now automated. As in the manual case, circuit simulations with nominal device parameters are used to obtain the circuit performance for a given set of design variables. A multi-objective solver is used to find an optimal set of design variables that results in the desired circuit performances. The following Subsections 2.3.3.2 and 2.3.3.3 will discuss the selection of the solver and how robustness can be incorporated as an objective in more detail.

Before moving on to these more detailed discussions, it is worth discussing some more generic points about the design process outlined in Figure 2.17a. First, in this particular example, the choice of circuit and initial sizing are done manually by the designer and optimisation left to the software tools. Whilst seemingly intuitive, there is research which aims to provide tools that automate a larger part of the process [59, 60]. For instance, a database of building blocks can be used to provide a circuit topology suitable for the application based on high-level parameters, e.g. gain and input and output impedances for an amplifier. Similarly, the initial guess for the device sizes can also be provided by a computer, for instance through algebraic equations for a particular circuit topology. A second, perhaps more subtle point about the design flow is the importance of the initial guess of the device sizes. Since it serves as the starting point for optimisation, it needs to be reasonably realistic in order for the optimiser to lead to reasonable results. In particular, care must be taken that with the initial guess, all devices operate in the desired operating regions. If this is not the case, depending on the behaviour of the circuit, the optimiser may not be able to move the operating points into the right region and therefore not find an acceptable solution. Stochastic optimisers, as will be discussed in Section 2.3.3.3, greatly reduce this requirement as they do not explore the design space on a deterministic trajectory.

2.3.3.2 Optimising for Robustness

When employing the automated variant of the circuit sizing approach outlined in Figure 2.17a, it is possible to add circuit robustness as one of the optimisation goals. There are several methods to achieve this, but two broad areas will be discussed in more detail as they are relevant in later part of this thesis: Direct optimisation and post-design filtering.

As stated at the beginning of Section 2.3.3, the typical approach to determine the robustness of a circuit against device parameter variation is Monte Carlo analysis. Monte Carlo analysis is the method that can provide the most complete picture of expected performance variation, provided that the number of iterations is sufficiently large. This is incidentally also the most significant drawback of Monte Carlo analysis, since the resultant long simulation time requires a careful choice of when to apply it. Direct optimisation of circuit robustness, illustrated in Figure 2.17b, would run a full Monte Carlo analysis at each step of the optimisation process and use the resultant sensitivity of circuit performances as objectives to be minimised. Whilst this would be desirable because it would truly optimise circuit robustness, the long simulation times would make it impractical for all but the most simple circuits. It should be mentioned that there are other variation analysis techniques apart from Monte Carlo, such as corner analysis or device-based worst-case analysis. One common problem with all of these techniques when applied to analogue circuits is that the interaction between parameters of different devices is non-trivial. For instance, whilst a corner analysis is usually adequate to determine worst-case variation of propagation delay of a digital circuit, it is nowhere near sufficient for determining the worst-case gain variation that an amplifier might suffer.

Unlike direct optimisation, which considers variation during the sizing process, post-design filtering defers optimisation for robustness. It is applicable to optimisation methods that work on populations of solutions, in particular Genetic Algorithms, which will be discussed in Section 2.3.3.3. For the moment it is sufficient to understand that instead of optimising a single pair of design variables and objectives, genetic algorithms survey a large and varied population of points in the design space to find a global optimum. This also means that once the stopping condition has been reached, depending on the problem, there is not a single solution but a number of different candidate solutions. This set of solutions may contain clusters of very similar solutions, corresponding to local minima in the optimisation objective. Normally only the solution with the lowest objective value would be chosen as the final solution. However, this may no longer be an optimal solution once variation has been considered. Figure 2.18 illustrates how filtering fits into the design process. Since many of the solutions correspond to the same minimum, they can be eliminated so that only one parameter set representative for each minimum remains. Then, a Monte Carlo analysis is performed for each of these points, adding a measure for the robustness of the corresponding parameter set. This allows to trade off robustness against circuit performance and the designer can choose a solution that satisfies the

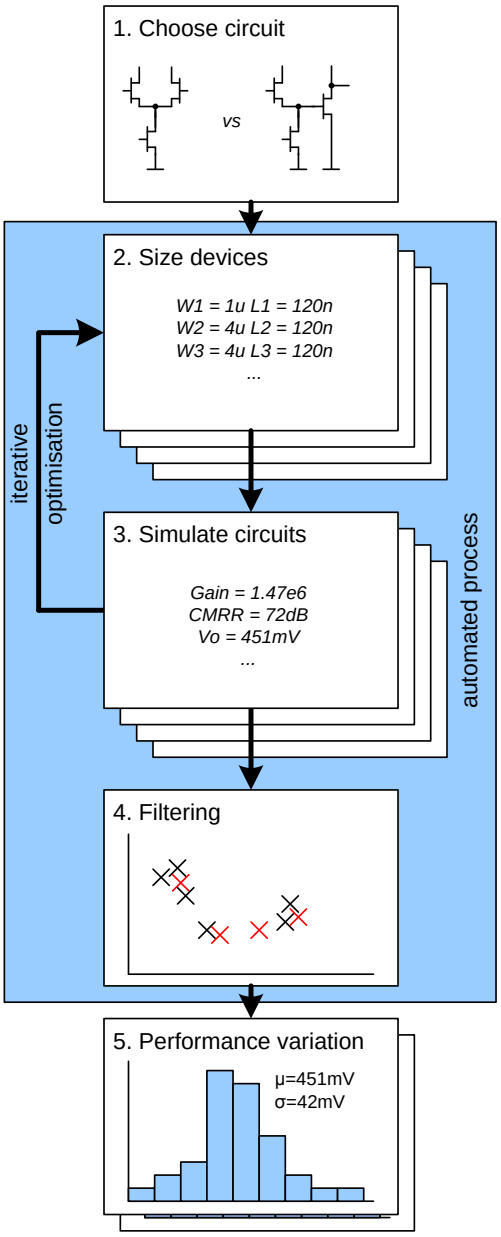


Figure 2.18: Design flow using post-design filtering.

specifications. Post-design filtering provides a viable alternative to direct optimisation, but it does not come without its own drawbacks. First, the genetic algorithm itself may be quite expensive in terms of computing time, requiring tens or hundreds of thousands of simulations, although likely less than direct optimisation. Second, it relies on the optimisation problem having a range of parameters that lead to similar performances so that the most robust one can be chosen. This may not necessarily be the case for every circuit and it may not be obvious at the beginning whether or not a given circuit shows the desired behaviour.

2.3.3.3 Optimisation Techniques

Section 2.3.3.2 discussed fundamental principles computer aided circuit sizing techniques, but did not consider the available choices of optimisation tools. Because circuit parameter optimisation is an important part of the following chapters, this section will give an overview of common techniques. It is important to note that the methods described in this section are not limited or specialised to circuit optimisation, but are generic numeric optimisation methods. Since optimisation is a very broad field of mathematics, the focus of this section is again on the principles and different approaches rather than detailed descriptions of countless methods.

There are several ways to classify optimisation methods. From an user's point of view, the following are some useful criteria of classification:

- Single- or multi-objective
- Single- or multi-variate
- Deterministic or stochastic

Single-objective optimisers are the most fundamental. Their objective is to find the arguments of a scalar function, e.g. $f(x)$ or $f(x, y, z)$ which minimise that particular function. The first example, $f(x)$, is also an example of a single-variate problem, whilst the second example is multi-variate, since there are three independent arguments to the objective function. Single-objective problems are undoubtedly the most fundamental ones and had the earliest solutions developed, e.g. Newton's Method [61].

Multi-objective problems are characterised by more than one objective, which must be minimised simultaneously, e.g. to find a set of arguments x, y, z, which minimises both $f(x, y, z)$ and $g(x, y, z)$. The important difference to note here is that unlike in the case of single-objective problems, determining whether a solution is optimal is not necessarily trivial. In the single-objective case an optimal solution is the one that results in the smallest function value, which is an objective and unambiguous criterion. In contrast, in many multi-objective problems the two objectives will have minima at different points, which means that the final solution will have to be a "trade-off" between the individual objectives. Depending on the characteristics of the problem, there are different approaches to determine how this trade-off should be handled.

One way is to transform the multi-objective problem into a single-objective problem by defining a combined objective. An excellent example for this would a multi-objective data fitting problem, where the sum of squared errors from each objective function would be an appropriate single objective. By doing this, the problem is no longer a multi-objective problem and all tools and properties that apply to single-objective problems can be applied. However, not all multi-objective problems can be reduced in this fashion. For

problems which require their objectives to be treated independently, Pareto optimality is an often-used metric [62].

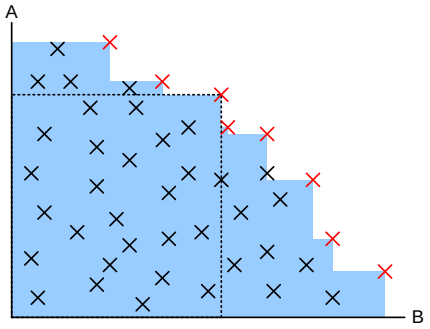


Figure 2.19: Illustration of a Pareto Front

The idea behind Pareto optimality is illustrated in Figure 2.19. The x- and y-axes are two independent objectives, A and B, and each point represents a possible solution. In this example, larger objective values correspond to a better solution. The points were obtained by surveying the design space, for instance through Monte Carlo simulation or a population generation mechanism in genetic algorithms. By considering optimality for each objective separately, a Pareto front can be established. This is the set of solutions for which no other solutions are objectively better, i.e. each solution in the Pareto front is better than at least one other solution in terms of at least one objective. This can be visualised as "shading", where solutions that are not part of the Pareto front are shaded by the ones that are. This is illustrated with dotted lines for one of the solutions. Each other solution that is within the dotted rectangle is worse in terms of at least one objective. The shaded area in the graph is the union of all such rectangles, and any points not within this area are Pareto optimal. This means that they can be considered equally good, and it is up to the person interpreting the result to find an appropriate trade-off between the two objectives. It must be noted here that Paerto optimality provides a method to establish an optimal set, but not a single optimal solution. This must be done by imposing further constraints that help choose a solution from the Pareto set. Often the entire Pareto set is used in population-based optimisation systems, e.g. genetic algorithms. In that case, the Pareto set of a population serves as the seed for the population in the next generation, thus taking advantage of the set of optimal solutions.

The final distinction criterion for optimisers is whether they are deterministic or stochastic. Deterministic methods cover virtually all classic methods and are characterised in that their solution and the path taken to find the solution are determined only by the problem, starting conditions and parameters. Stochastic methods, on the other hand, add random elements to the search for an optimum. Such random elements can be variations around the current point or probabilistic decisions along the path to the solution.

Deterministic solvers are well-suited for optimisation problems where the problem is well understood by the person applying the algorithm and choosing its parameters. Depending on the problem and parameters, many deterministic solvers are capable of determining solutions that are as close to the "true" mathematical solution as is numerically possible. However, if the problem is not well understood or an unsuitable algorithm chosen, it is possible to not find an optimal solution at all, e.g. because the solver repeatedly solves for local minima.

Stochastic solvers, on the other hand, may not be able to find the optimal solution to within the same degree as deterministic solvers, but are less likely to settle for local minima. The two most well-known stochastic solving algorithms are genetic algorithms [63] and simulated annealing [64]. The basic principles of genetic algorithms were outlined already earlier in this section. The non-deterministic element is the variation and evolution of one population to the next, with the intention of covering a large area of feasible solutions. Simulated annealing follows a different route, where a deterministic solver is re-set to nearby starting points according to a time-dependent, probabilistic function with the aim of avoiding local minima.

Whilst this overview of optimisation techniques may have been very broad, it is important to realise that they are all equally valid tools in automated and computer-aided circuit design: From simple single-objective Newton-Ralphson solvers in circuit simulators such as SPICE [65] to genetic algorithms and simulated annealing in automated circuit sizing systems [66, 67]. This therefore forms the basis for understanding the methods described in the subsequent chapters of this work. The remainder of this chapter will now go further from general calibration and design techniques and focus on particular techniques that the original work in this thesis is based upon.

2.4 Design and Calibration Techniques for DACs

2.4.1 Introduction

Having discussed a number of general design and calibration techniques in the previous section, this section will cover some circuit-specific techniques. In particular, the focus will be on techniques targeting segmented current-steering DACs. The operation of such DACs is explained in detail in Chapter 4, but for the purpose of this section it is sufficient to know that one of the key challenges in their design is to provide a series of differential current sources with well-matched, equal currents. These current sources are driven by a thermometer decoder to produce an output current which is ideally a linear function of the input code.

Among the most important performance metrics of digital-to-analogue converters are integral and differential non-linearity, which are abbreviated INL and DNL, respectively.

INL is the deviation of the DAC transfer function from its linear ideal, while DNL is the deviation from the ideal step size between two consecutive codes. In segmented DACs, the main source of INL is mismatch in the unary current cell array, while DNL is mainly caused by mismatch within the binary cells or mismatch between the unary and binary cells.

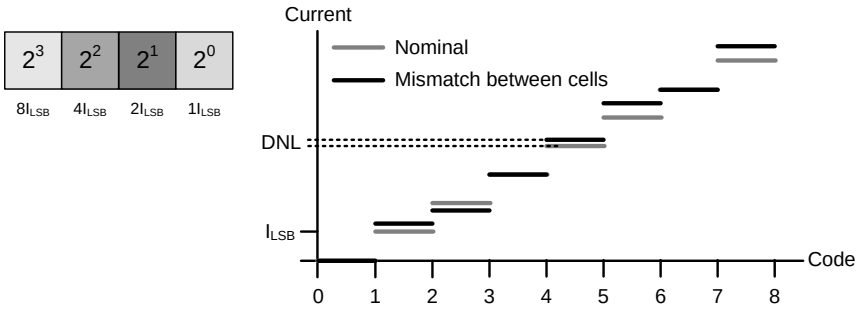


Figure 2.20: DNL caused by mismatch in binary current cells.

Figure 2.20 illustrates the origin of DNL, by considering the binary current cell array and the output current for the first few codes only. If the binary cells were perfectly matched, incrementing the input code would result in a step of exactly I_{LSB} in the output as the input code increments. In reality, the binary current cells shown to the left all have slightly mismatched currents, so that each cell's current is not exactly twice that of its lower neighbour, illustrated by shading in the binary cells. The resultant output steps between codes are then not exactly I_{LSB} and this deviation differs between different codes, depending on which cells are active. This deviation in each step in output current from I_{LSB} is DNL. Since in this example the binary segment of the DAC consists of four bits, the exact DNL pattern will repeat every 16 input codes. Furthermore, mismatch between the binary and unary cells also contributes to DNL. At each 16^{th} code, the output current from the binary segment resets to 0 and the next unary cell is activated, ideally resulting in a step of I_{LSB}. However, if the unary cell current is not exactly 16 I_{LSB}, this step may be smaller or larger and therefore contribute to DNL.

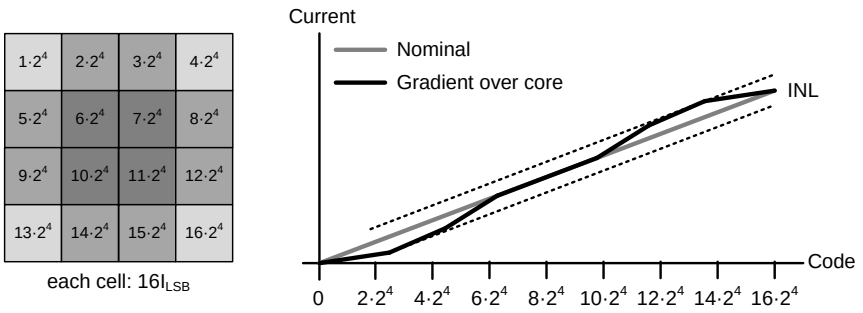


Figure 2.21: INL caused by mismatch in unary current cells.

Figure 2.21 illustrates the origin on INL. The illustration on the left shows the unary current cell array, complementing the binary array from Figure 2.20. Ideally, the current of each cell would be $16I_{LSB}$, and one cell would be activated for every 16^{th} input code, resulting in the linear transfer function shown at the right. In reality, the currents of the cells in the array will suffer from mismatch, as illustrated by the differences in shading. This mismatch can either be due to intrinsic variation or a gradient across the array. In the illustration of Figure 2.21, the cells at the centre of the array provide a higher current than the ones around the perimeter. This results in a transfer function that deviates significantly from the linear ideal at certain codes, as shown at the right. At any point, this deviation from linear is INL, and the maximum or peak-to-peak INL over the entire range is a typical "headline" performance metric.

In terms of overall performance, INL is usually a greater concern than DNL, mainly for two reasons. First, in a segmented DAC, integral non-linearity accumulates all absolute errors over the current cell array. On the other hand, differential non-linearity is defined by the matching of consecutive code outputs and does therefore not accumulate. The second reason follows as a direct consequence: INL can become large very easily, while the DNL is effectively limited. Therefore, the limiting factor for accuracy is usually INL, and the overall performance of a DAC can be more readily improved by improving accuracy through a lowered INL than by improving resolution through a lowered DNL. Typical values of INL for commercial DACs range from fractions of LSB up to a few LSB, whilst DNL is typically of the order of 1LSB.

In addition to mismatch between current cells, there are also other factors that contribute to INL and DNL, such as nonlinearities of the output amplifier. However, mismatch is by far the most important and has been the focus of significant research which has led to numerous different solutions. These can be roughly divided into two categories: Improvements to design techniques and architecture, where good matching is achieved by design, and calibration, where there is some kind of calibration facility built into the DAC.

2.4.2 General Design Techniques

One common technique to improve matching of current sources across an array is to split each source into four parts and locate them in different quadrants of the array, as illustrated in Figure 2.22. If the current sources are then activated through a thermometer code, the errors in current generated by each part of a current source are supposed to cancel out. This method works well if the errors in current are linear over the array. However, in practice the errors across an array are not necessarily linear [68]. This idea is in principle the same as common-centroid transistor layout to improve matching between two or more transistors, as described in Section 2.3.2. In theory, this technique would not result in a great increase in chip area because each current source transistor

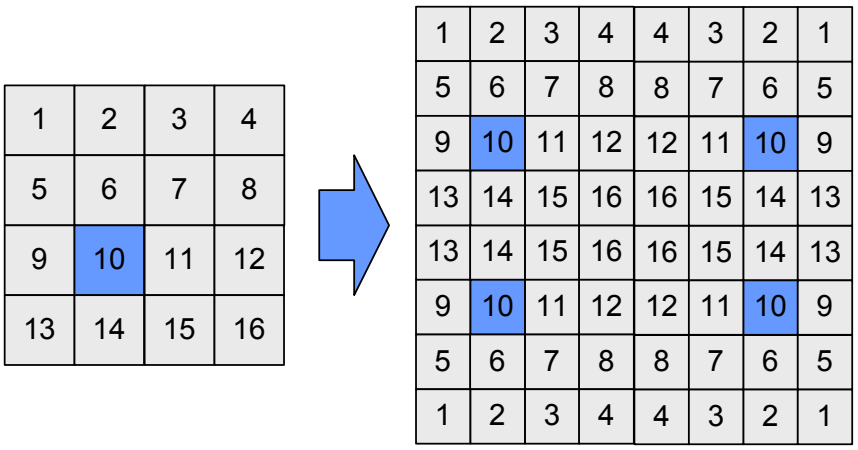


Figure 2.22: Distributing current cells across the array.

will only have to be one quarter of its original size. However, decreasing the current source transistor by such an amount can degrade device-to-device matching, which may again contribute to an increase in INL. This segmentation also necessitates additional wiring between the individual quadrants. Therefore, in practical implementations this technique will incur an area penalty.

Another way to improve the INL of the current cell array is to modify the switching sequence. Instead of using a simple thermometer sequence, the current cells can be activated in a sequence that reduces the effects of parameter gradients. This can be achieved by selecting the switching sequence such that the consecutive cells' physical locations will cause the errors in current to cancel out. For example, it was shown that by using a random switching sequence, where consecutive current cells are distributed across the array, the effects of both linear and non-linear errors in the current source array can be reduced [69]. While for this approach the current cells need not be modified, the cell select logic can potentially become very complex. If the current cell matrix and the switches are physically separate, the switches can still be controlled through thermometer decoders and the switching sequence can be implemented in the wiring. Modified switching sequences can also be combined with distributed current cells, in which case the modified switching sequence applies to each quadrant and is then replicated across.

2.4.3 Calibration DACs

Techniques that calibrate the current cells are very diverse, and only a representative selection will be discussed here. In fact, a number of the generic calibration methods discussed in Section 2.2 were developed for DACs and can therefore be applied to current cells of current-steering DACs. An additional area of techniques concentrates on adding what is conventionally termed Calibration DACs (CalDAC) to the current cell

matrix [70]. These calibration DACs add a controllable current to the output of the primary DAC, thus facilitating calibration. The input word for the calibration DAC can be derived from the primary DAC input and calibration memory [71]. However, this is typically not suitable for high-speed converters since the calibration DAC and its control logic must be able to run at the full conversion rate. Another use of calibration DACs which is more commonly found in high-speed DACs is to calibrate all or some of the primary current cells. A typical implementation of a current cell with calibration capability consists of a primary current source and several parallel current sources that can be enabled digitally. Examples exist in the literature where either each current calibration source is switched at its gate [72] or where each calibration source has its own cascode transistor which is then switched [73]. In effect, CalDACs are similar to Configurable Analogue Transistors (Section 2.5) when applied to DACs, as is done in Chapter 4.

A further very similar technique is presented in [39], which propose to use minimum-sized transistors as calibration devices in current cells of current-steering DACs. A current cell employing this technique is depicted in Figure 2.23 [39]. It consists of a current mirror M_r and M_0, which would be the only component in a conventional current cell. In addition, it also contains a number of minimum-sized transistors in parallel to M_0, which can be enabled or disabled by a series pass transistor.

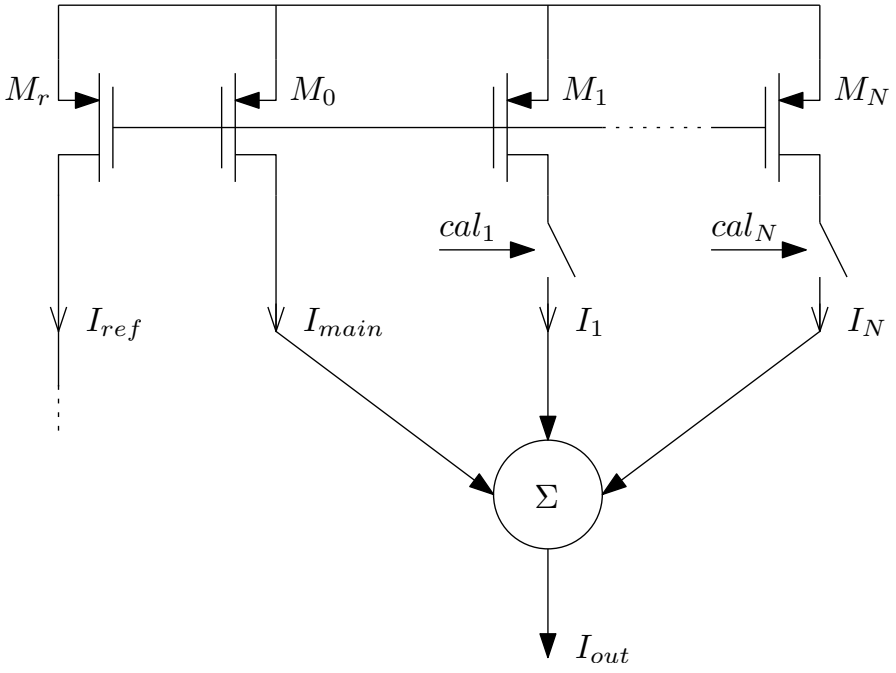


Figure 2.23: Schematic diagram of a calibratable current cell.

Such a current cell is calibrated by feeding the output current I_{out} to a current comparator and comparing it to an upper and lower accepted reference current. The calibration transistors are then enabled in sequence until the output current is within the defined

limits. To store the configuration word, the authors propose a one-bit SRAM memory cell for every transistor and an automated calibration process at power-up.

In [74], the authors present a DAC equipped with this calibration technique, where each current cell consists of the main transistor M_0 and 9 calibration stages. Overall, the output current of the main transistor has a standard deviation of 1.5% and each calibration current has a standard deviation of 30%. After the calibration process, the standard deviation of the total output current is improved to 0.018% and the 12-bit DAC achieves an INL of $\pm 0.25LSB$ and a DNL of $\pm 0.3LSB$.

The designator CalDAC is also applied to techniques where an additional calibration DAC does not provide a calibration current that is added to the main current, but instead biases the body of the current source transistors [75]. Such a scheme employs the principle of Section 2.2.3 to match all of the current sources within the unary array. The advantage over CalDACs that manipulate the output current directly is that the CalDAC in this case only needs to provide a static value during normal operation. However, it does require one CalDAC for each current source and also requires the capability to bias the body of the current source transistors.

In addition, trimming techniques discussed in Section 2.2.2 that have also been applied to DACs are floating gates and laser-diffused resistor trimming. In [76, 77], the authors demonstrate how floating gate trimming can be applied to trim current-steering DACs and report a post-calibration INL of $\pm 0.3LSB$ and DNL of $\pm 0.4LSB$. [78] presents a DAC that is equipped with diffusion resistor trimming, which achieves an INL of $\pm 0.7LSB$ and a DNL of $\pm 0.2LSB$ after trimming with an area overhead for the trimming circuit of 4.4%.

2.4.4 Switching Sequence Post Adjustment

In addition to static optimisation of the switching sequence at design-time, it is also possible to make the switching sequence dynamically configurable. Again, there are several implementations of this concept, but one particularly interesting technique is Switching Sequence Post Adjustment (SSPA), presented in [79]. In this case, the current sources are left uncalibrated and a configurable switching matrix is used to yield a switching sequence that reduces the effect of errors on converter performance.

In SSPA, the switching sequence of the DAC is rearranged in a two-step process, which is illustrated in Figure 2.24. First, all current sources are sorted by current and paired to match up current sources with the largest positive and negative deviations. Then, a similar algorithm is used to sort the pairs, interleaving pairs with positive and negative deviation. The final sequence of unary current sources after this resequencing significantly improves INL compared to the uncalibrated circuit. In addition to the required number of current sources, such DACs may also contain a number of spare current

sources, which can be used to substitute for a certain number of faulty current sources. Furthermore, the array of current sources will be enclosed by a guard ring of unused current sources, which ensures that the active current sources are in an area of homogeneous layout to improve basic matching, which is an application of the basic layout techniques discussed in Section 2.3.

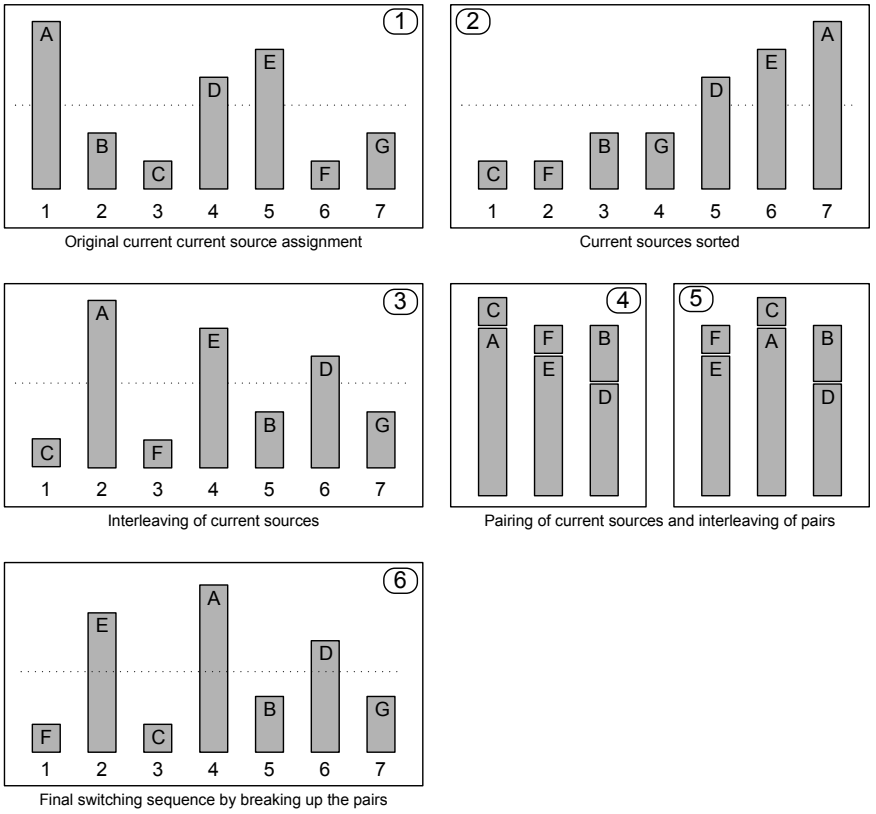


Figure 2.24: Operating principle of switching sequence post adjustment.

Compared to direct calibration of current sources, this approach only requires fairly modest analogue support circuitry, which mainly consists of a current comparator, and the digital switching matrix. The latter, although easily implemented with digital transmission gates, may introduce a significant silicon area overhead. The authors of [79] report a reduction of DAC INL by a factor of 2.2, where the calibration controller consumes approximately 58% of the chip area and the remainder equally divided between the current source array and the switching matrix.

The techniques presented in this section will help to see the application of the Configurable Analogue Transistor to a DAC in chapter 4 in the context of current research. This chapter will now finish with a detailed review of the existing concepts and methodologies for the Configurable Analogue Transistor.

2.5 The Configurable Analogue Transistor

2.5.1 Introduction

The Configurable Analogue Transistor (CAT), first proposed by Wilson and Wilcock [80], is another calibration technique, as introduced in Figure 2.1. Whilst it is on a device level similar to some of the calibration techniques discussed already, it is unique in that it is envisaged to be supported by a suite of software tools and algorithms and aims for general applicability. The aim of these tools is to help the designer integrate CAT into their circuits, and optimise the CAT device sizes and calibration values with ideally very little manual input. Since a significant portion of the original work in this thesis is based on the CAT, it is discussed in this section in more detail than the calibration techniques in the previous sections. Some of the figures used in this section are adapted from a worksheet created by Dr Peter Wilson and Dr Reuben Wilcock.

2.5.2 Operating Principle

In order to improve parameter variation of a single device, the CAT device employs a number of calibration transistors in parallel with a main transistor to allow post-silicon calibration of the overall transistor width. Although these calibration slices themselves are subject to normal device variability, the variability of the overall device can be significantly improved, as will be described below.

On the hardware side, the CAT consists of a main transistor in parallel with a number of smaller transistors, as depicted in Figure 2.25 [80]. The gates of the calibration transistors are either grounded to disable them, or connected to the gate of the main transistor, thus resulting in a parallel combination of main and calibration transistors. This gives in effect a transistor which can be adjusted in width within a certain range with a certain resolution to compensate for variation. In addition to the physical device, the CAT technique also relies on a suite of software tools that accompany the entire design flow to optimize performance improvement. The typical IC design flow when using a CAT is illustrated in figure 2.26 [81].

The task of the first tool is critical device identification (CDI), which determines the devices in a circuit that should be swapped for CATs in order to maximise yield or to minimize the variation in a set of circuit parameters. This step is vitally important, as especially in large or unintuitive circuits the ideal transistors for adjusting circuit performance may not be obvious. Furthermore, the devices whose parameter variation contributes most to the circuit's performance variation may not be the ideal adjustment devices. Therefore, the development of a reliable automated algorithm for Critical Device Identification is one of the cornerstones of the CAT technique. In Section 3.2, an algorithm for Critical Device Identification is presented.

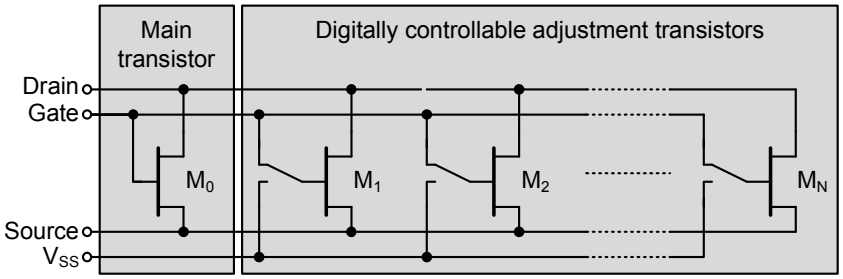


Figure 2.25: Schematic of the CAT device.

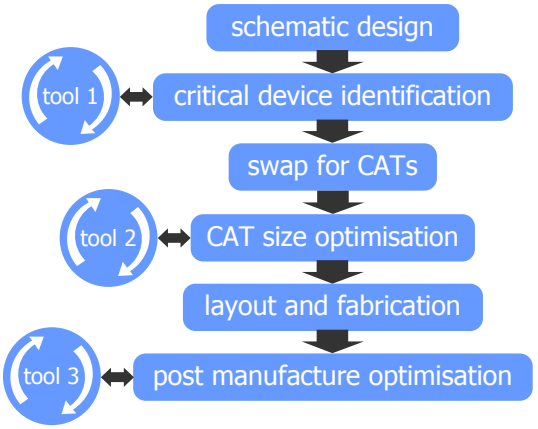


Figure 2.26: Design flow when CAT is employed. [81]

Once transistors have been chosen to be replaced by CATs, another tool is used to optimise the calibration device sizes. There is an established [81] algorithm for optimised sizing, but in its original form it can only be used for constant-current sources. Section 3.3 describes how this algorithm can be employed to optimise device sizes in a wide range of circuits. The constraints that exist in device size optimisation are the desired resolution and calibration range, which is a trade-off between attainable yield and area overhead. At any point after fabrication the calibration words for the devices of any given chip must be determined, which is the task of the third tool. This may either be a one-time operation immediately after fabrication or occur periodically during normal operation in more complex SoCs that can provide the appropriate auxiliary hardware. The algorithm or approach that is used for calibration depends highly on the circuit. In an amplifier, for example, a single CAT may adjust the gain in a fairly linear manner, so a fast and simple successive approximation may be suitable to determine the optimal configuration. However, as circuits become more complex and the number of CATs increases, multi-objective optimisation algorithms will be required. For example, for the DAC in Chapter 4, the post-fabrication optimisation tool is based on a mathematical model of INL and a fast Levenberg-Marquardt least-squares fitting algorithm, which is

described in Section 4.5.3.

2.5.3 Improvement of Current Distribution

Having introduced the CAT in the previous section, it will now be shown how it can be used to improve a circuit's response to parameter variation. For this, the drain current I_D of a MOS transistor in saturation will be used as an example. As was outlined in equation 2.2, the drain current is subject to device size, oxide thickness and doping (via the threshold voltage), which makes it a good indicator for the most relevant types of variability.

The principle of the CAT device can be explained intuitively: Since the main transistor is subject to variability, the sampled value of the drain current may be anywhere on the normal distribution defined by the ideal (mean) drain current and a process-dependent standard deviation. In order to reduce the probability of the drain current being outside an ideally narrow window around the mean, calibration slices can be activated to effectively alter the width and thus the drain current of the device. This basic operation is illustrated in Figure 2.27, which visualizes the relative sizes of main transistor and calibration slice currents. The sizes of the slices are smaller than the main transistor, which allows the resulting drain current to be adjusted with reasonable granularity. Furthermore, the width of the main transistor is reduced with respect to a single transistor design by an amount equal to half the total calibration slice width, which allows the equivalent width of the CAT to be set to values both above and below the nominal width.

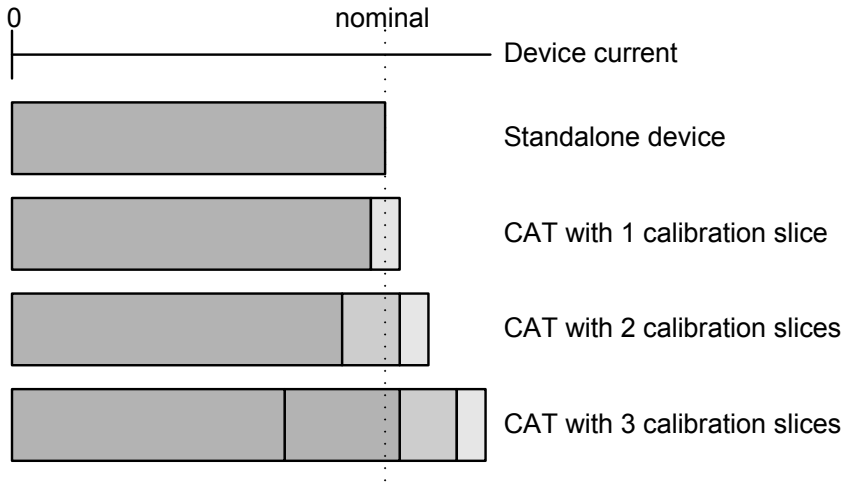


Figure 2.27: Illustration of CAT devices with different number of calibration slices.

The resulting distribution of drain currents is illustrated in Figure 2.28. The design is for a single transistor with a nominal drain current of $400\mu A$ and a standard deviation from

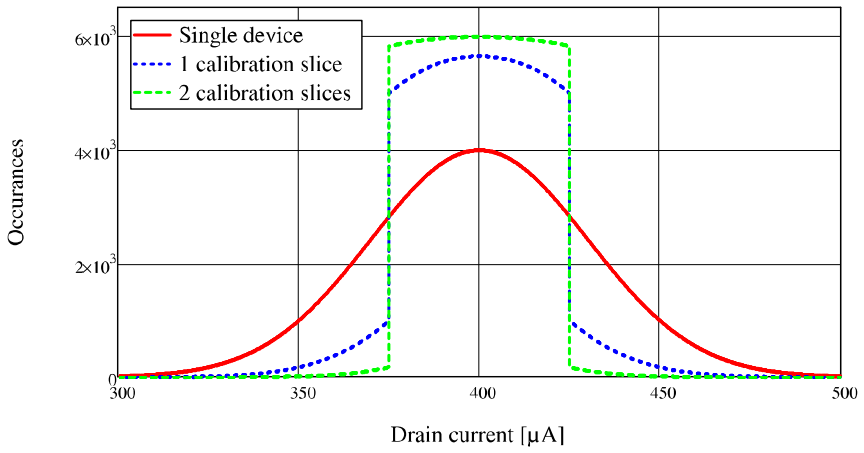


Figure 2.28: Effect of adding calibration slices.

process variation of $30\mu A$. In a first step, one calibration slice with a nominal current of $50\mu A$ is added and the main transistor current consequently reduced to $375\mu A$. This allows the CAT width to be selected for either $375\mu A$ or $425\mu A$, where the selection that results in a drain current closest to the nominal value is chosen. This adaptive adjustment of transistor width changes the distribution of drain currents; instead of a normal distribution, there is now a "peak" of currents around the mean, whereas fewer currents come to lie just outside of that window. By employing this adaptive method, standard deviation of the drain current is reduced to $18.1\mu A$. This can intuitively be illustrated as follows: For every sample, the current both with and without calibration slice is determined, and the best solution accepted. This means that currents that would nominally be too low can be "pulled" into the accepted range with the calibration slice active, and thus results in this accumulation around the mean. Likewise, currents can only be adjusted by whichever value the calibration slice permits, so there are still currents outside that windows. For very large deviations, the calibrated distribution follows the normal distribution again, as there is no way of improving such errant samples.

This performance can be further improved by adding more calibration slices. Also illustrated in Figure 2.28 is the result of a CAT with one main transistor and two calibration slices. The nominal drain current of the main transistor is reduced to $325\mu A$ and the calibration slices are sized in a binary fashion, thus allowing the effective width of the resulting CAT to be adjusted in steps of $50\mu A$ from $325\mu A$ to $475\mu A$. As would be expected, the introduction of further calibration slices further improves the CAT's ability to correct for device variability, as currents with a greater deviation can be pulled into the window of best adjustment, which is equal to the LSB current of the CAT. The standard deviation of the overall device current in this configuration is reduced to $14.8\mu A$

A further point to note is the size of the calibration devices. In the previous examples, the calibration slices were sized arbitrarily to illustrate the operation of the CAT. The relative sizing of the calibration slices to the main transistor has a significant effect on the overall performance. This is illustrated in Figure 2.29, which shows the distribution of current in a CAT with two calibration slices and minimum slice sizes of $25\mu A$ and $75\mu A$, respectively. The respective resulting standard deviations are $15.0\mu A$ and $21.1\mu A$. In the case where the size of the calibration device is increased, the window into which the currents are pulled becomes wider, meaning that the value to which the drain current can be calibrated becomes less precise. This is only to be expected, as the adjustment current increases with the device width. Conversely, reducing the width of the calibration slice narrows the calibration window, meaning that only currents that are already close to the mean can be calibrated, but this calibration is more precise. The optimization of sizing and number of calibration slices will be discussed in the following section.

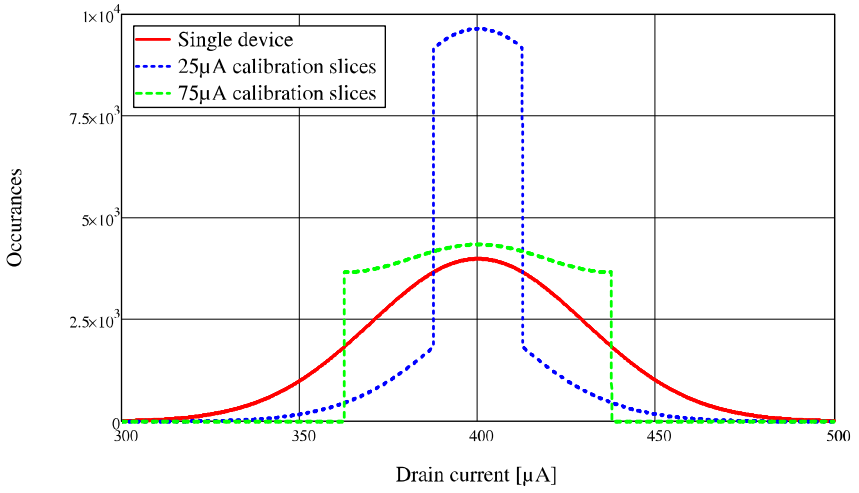


Figure 2.29: Effect of different calibration slice sizes.

The shape of the overall current distribution of the CAT is no longer Gaussian, as mentioned previously. Although the distribution at this point has to be computed numerically, its shape is easy to understand. For the purpose of illustration, consider a CAT with a single calibration slice that is relatively large compared to the main transistor. The resulting current distribution is depicted in Figure 2.30 and features an almost parabolic shape in the centre window, with an abrupt change to the normal standard distribution outwards. This shape is a direct consequence from the operation of two individual transistors. In order to facilitate calibration in both positive and negative direction, the nominal current of the main transistor is reduced by half the nominal current of the calibration slice, i.e. the original current distribution of the main transistor is shifted to the left. If the current is greater than this nominal current, the calibration slice remains deactivated as this would only increase the overall current. Thus, currents that are slightly greater than the mean current of the main transistor follow its distribution closely. For currents lower than this nominal value, the calibration slice is activated,

which shifts all currents lower than the nominal value by the calibration slice value, thus creating the left-hand distribution side at the upper end of the window. As the size of the calibration slice decreases relative to the mean, the two constituent distributions move closer together. As they overlap more and more, the "dip" in the CAT distribution becomes smaller, eventually moving towards the shapes in Figure 2.29.

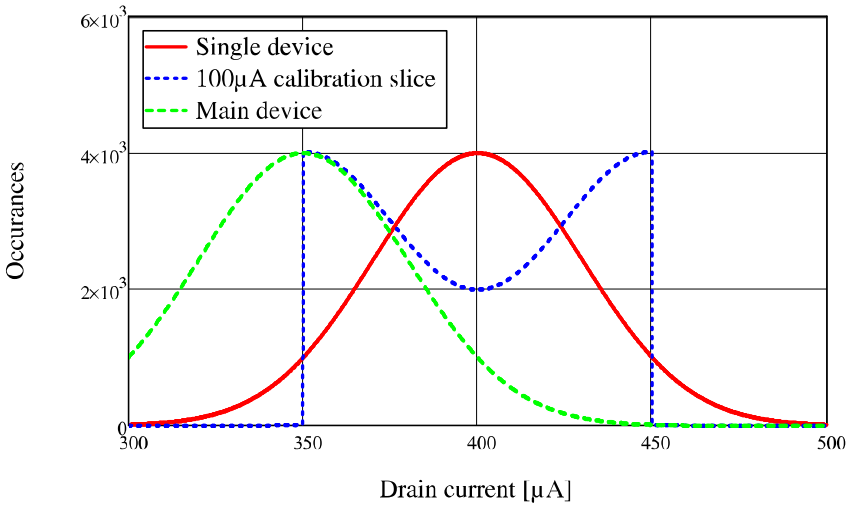


Figure 2.30: Effect of a very large calibration slice.

2.5.4 Device Size Optimisation

As established above, there are two main factors that affect the efficiency and performance of a CAT device: The number of calibration slices and the sizing of the calibration slices. This section outlines how both values can be optimized to achieve a maximum improvement over a standard transistor. A good metric for comparison of CAT and standard transistors is the standard deviation of the drain current. Although the current of a CAT is not normally distributed, the definition of standard deviation can still be applied, which allows a direct comparison.

The first objective to be optimized is the calibration slice size. For a given number of binary calibration slices and the smallest slice size, the resultant distribution can be computed numerically as explained in Section 2.5.3. The standard deviation of this distribution can then be related to the original Gaussian distribution, and the improvement in standard deviation calculated. This improvement in variability over a single transistor can then be plotted against slice size for a given number of slices. The resulting curve, shown in Figure 2.31, clearly has a maximum, which allows to select a slice size that will result in the lowest standard deviation. The general shape of this curve is to be expected, as very small slice sizes do not allow for a wide calibration range, whereas large slices

are too inaccurate to bring currents to within the desired window. Thus, there must be a certain slice size which results in the largest improvement in standard deviation.

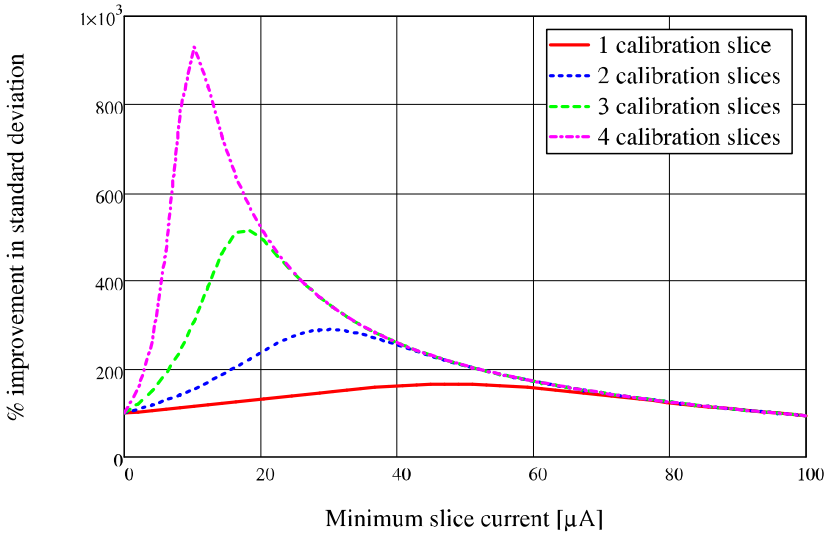


Figure 2.31: Improvement in standard deviation for different slice sizes and number of slices.

Next, the number of calibration slices can be considered. The graph of Figure 2.31 also contains plots for different numbers of slices. An increased number of slices increases the range from which currents can be pulled into the calibrated range, which leads to greater improvement in standard deviation. A more subtle point to note is that the slice size at which the maximum improvement occurs depends on the number of calibration slices. With this in mind, both the maximum achievable improvement and the slice size can be derived from the number of slices alone. This is illustrated in Figure 2.32, where the maximum improvement in standard deviation (corresponding to the height of the peaks in Figure 2.31) and the optimised slice size (the current at which each peak in Figure 2.31 occurs) are plotted against the number of calibration slices. Using this graph or a numeric solver, the designer can determine the optimised slice size and the achievable improvement in standard deviation by only specifying the desired number of slices. Likewise, a desired minimum improvement in standard deviation can be used to determine the required number of slices and their size.

The entire process of device size optimisation is purely numeric. The reason for this is that the algebraic expressions for the current distribution of a CAT contain conditional statements, which significantly complicates finding inverse functions or applying calculus to optimise the calibration device size. Since the CAT distribution function is merely a numeric integral, it is computationally cheap and therefore there are no significant drawbacks in using a numeric method over an algebraic one.

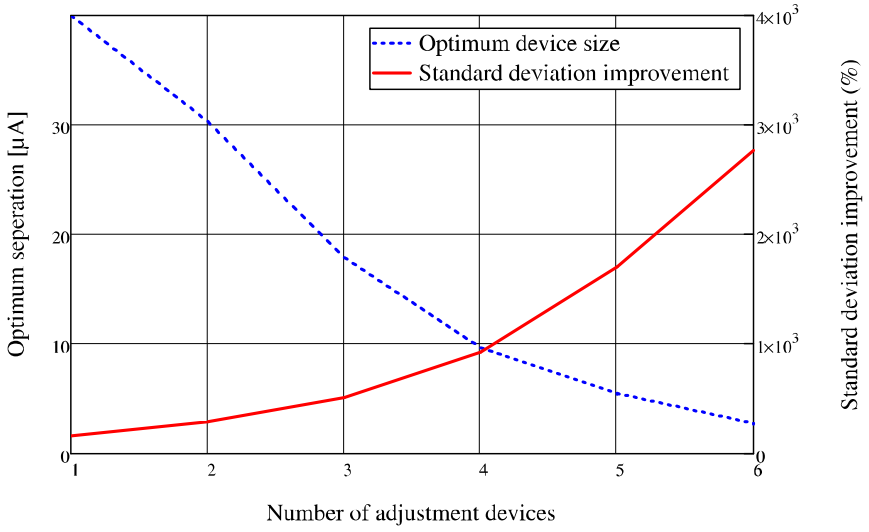


Figure 2.32: Theoretical maximum improvement and optimised slice size as a function of calibration slices.

2.5.5 Applicability

In principle, the CAT technique is applicable to any circuit that can be calibrated by adjusting the aspect ratio of a transistor. Although most device parameters (e.g. transconductance) are directly affected by a change in aspect ratio, the CAT technique operates at a slightly higher level and does not depend on a well-defined relationship between aspect ratio and device parameters. Instead of optimizing a single device parameter, the CAT is concerned with optimization of circuit parameters. For example, when calibrating the offset voltage of a comparator, one might intuitively want to calibrate the threshold voltage of the input transistors. Although threshold voltage fundamentally cannot be adjusted with the CAT technique as it does not depend on the W/L ratio, insertion of a CAT in the load current mirror allows for the effects of input offset voltage to be compensated. Thus, although the source of the disturbance was not eliminated by the CAT, the effects are the same on the circuit level.

Furthermore, application of the CAT is not limited to any specific circuit type or device placement within a circuit. Given a set of desired circuit parameters and the simulation tools necessary to determine these, the CAT support tools will ideally be able to determine a set of optimal CATs for any given circuit. The method of Section 3.2 achieves this through exhaustive simulations and therefore does not need any prior information on the circuit type or structure. However, it is limited by computing resources, which may require larger circuits to be split into functional blocks that are then considered separately.

Typically, the CAT configuration would not be fixed during calibration (e.g. with fuses), but rather be stored in memory, such as flash or a scan chain. This allows the CAT configuration of a system to be altered during normal system operation to compensate for environmental effects. Such effects can be virtually anything that changes circuit performance over time, such as ageing, temperature, radiation, etc. The CAT configuration that optimally compensates such effects can either be determined on-the-fly by forming a test loop over the system or through a static lookup table which relates an environmental parameter to the optimised CAT configuration. An application of CAT for combined process parameter and online temperature calibration is described in Section 3.4.

2.6 Summary

This chapter categorised and reviewed a number of state-of-the art techniques for calibration and robust design of analogue and mixed-signal circuits.

In Section 2.2, a number of generic calibration and robust design techniques have been discussed. The following is a brief summary of the techniques discussed:

- Trimming: Applicable to passive devices, requires special equipment and processes in most cases
- Substrate biasing: Adjusts threshold voltage, requires triple-well or silicon-on-insulator process if n-MOS devices need to be tuned
- Floating gates: Adjusts threshold voltage, requires special processes in most cases
- Calibration devices: Adjusts transistor aspect ratio, may introduce parasitic effects, no special process requirements

In addition to these specific techniques, system-level considerations were also discussed under the label of Digitally Adjustable Analogue Circuits in Section 2.2.6.

Complementing calibration techniques, Section 2.3 introduced the concept of inherently robust design. First, a number of general layout techniques that improve variation tolerance were discussed. Then, it was shown by means of an example that different circuits can improve variation tolerance while still performing the same function. Lastly, model-based robust design was introduced, where computer models and tools are used to optimise a circuit for variation tolerance.

Section 2.4 introduced the matching challenges faced in the design of segmented current-steering DACs and an overview of select solutions. In Section 2.4.2, a selection of design techniques that improve matching or relax matching requirements have been discussed.

Section 2.4.3 then introduced the concept of CalDACs as one way of calibrating the DAC's current sources to improve linearity. In contrast, the Switching Sequence Post Adjustment method discussed in Section 2.4.4 is an example of how a similar linearity improvement can be achieved solely through modification of the switching sequence. All of these techniques serve as background or reference for Chapter 4, where the CAT is applied to a segmented current-steering DAC to improve linearity.

Finally, Section 2.5 introduced the Configurable Analogue Transistor (CAT) and described in detail its operation and properties. The existing algorithm for calibration device size optimisation was also described in Section 2.5.4. All this serves as the basis for the following two chapters. Chapter 3 will introduce further development of the CAT technique, while in Chapter 4 the CAT is applied to a complex fabricated circuit and its characteristics are verified through measurements.

Chapter 3

Design-Time Methodologies for Configurable Analogue Transistors

3.1 Introduction

After having reviewed a number of methods and techniques to cope with variation in Chapter 2, this chapter presents three contributions to a particular calibration technique, the Configurable Analogue Transistor (CAT). One aspect of the CAT is that it is accompanied by a suite of automated design tools, which assist with equipping a circuit with CAT and post-fabrication calibration. However, most of the work on the CAT up to now has been concerned with its fundamental operation on a device level, with large parts of the automated design aspect still not complete. In this chapter, the methods and tools required for the automated design flow are developed. They are three closely related design-time methodologies that facilitate or extend the automated design flow in these specific areas:

- Automated Critical Device Identification (CDI): As discussed in Section 2.5.2, CDI is a key aspect of the automated CAT design flow, but no algorithm has thus far been developed. In Section 3.2, a methodology for CDI based on circuit sensitivity and adjustment independence is presented.
- Device Size Optimisation: The established algorithm for CAT size optimisation described in Section 2.5.4 is based on knowledge of drain current distributions. However, when CATs are applied to a circuit, they are used to calibrate overall circuit performance and not individual device currents. Therefore, CAT sizing must be based on statistical information of the overall circuit rather than individual devices. Section 3.3 details how optimised CAT sizes based on circuit performance can be obtained through Monte Carlo Simulation.

- Online Calibration: Whilst the CAT technique was originally envisaged as a means of post-fabrication adjustment against the effects of device variability, it can in principle also be used for on-line compensation of environmental variables. Section 3.4 shows how the CDI and size optimisation tools can be extended to consider environmental effects in addition to intrinsic variability.

In Sections 3.5 and 3.6 CDI and device size optimisation algorithms and the on-line calibration scheme are applied in simulation to two example circuits to demonstrate their viability and illustrate their operation. Finally, Section 3.7 describes the integrated software tool developed to automate the simulations and perform CDI and device size optimisation. The work in this chapter led to two publications. The combined work of Sections 3.2 and 3.3 was published at DATE 2012 [1] and the work of Section 3.4 was published at the 2012 IEEE Aerospace Conference [2].

3.2 Automated Critical Device Identification

3.2.1 Introduction

Referring back to the automated CAT design flow of Figure 2.26 in Section 2.5.2, Critical Device Identification is the first computer-aided step in the CAT design flow and is followed immediately by calibration device sizing. The two methods of CDI and device size optimisation presented in this work are linked closely together and depend on a large common set of operations and data. Therefore, CDI and device size optimisation can be explained as a single process, although each can be used independently. Figure 3.1 gives an overview of the steps necessary to perform CDI and device size optimisation, the details of which will be explained in this section and Section 3.3.

3.2.2 Method

- Step 1: A conventional sensitivity analysis is performed in simulation. By varying the width of each transistor by a small amount, the dependence of circuit performance to changes in individual devices is determined and recorded in a sensitivity table.
- Step 2: The sensitivity information is used to perform CDI, resulting in a list of transistors that are most suitable for adjusting circuit performance after fabrication. The first and second steps are discussed in Section 3.2.3.
- Step 3: A Monte Carlo (MC) simulation of the circuit is carried out to simulate the circuit under parameter variation. In each iteration, the critical devices are adjusted to minimise performance variability. The required adjustment for each

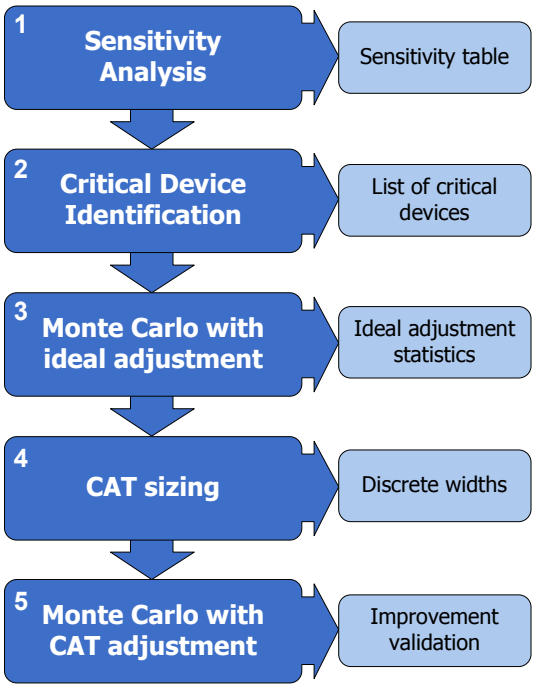


Figure 3.1: Process flow of critical device identification and application of configurable analogue transistors.

device is computed from the circuit performance and sensitivity information, which is then used in Step 4 to determine to optimised CAT sizes.

- Step 4: The critical devices are replaced by CATs sized according to Step 3 with a finite number of calibration transistors. The sizing of these CATs is based on statistical information from the ideal adjustment performed in Step 3. Introducing realistically sized CATs allows modelling of the post-fabrication calibration process and the resulting performance. Steps 3 and 4 are discussed in Section 3.3.

- Step 5: A final MC simulation is performed and in a similar manner to Step 3 the required adjustment for each critical device is calculated. However, instead of adjusting devices without constraints, transistor widths can now only be adjusted to the discrete widths of the CATs. By simulating each MC step again with the properly configured CATs, post-fabrication calibration of the circuit is simulated.

3.2.3 Sensitivity Analysis and CDI

The method for Critical Device Identification presented in this work is based on information about the sensitivity of circuit performances to transistor size. First, this information is obtained through conventional sensitivity analysis, where the width of each transistor in the circuit is in turn varied by a small percentage (e.g. 5%) of its nominal value and the impact on the circuit performances recorded. Sensitivity in this

context is a measure of how much a particular circuit performance is affected by changes in a particular transistor's parameters. The resulting sensitivity of each performance to a change in each transistor's width is normalised as a relative change to its nominal value, as defined in Equation 3.1

$$s_{A,n} = \frac{A|_{w_n=nom+5\%} - A|_{w_n=nom-5\%}}{A|_{w_n=nom}} \tag{3.1}$$

where A is a circuit performance (e.g. gain) evaluated with the width of transistor n, w_n, at different values, resulting in the sensitivity of performance A on transistor n, $s_{A,n}$. By normalising sensitivity in this manner, the dependencies on absolute values of both performance and transistor size are removed. This makes it possible to compare sensitivities of different transistors and performances, e.g. "If the width transistor T1 is changed by 1%, performance A changes by 2% and performance B changes by -4%".

The sensitivities obtained from the above method make two assumptions about the circuit. The first assumption is that there is a linear relationship between the performance and transistor width within the range of interest. The consequence of this assumption is that the calculated sensitivity values may be incorrect if the real sensitivity deviates significantly from the assumed linear behaviour. To ensure this assumption remains valid, the variation around the nominal transistor width should be as small as possible. In addition, the transistor operating region should be monitored during sensitivity analysis. If a transistor moves from its nominal operating region during sensitivity analysis, it is a very strong indicator that the variation in width was too great and that the calculated sensitivity value is likely incorrect.

The second assumption is that the circuit is a linear system within this range and that superposition is therefore applicable. This means that the sensitivities computed by varying one transistor at a time are assumed to remain the same even if multiple transistors are subject to variation. It will be shown later in this chapter that this assumption is acceptable for the purpose of CDI. For device size optimisation and post-fabrication calibration, where multiple devices are adjusted simultaneously, this assumption may no longer be adequate. In these cases, a numeric solver integrated with a circuit simulator is used to remove the dependency on assumptions of linearity.

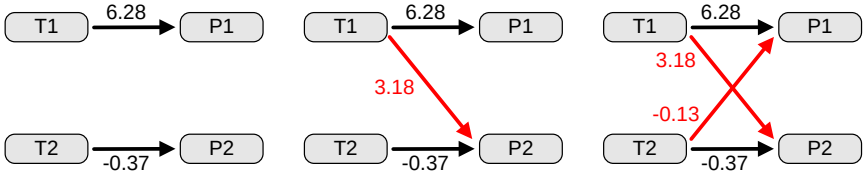


Figure 3.2: Different relationships between transistors and performances.

The second step in the CDI process is to determine the critical devices from the sensitivity information. A naive approach would be to find a set of transistors where each transistor affects a particular performance the greatest. Indeed, the parameter variation of these transistor is likely to contribute most to performance variation. However, there are several obvious flaws with this approach. The greatest problem in this case is the interdependence of adjustments, which is illustrated in Figure 3.2. In the scenario at the left of Figure 3.2, each adjustment transistor (T1 or T2) controls one, and only one, circuit performance (P1 and P2). Using the known sensitivities, the performances can be trivially adjusted by the transistors. However, such a behaviour is nearly impossible in any real circuit, because performances are typically tied to more than one transistor.

A more realistic case is illustrated in the centre scenario of Figure 3.2, where transistor T1 affects both performances P1 and P2. In this case, transistor T1 has to be adjusted to provide the desired calibration of performance P1. However, transistor T2 has to be adjusted to compensate for the effects of T1 on P2, as well as to provide the desired calibration of P2. While this does not seem a problem in principle since all relationships are known, the example sensitivities illustrate why this situation can be problematic: Suppose the effect of T1 on both performances is very large, while the sensitivity of P2 on T2 is small. In this case, it could be impossible to compensate for the unwanted effects of T1 on P2, because the resultant change in size of T2 would not be practical.

The case illustrated at the right of Figure 3.2 best describes most real circuits, where each circuit performance depends on a number of transistors. In this example, transistors T1 and T2 each affect both performances P1 and P2. This configuration suffers from the same problem as in the previous case where a strong cross-dependence may make successful adjustment of a performance practically impossible. Additionally, there is now a situation where a performance's sensitivity on one transistor is in the opposite direction of the other. This is illustrated again in the example sensitivities: Transistor T1 causes positive changes in both performances, while T2 results in negative changes. Suppose now that performance P1 needs to be adjusted upward and P2 downward, each by the same relative amount. First, transistor T1 is used to adjust P1 upward. However, this also results in an unwanted upward adjustment of performance P2. Now transistor T2 needs to be adjusted to achieve the initial adjustment of P2, but also compensate for the unwanted adjustment. This, in turn, results in an unwanted decrease in P1, which must be compensated by adjusting T1 again, and so on. It can be seen very quickly that under such circumstances, several unsatisfactory outcomes are possible. These range from impractical adjustment transistor sizes to cases where no solution can be found at all. While this example is extreme, similar situations are abound in real circuits. Also note that this issue is not related to the sequence of adjustments given in the example. With these numbers, simultaneously solving for both transistor sizes would also result in the same impractical solutions.

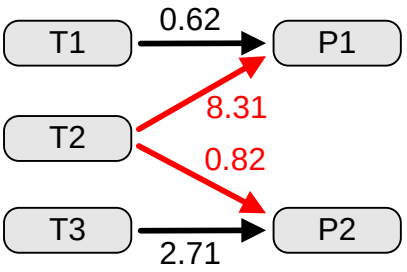


Figure 3.3: Illustration of adjustment independence.

It is therefore clear that merely choosing calibration transistors based on the magnitude of their effect on performances may not be the best choice. Knowing this, the opposite approach to CDI would be to pick the transistors that most independently affect each performance. Depending on the specific circuit, this may largely eliminates the problem of conflicting adjustment requirements. However, this method also suffers from a critical problem, which is illustrated in Figure 3.3. In this case, there are three transistors, T1, T2 and T3, and two performances to be adjusted, P1 and P2. From the shown relationships, T1 affects only P1 and T3 affects only P2. Thus, these two transistors allow completely independent adjustment of the performances. However, the sensitivity of P1 on T1 is comparatively low. In practice, this may mean that P1 cannot be adjusted adequately within practical limits for T1. A better choice in this case would be to use T2 to adjust P1, where the sensitivity is much greater. Unlike T1, T2 also affects P2, which was avoided previously. However, the effects of T2 on P2 are low compared to the sensitivity of P2 on T3, which means that T3 can likely be used to compensate for the unwanted effects of T2 on P2.

Therefore, an optimal set of calibration transistors will be one where each transistor maximally affects one specific circuit performance and minimally affects all others. The method of CDI presented in this section aims to combines the two goals of high sensitivity and high independence. Figure 3.4 will be used as an example as the process of CDI is explained below.

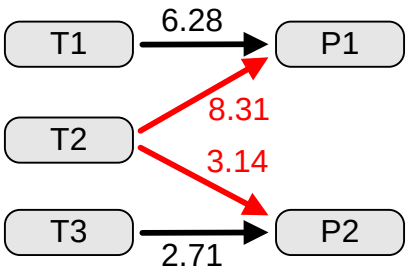


Figure 3.4: Example sensitivities of a circuit.

First, a normalised measure for independence is derived by calculating for each transistor the ratio of each sensitivity to all the sensitivities associated with that particular transistor, as per Equation 3.2

$$w_{A,n} = \frac{|s_{A.n}|}{\sum |s_n|} \tag{3.2}$$

Where A represents a certain performance, n a certain transistor and $w_{A,n}$ the relative amount of transistor n's total impact on performance A. The result, $w_{A,n}$, is a dimensionless number between 0 and 1, which is the fraction of transistor n's total effect on the circuit that affects performance A.

Sensitivities (s)

	P1	P2
T1	6.28	0
T2	8.31	3.14
T3	0	2.71

Weighting Factors (w)

	P1	P2
T1	1	0
T2	0.73	0.27
T3	0	1

Scaled Sensitivities (s')

	P1	P2
T1	6.28	0
T2	6.06	0.85
T3	0	2.71

Figure 3.5: Illustration of the Critical Device Identification process.

This serves as a measure for adjustment independence: If performance A's $w_{A,n}$ is much larger than all others for the same transistor, it indicates that adjusting this transistor will almost exclusively affect performance A and that therefore A can be adjusted very independently. Conversely, if several performances' $w_{A,n}$ are approximately equal for the same transistor, it indicates that these performances are affected almost equally by adjusting the transistor and that therefore a strong interdependence exists. Figure 3.5 illustrates this property for the given example. The table on the left contains the sensitivities, $s_{A,n}$, that were determined using Equation 3.1. Next to it is the table of weighting factors, $w_{A,n}$. In the case of transistors T1 and T3, the weighting factors are 1, because these transistors affect only one performance each. For T2, the weighting factors indicate how the effect of this transistor is divided amongst the performances. From this, it is immediately clear that although it has the largest effect, transistor T2 would be a bad choice for adjusting performances. For instance, if it were used to nominally adjust P1, it would also adjust P2 with about one third of the magnitude as an undesired side effect.

These $w_{A,n}$ values are then used as weighting factors for the original sensitivity table. Element-wise multiplication of the original sensitivity table and the weighting table is performed as per Equation 3.3

$$s'_{A,n} = s_{A,n} \cdot w_{A,n} \tag{3.3}$$

which results in a weighted sensitivity table, $s'_{A,n}$, that considers both absolute sensitivity and independence. For each performance, the device with the greatest weighted sensitivity is then selected as the critical device for adjustment. By scaling the original sensitivity table with the weighting factors, the sensitivity values for transistor-performance combinations that have a high interdependence with other performances are reduced, thereby promoting more independent transistor-performance combinations to be chosen as Critical Devices, even if their absolute sensitivities are lower. This is illustrated on the right of Figure 3.5, which shows the table of scaled sensitivities for the example. Due to scaling, the absolute magnitude of sensitivities to transistor T2 have been reduced and transistors T1 and T3 now have the largest scaled sensitivities for performances P1 and P2, respectively. Choosing the Critical Devices by largest scaled sensitivity will now recognise T1 and T3 as the ideal calibration devices that they are. Therefore, this method can achieve the original goal of finding Critical Devices that are suitable for post-fabrication adjustment.

This method for critical device identification has been designed to mimic the intuitive manual device selection process in a way that can be readily implemented in an EDA tool to allow systematic application to circuits. The above algorithm can be modified to incorporate a number of special requirements. For example, if certain performances needed to be assigned priority, the corresponding columns of the weighting table can be scaled by chosen factors. Similarly, if the impact of independence on device choice had to be changed, the entire weighting table can be raised to a certain power, which will further reduce the weighting factors for transistors which affect multiple performances. If certain devices should not be considered as CAT candidates, an exclusion list can be used during the CDI process. Additionally, the number of adjustment devices does not need to be the same as the number of performances. Tuning of a performance could potentially be improved if more than one calibration device for that performance is used.

3.3 Device Size Optimisation

3.3.1 Introduction

After Critical Device Identification, the next step in the computer-aided CAT design flow is to optimise the size of the CAT devices. The implementation of device size optimisation presented in this work is based on information obtained from the CDI algorithm of Section 3.2.3 and follows on directly from it. Although an algorithm for calibration device sizing already exists [81], it is based on statistical information of device currents, which is not applicable when adjusting overall circuit performance. The method presented in this section solves this problem by showing that the device size optimisation algorithm can be directly based on statistical information of the "ideal" transistor size adjustments that would lead to a perfectly adjusted circuit.

3.3.2 Method

Based on the sensitivity information obtained in Section 3.2.3 and the corresponding assumptions about linearity, the resulting change in circuit performance from adjusting the widths of the critical devices can be described by a system of linear equations, e.g. 3.4

$$\begin{aligned}\Delta A &= s_{A1}\Delta w_1 + s_{A2}\Delta w_2 + \cdots + s_{AN}\Delta w_N \\ \Delta B &= s_{B1}\Delta w_1 + s_{B2}\Delta w_2 + \cdots + s_{BN}\Delta w_N \\ \Delta C &= s_{C1}\Delta w_1 + s_{C2}\Delta w_2 + \cdots + s_{CN}\Delta w_N \end{aligned} \tag{3.4}$$

where ΔA is the change in performance A (e.g. bandwidth) from adjusting the width of transistor 1 by Δw_1, defined by the sensitivity of parameter A to a change in transistor 1, s_{A1}. Note that this sensitivity is not the same numerical value as in the sensitivity table of Section 3.2.3, which is a dimensionless value normalised to the nominal performance and to a certain relative change in transistor width (e.g. $\pm 5\%$). In order to evaluate the above equations the sensitivities must be de-normalised to obtain a gradient with units, e.g. $Hz/\mu m$.

Likewise, Equations 3.4 can be used to solve for the required changes in transistor size to achieve a certain adjustment of performances. In the first step of device size optimisation, this property is used to obtain statistical information about the required "ideal" transistor adjustments for a circuit that is subject to parameter variation. For this, a Monte Carlo simulation of device parameter variation is run on the circuit. In every iteration, the circuit performances are compared to their nominal values and the differences (e.g. ΔA) are calculated. In conjunction with the previously found sensitivities (e.g. s_{A2}), these are used to solve Equations 3.4 for the adjustment in transistor sizes (e.g. Δw_2) required to return the performances to their nominal values. This process is repeated for every MC iteration and the and required transistor adjustments are recorded to build statistical information about them.

It must be noted at this point that the above description still assumes that the circuit is a linear system with regards to transistor width adjustments, which is very unlikely the case for practical circuits. Therefore, straightforward application of the solutions to Equations 3.4 will likely not result in an optimised set of transistor adjustments. It is therefore necessary to employ a more suitable technique of finding these solutions. From the review in Section 2.3.3.3, this is a multi-variate, multi-objective optimisation problem, to find a set of transistor sizes that result in a set of performances that are closest to their nominal values. There is a large variety of solvers available for such a problem, but there is one particular constraint in this case: Since the solver must run in

each iteration of a Monte Carlo simulation, it must be reasonably fast. This rules out many classes of solvers, particularly stochastic ones.

However, in this particular case the optimisation problem can be greatly simplified: First, an approximation of the required transistor adjustment can be simply calculated from Equations 3.4, which are a good starting point for the solver. Second, this problem can be transferred to a single-objective problem by considering the total sum of errors in performance as the combined objective. This opens up a wide range of single-objective, multi-variate solvers to choose from. The choice was made to use a simple steepest descent solver, as it provides a reasonable trade-off in terms of number of iterations required versus time required for each iterations. Solvers that take advantage of the second derivative, e.g. the Levenberg-Marquardt algorithm [82], would require slightly fewer iterations, but spend significantly more time in each iteration, as the number of times the base simulation has to be run increases significantly. In the end, the choice of solver and algorithm is more of an engineering rather than an academic problem and the choice made provides adequate performance and robustness for the time being.

The existing device size optimisation process described in Section 2.5.4 is based on the mean and standard deviation of transistor drain current. However, in the higher-level view of CAT taken in this work, only circuit performances and not individual device currents are considered. Indeed, there may not even be a direct relationship between a device's drain current and circuit performance. Furthermore, the idea of constant current calibration slices that translate directly to given device widths is only applicable if the transistor is used in a constant-current sink or source application, which is not even the case in many circuits (e.g. cascode transistors) Therefore, device size optimisation must not be based on device currents, but on device size and circuit performance instead.

However, the fundamental principle of the device size optimisation algorithm of Section 2.5.4 is in fact not specific to transistor sizes and device currents. Instead, it is a numeric method that can be applied to any population of Gaussian variables with a choice of discrete means. It will determine the optimum spacing between the means such that the resultant standard deviation is minimised when the correct mean is chosen in each instance. In other words, the device size optimisation algorithm is applicable to any distribution where a CAT-like manipulation is possible.

Based on this crucial property, it is possible to directly obtain the calibration device sizes by using the distribution of ideal transistor adjustments as input to the device size optimisation algorithm. Under the premise that the system of linear Equations 3.4 and superposition are applicable, the resultant calibration device sizes are optimised because there is a linear relationship between the calibration device sizes and circuit performance. The distribution of ideal calibration device sizes is directly related to the performance distribution from process variation and therefore optimising the post-CAT device size distribution results in an optimised post-CAT performance distribution. In practice,

these relationships are of course not perfectly linear and it will be shown in Section 3.5 that there is a slight discrepancy between the performance achieved through optimised CATs and hypothetical optimised performance distributions.

The method for calibration device size optimisation presented in this section completes the set of design tools necessary to apply CAT to any given circuit. Section 3.5 contains a complete example of the CDI and device size optimisation process, which will give more insight into some of the practical aspects of the methods. Section 3.4 will now demonstrate how the CDI process of Section 3.2 and the device size optimisation process of this section can be extended to be not only based on device parameter variation, but to take into account external variables like temperature.

3.4 On-Line Temperature Calibration of Analogue Circuits

3.4.1 Introduction

The primary design goal of a CAT is to allow post-fabrication calibration to compensate for errors introduced by process variation. After the CAT design flow described in Section 2.5, each chip is individually tested and the optimal CAT settings to achieve best performance are determined. This optimal configuration is typically stored in non-volatile on-chip memory so that it can be restored whenever necessary, e.g. after the chip is powered up. Since both process and mismatch variation are largely time invariant, a static CAT configuration is sufficient to counteract any errors introduced by these mechanisms to achieve optimal performance. However, in this configuration the circuit is still subject to environmental influences, such as temperature, radiation and ageing. Performance degradation introduced by these means cannot be compensated with a static CAT configuration, which calls for an online calibration approach.

Online calibration of a circuit equipped with CAT is easy to understand, and requires the CAT configuration to be altered during run-time according to certain rules, much like the techniques discussed in Section 2.2.6. In principle, this involves measuring the current system performance and, if necessary, switching to a different CAT configuration that will improve performance. However, there are at least two complications in this generic case. First, to determine the current performance of the circuit, it may be necessary to suspend normal operation and put the circuit in a test mode. Second, determining the optimal CAT configuration can be an iterative process, during which the circuit is not likely to operate at optimal performance. The result from these issues is that the circuit will not be able to perform its normal operation continuously and that it may operate outside specifications during reconfiguration. In addition, the added complexity required for measuring circuit performance and generating the required test conditions may be unsuitable for many circuits. This section introduces a method for employing

CAT to calibrate a circuit for the effects of temperature variation. In order to overcome the aforementioned problems during reconfiguration, this method uses a pre-computed lookup table which directly links temperature to the required CAT configuration. Therefore, this method allows true on-line calibration without interruption of normal circuit operation, and with minimal support circuitry.

3.4.2 Method

Online calibration of CATs with respect to temperature is a special case that lends itself well to practical implementation. The dependence of circuit performance on temperature is well described through SPICE models and the temperature of the chip can be easily measured continuously, which allows the system to conduct the appropriate reconfiguration before the performance has dropped below a threshold. Additionally, the temperature behaviour of the circuit can be accurately modelled before fabrication or measured for each chip after fabrication, which reduces the reconfiguration process to a lookup table. This type of online reconfiguration can be carried out without any interruptions in the operation of the circuit, because the current performance does not need to be measured and the optimal configuration is predetermined. However, signals processed in the system may still be subject to short glitches at the moment when the CAT configuration is changed.

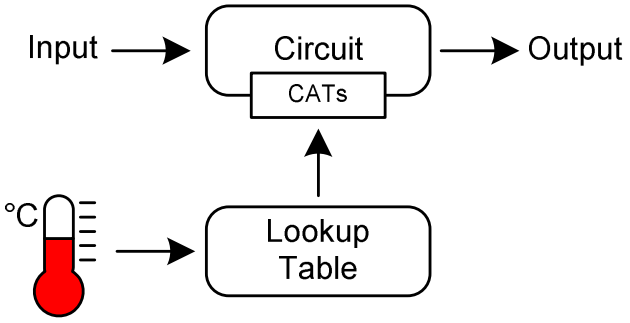


Figure 3.6: System structure for online temperature calibration using CATs.

Figure 3.6 illustrates the required system architecture for online CAT reconfiguration. The temperature of the chip is continuously monitored, and the corresponding optimal CAT configurations obtained from a lookup table. There are several points to note about this concept. First, in most practical applications, temperature does not need to be measured continuously. Instead, it may be sufficient to sample its value at given intervals or only under certain conditions. Discontinuous sampling of temperature also reduces power consumption, since the temperature sensor and the associated reconfiguration hardware operate only for short periods. Secondly, the task of digitizing temperature readings and looking up the corresponding configuration words in memory bear very little computational load. It is therefore practical to handle this task in an already

existing digital processing system, rather than a dedicated computer for CAT reconfiguration. Again, this is especially beneficial for applications in which energy conservation is a primary requirement. In summary, the hardware overhead for incorporating online CAT reconfiguration is potentially very low. Apart from the CATs themselves, the only other required on-chip component is a temperature sensor, which may be as simple as an appropriately biased PN junction. All remaining components, such as the ADC, computation, lookup table and configuration memory may be incorporated into an existing signal processing system at little additional cost.

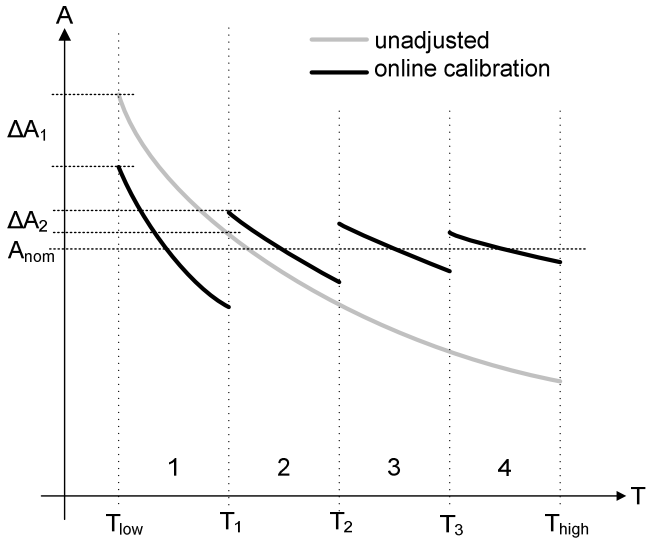


Figure 3.7: Principle of temperature compensation using CATs.

3.4.3 Design of CAT for Online Temperature Calibration

The CAT design processes when considering temperature variation is in principle no different than these introduced in Sections 3.2 and 3.3. However, instead of performing a Monte Carlo simulation across the process parameter space to gain stochastic information about the circuit's performance, a temperature sweep across the specified range is sufficient. Figure 3.7 shows the temperature dependence of an example circuit performance, A, exhibiting a negative temperature coefficient. To find the ideal sizes of the CAT devices, the established device size optimisation algorithm can be used by again finding the ideal critical device adjustment at each temperature step. The resulting CATs will be sized such that the mean deviation from the nominal value over the entire temperature range is minimized. The configuration lookup table, which maps temperature to the CAT configuration can be derived directly from the ideal critical device adjustments by appropriately assigning them to the closest discrete CAT configuration for each temperature.

For the purposes of illustration, a possible outcome of calibrating the example performance with a 2-bit CAT device is also shown in Figure 3.7. The CAT configuration that is active in a certain temperature range is indicated by numbers along the temperature axis. For very low temperatures, configuration 1 is chosen, which reduces the numerical value of the performance by ΔA_1. This reduction in value brings the mean of the performance between T_{low} and T_1 closer to the nominal performance, A_{nom}. If the temperature rises above T_1, configuration 2 is chosen. This reduces the performance only by ΔA_2, thereby bringing the performance closer to the nominal value, and so on. This example should also reinforce the point that neither the temperatures at which the configurations change nor the sizing of the CAT devices, corresponding to the change in performance, are arbitrary, but are optimized during the design stage.

While this approach to temperature compensation is valid for a single chip at nominal device parameters, it does not consider the various parameter variation processes that occur in real circuits. A real circuit design, which includes optimised CAT devices, is replicated several times on a wafer to yield a large number of chips. While ideally all chips of the same design have identical behaviour, in reality the performances of any two chips and indeed identical devices on the same chip are not the same. This is due to a number of different variation processes as discussed in Chapter 1.

The consequences of these variation mechanisms on the application of CATs to compensate temperature variation are two-fold. Firstly, because the designed CAT must provide good results on all produced chips of a given circuit, optimised sizing of the CAT must now consider both temperature and parameter variation. This brings the CAT from simple temperature sweeps back to its original stochastic domain, where the temperature can be considered as an additional random variable. Secondly, because the CAT must now compensate parameter and temperature variations, the achievable level of calibration will be lower than in the case where only temperature or parameter variation is considered. Nevertheless, the expected improvement in performance variation is still well defined through the stochastic processes.

A crucial difference between the temperature-only and variation-aware CAT application lies in the post-fabrication stage. In the case where only temperature is considered, it is sufficient to generate a single configuration lookup table from the simulations that is valid for all chips of a particular circuit. When considering additional parameter variations, not only must the initial CAT configuration be determined on a chip-by-chip basis, but also an individual lookup table generated for each chip. This is necessary because both the initial CAT configuration and the temperature behaviour are likely different between chips. However, using the existing circuit models, the lookup table for each chip can be generated in software from post-fabrication measurements without the need for a full temperature sweep. Therefore, generating the per-chip lookup table does not require any additional post-fabrication test equipment compared to regular CAT application.

In the design stage, the process for regular variation-only CAT and variation-aware temperature compensation CAT is virtually the same. CDI is again achieved by means of a weighted sensitivity table without any changes at all to the process. For device size optimisation, temperature is added as a further random variable to Monte Carlo simulation. This means that each Monte Carlo iteration represents a particular circuit at a particular temperature. Since the CAT technique is based on stochastic information about circuit performance, it is possible to derive all necessary information for CAT sizing from this, without the need for a full temperature sweep for every set of Monte Carlo parameters. Indeed, such information would be meaningless for CAT sizing because the CAT is sized for a particular circuit, of which multiple copies are produced on a single wafer. CAT sizes can therefore not be optimized for a single chip, but for a particular circuit, for which only stochastic information is relevant. Once the critical devices have been identified and sized, the remainder of the design process is carried out as normal.

Whilst based on the same tools and methods, the regular offline CAT and the online calibration scheme described in this section cover two distinct application areas of calibration. When used offline, the CAT configuration is determined once after fabrication and remains fixed thereafter. This facilitates improvement of production yield, which is the relative number of fabricated devices whose performance meets specifications. By using CAT for post-fabrication offline calibration, the performance of some devices which would normally not meet specifications can be adjust to meet specifications, and thereby improve yield. On-line calibration of CAT does not affect yield at all, but improves reliability. Reliability in this sense means that a device will be able to operate within specifications despite external or internal effects which would normally lead to degradation of performances. By adjusting the CAT configuration based on measured inputs, the performance of the circuit can be altered to respond to respond to such external effects and therefore ensure its performance remains within specifications. The example chosen in this section is temperature, but it would be conceivable to tailor an online calibration scheme which considers other effects, such as ageing or radiation.

In order to illustrate this design and calibration process, Section 3.6 contains an example where a circuit is equipped with CAT to compensate both parameter variation and temperature. This example therefore contains implements both offline and online calibration. The example also discusses a few more detailed points which were only touched upon in this description.

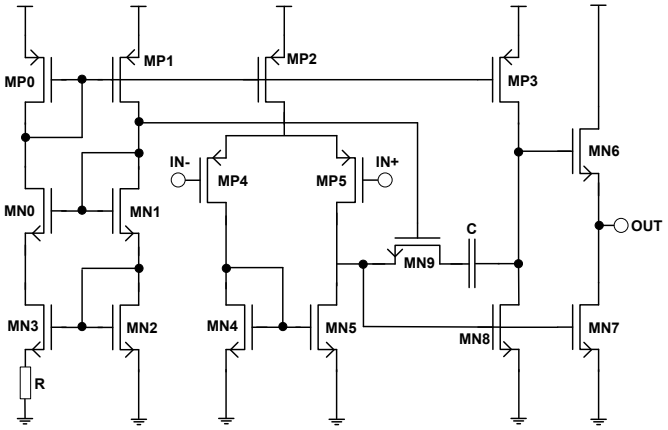


Figure 3.8: Circuit diagram of the operational amplifier.

3.5 Case Study 1: Application of CDI and Device Size Optimisation

In this section, the CDI and device size optimisation methods presented in Sections 3.2 and 3.3 are applied to an operational amplifier in simulation to illustrate the process and prove their viability. Figure 3.8 shows the circuit of the operational amplifier, which is used as the case study circuit in this section. The operational amplifier consists of a differential input stage, MP4 and MP5, a common-source gain stage, MN8, and a common-drain output buffer, MN6. MN9 and the capacitor form the internal compensation network and MP0-MP1, MN0-MN3 and the resistor form the bias circuit. The performance characteristics of the operational amplifier that are considered in the case study are DC voltage gain, open-loop bandwidth and common-mode rejection ratio. The operational amplifier is designed in a standard $0.35\mu m$ CMOS process at 3.3V, with active and passive component values as listed in Table 3.1.

It must be pointed out at this stage that these device sizes were designed with practical layout considerations in mind, which must be re-assessed when CAT is applied. For example, all transistor channel widths are multiples of $0.35\mu m$, which allows matched devices to be laid out using interdigitated transistors, as discussed in Section 2.3. However, when transistors are replaced by CAT, the resultant optimised calibration device sizes may not necessarily fit into the existing sizing scheme. For the purpose of this example, these layout constraints are ignored and devices of arbitrary size within reasonable bounds accepted. In practice, the CAT devices may need to be altered from their calculated values to integrate them in the desired layout.

As outlined in Section 3.2, the first step is to perform a sensitivity analysis. The left section of the table in Figure 3.9 shows the normalised sensitivities of the three circuit performances for a $\pm 5\%$ change in transistor width, as per Equation 3.1. This sensitivity table clearly illustrates the problem with adjustment interdependence: If only the

Device	Dimensions	Device	Dimensions
MP0	$8.75\mu m$ / $0.35\mu m$	MN2	$35\mu m$ / $0.35\mu m$
MP1	$8.75\mu m$ / $0.35\mu m$	MN3	$35\mu m$ / $0.35\mu m$
MP2	$105\mu m$ / $0.35\mu m$	MN4	$52.5\mu m$ / $0.35\mu m$
MP3	$105\mu m$ / $0.35\mu m$	MN5	$52.5\mu m$ / $0.35\mu m$
MP4	$105\mu m$ / $0.35\mu m$	MN6	$175\mu m$ / $0.35\mu m$
MP5	$105\mu m$ / $0.35\mu m$	MN7	$175\mu m$ / $0.35\mu m$
MN0	$8.75\mu m$ / $0.35\mu m$	MN8	$98\mu m$ / $0.35\mu m$
MN1	$8.75\mu m$ / $0.35\mu m$	MN9	$50\mu m$ / $0.35\mu m$
Component	Value	Component	Value
R	$823k\Omega$	C	623fF

Table 3.1: Device sizes and component values.

absolute magnitude of sensitivity is considered, MP0 and MP1 are identified as critical devices because they have the greatest impact on bandwidth and gain, respectively. Both transistors affect both performances with a similar magnitude, but in opposite directions. Trying to adjust both gain and bandwidth in the same direction using these devices would result in a conflict because the effect of adjusting one transistor will be largely negated by adjusting the other. Although a mathematical solution can still be found in this case, the resulting numerical values for width adjustment are likely to be unrealistically large and therefore impractical. Likewise, only considering adjustment independence will not result in a satisfactory device selection, either: The bandwidth can be most independently tuned by adjusting MN9, which has no effect on other performances. However, the sensitivity of bandwidth to changes in MN9 is very small, meaning that to significantly alter this performance the required device width change would be extremely large and possibly not feasible. It is therefore clear that neither absolute sensitivity nor adjustment independence are good metrics to determine the calibration devices. Instead, the CDI process developed in Section 3.2 is used.

The right section of Figure 3.9 shows the weighted sensitivities, as per Equation 3.3. When selecting the devices with the largest weighted sensitivity for each performance, the critical devices are MP0 for bandwidth, MP3 for gain and MN7 for CMRR. The rows relating to these transistors have been shaded and the corresponding sensitivities set in bold throughout the table as a visual aid. Note that MP3 and MN7 adjust their respective performances almost exclusively. Although MP0 significantly affects both gain and bandwidth it is still chosen as a critical device for bandwidth due to its very high absolute sensitivity to this performance. It will be seen in the simulations that for this particular circuit, this interdependence is not a problem, as MP3 is able to compensate enough of the undesired effects of MP0. However, this may not be the case for other circuits. If this interdependence does constitute a problem, the CDI algorithm can be tuned by giving more weight to adjustment independence, as outlined in Section 3.2.

After having determined the critical devices for this circuit, CAT device size optimisation

Condition	Performance			
		Gain	BW (Hz)	CMRR
Nominal	Mean	$1.83 \cdot 10^3$	$471 \cdot 10^3$	$30.5 \cdot 10^3$
Monte Carlo	Mean	$1.83 \cdot 10^3$	$474 \cdot 10^3$	$30.5 \cdot 10^3$
	Standard deviation	264	$88.3 \cdot 10^3$	$2.96 \cdot 10^3$
Monte Carlo with ideally sized CATs	Mean	$1.83 \cdot 10^3$	$471 \cdot 10^3$	$30.5 \cdot 10^3$
	Standard deviation	13.1	201	$88.7 \cdot 10^{-3}$
	Standard deviation improvement	95.0%	99.8%	99.9%

Table 3.2: Circuit performance under variation before and after ideally sized CATs are applied.

		MP3	MP0	MN7
Nominal width		$105\mu m$	$8.75\mu m$	$175\mu m$
CAT width step		$11.64\mu m$	$0.743\mu m$	$10.20\mu m$
CAT device	Main	$64.26\mu m$	$6.150\mu m$	$139.3\mu m$
	1^{st}	$11.64\mu m$	$0.743\mu m$	$10.20\mu m$
	2^{nd}	$23.28\mu m$	$1.486\mu m$	$20.40\mu m$
	3^{rd}	$46.56\mu m$	$2.972\mu m$	$40.80\mu m$

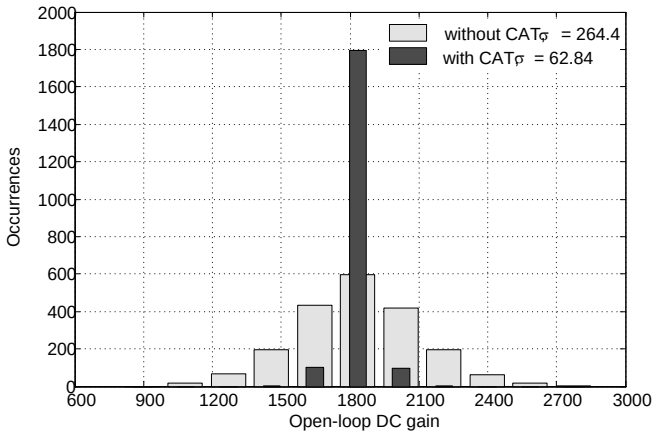
Table 3.3: Optimised sizes of the Configurable Analogue Transistors.

can be performed, as described in Section 3.3. Performing the Monte Carlo simulation for 2000 iterations with perfect device adjustment in each step yields the data in Table 3.2. It can be clearly seen that by adjusting the critical devices according to the above algorithm, standard deviation in circuit performance can be reduced by between 95% and 99% compared to the unadjusted values. Theoretically, the standard deviation of all performances after tuning could be reduced to 0 if the adjustment is convergent within device size limits in each case. In addition, the histograms of the required CAT size adjustments are shown in Figure 3.10.

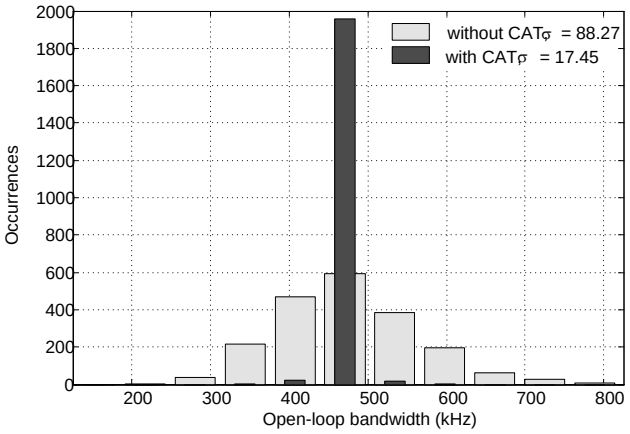
Using the information from Figure 3.10 as the input for the sizing algorithm results in the optimised CAT sizes shown in Table 3.3. As expected, MP3 shows the largest relative CAT devices, since it must be able to compensate for the undesired effects on gain from adjusting MP0. As pointed out at the beginning of this example, these resultant device sizes must be adjusted to make them practical should the circuit be fabricated. For example, it is not realistic that the precision of the device sizes given in Table 3.3 could be achieved in practice. Furthermore, these sizes are incompatible with the base unit size of $0.35\mu m$ used for the nominal device sizes, precluding integration of these devices into the circuit. Whilst such constraints are not considered by the CAT sizing tool, the designer can manually adjust the device sizes to suit the layout requirements. For example, the main device of MP0 could be rounded to $18 \cdot 0.35\mu m = 6.3\mu m$, and the calibration sizes to multiple of $2 \cdot 0.35\mu m = 0.7\mu m$, which would once again allow interdigitated transistors of width $0.7\mu m$.

After obtaining these transistor sizes, a further MC simulation is performed to simulate

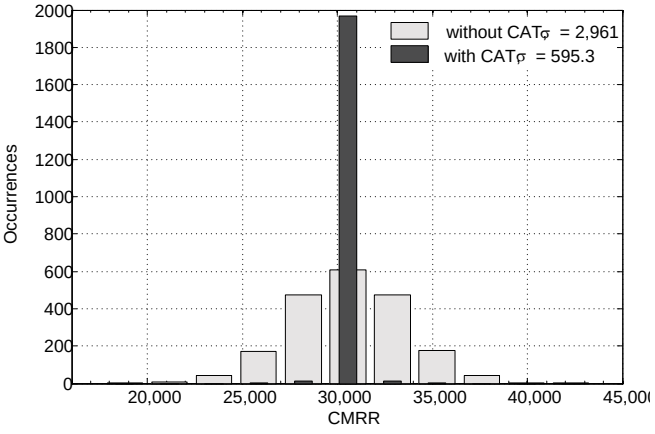
the expected performance of the circuit with CAT. As before, the optimal adjustment in width is calculated for each of the critical devices in each iteration. However, instead of adjusting devices with infinite granularity, the CAT configuration is now used to constrain the adjustment to the eight selectable widths. This is equivalent to the process used to tune the CATs on a chip after fabrication. The results of this simulation therefore represent the performance obtained following post manufacture adjustment with optimised CAT sizes.



(a) Gain



(b) Bandwidth



(c) Common-Mode Rejection Ratio

Figure 3.11: Histograms of performances before and after CAT adjustment.

Figure 3.11 shows the histograms of the performances before and after application of the

Condition	Performance			
		Gain	BW	CMRR
Monte Carlo	Mean	$1.83 \cdot 10^3$	$474 \cdot 10^3$	$30.5 \cdot 10^3$
	Standard deviation	264	$88.3 \cdot 10^3$	$2.96 \cdot 10^3$
Monte Carlo with CATs sized as per Table 3.3	Mean	$1.83 \cdot 10^3$	$471 \cdot 10^3$	$30.5 \cdot 10^3$
	Standard deviation	62.8	$17.5 \cdot 10^3$	595
	Standard deviation improvement	76.2%	80.2%	79.9%
	Maximum possible standard deviation improvement	80.7%	80.8%	80.3%

Table 3.4: Circuit performance after applying CAT.

CATs. It can be clearly seen that the spread in all three performances is reduced significantly, resulting in a lower standard deviation and greater yield. Table 3.4 compares the standard deviations of the performances before and after application of the three CAT devices. The standard deviations are improved by 76.2%, 80.2% and 79.9% of their original values for gain, bandwidth and CMRR, respectively.

The last row of Table 3.4 also shows the theoretical best attainable improvement for every performance by applying the CAT sizing algorithm from [81] to the MC performance distributions from Table 3.4. The theoretical maximum is almost impossible to achieve in practice because the relationship between CAT width and the performance it adjusts is unlikely to be perfectly linear. Therefore, evenly spaced adjustment slices in the performance domain would lead to irregularly spaced adjustment widths, which cannot be implemented with binary sized calibration transistors. The results from the circuit equipped with CAT, however, come remarkably close to the theoretical maximum values of improvement, indicating that the assumption of linear sensitivity is adequate at least for this particular combination of circuit, parameters and variation.

3.6 Case Study 2: Variation-Aware Temperature Compensation Using CAT

For the second case study, the CAT online calibration scheme of Section 3.4 is applied to a standard instrumentation amplifier comprised of three identical operational amplifiers and passive components. The amplifier schematic diagram is shown in Figure 3.12. The resistors in the instrumentation amplifier are modelled as p-type diffusion resistors.

The operational amplifier used in the instrumentation amplifier is the same circuit as in Section 3.5 with the transistor sizes already given in Table 3.1. Table 3.5 shows the resistor values for the instrumentation amplifier. The nominal circuit performances are listed in Table 3.6. The circuit is implemented in a standard $0.35\mu m$ CMOS process,

Component	Value	Component	Value
R1	$99k\Omega$	R5	$100k\Omega$
R2	$99k\Omega$	R6	$100k\Omega$
R3	$100k\Omega$	R7	$2k\Omega$
R4	$100k\Omega$		

Table 3.5: Instrumentation amplifier component values.

Performance	Symbol	Nominal value
Gain	G	94.35
Bandwidth	BW	$1.99 \cdot 10^{6}$
Offset voltage	VOS	$5.55 \cdot 10^{-5}$

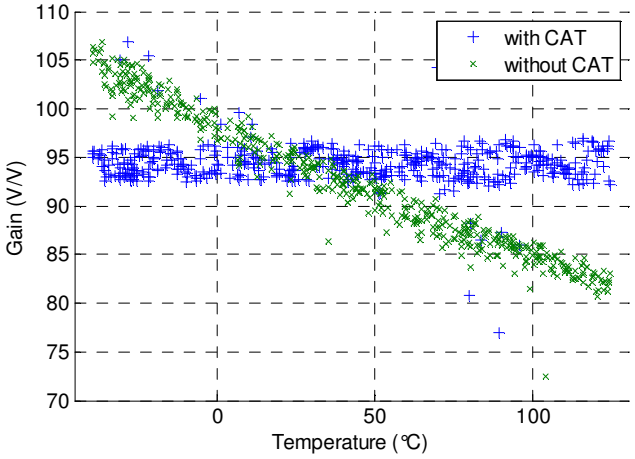
Table 3.6: Instrumentation amplifier nominal performances.

where the foundry device SPICE models are valid within the temperature range from $-40°C$ to $+125°C$. The critical devices in this circuit, found following the algorithm described in Section 3.2, are R7, MP0B and MP5B for gain, bandwidth and offset voltage, respectively. This resistor and the MOSFETs of operational amplifier B are coloured red in Figures 3.12 and 3.13.

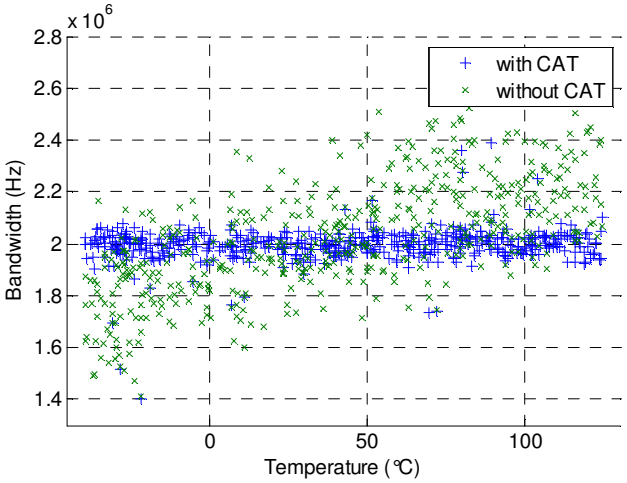
It is worth noting that one of the resistors has been identified as a critical device. Because the operational amplifiers operate under negative feedback, gain is primarily determined by the passive components with very little effect from active components. Therefore, the only viable way to adjust gain is to equip certain passive components with a CAT structure. While CAT cannot be directly applied to most passive components, an equivalent result can be achieved by a series or parallel combination of main and calibration components. In the case of resistors, the calibration devices need to be connected in series so that equally spaced calibration slices are possible. Each series resistor has a transmission gate in parallel through which the resistor can be by-passed, as illustrated in Figure 3.14. In effect, this allows the CAT design and calibration scheme to be applied to passive components. Bandwidth, on the other hand, is determined by the passive components and the gain-bandwidth product of the operational amplifier and can therefore be adjusted through transistors. Finally, offset voltage is not dependent on the passive components at all and can be adjusted by varying the transistors in the differential input stages.

Then, to perform device size optimisation, a Monte Carlo simulation of the instrumentation amplifier was performed as described in Section 3.3. In addition to process and mismatch variation, temperature was added as an additional random variable in the Monte Carlo simulation. This means that each Monte Carlo iteration represents a particular circuit at a particular temperature. Figure 3.15 shows the results of the Monte Carlo simulation, with the circuit performances plotted against temperature and marked with the symbol $\times$ and coloured green. Gain and bandwidth clearly show a dependency on temperature, indicated by the trend change over temperature. The individual points

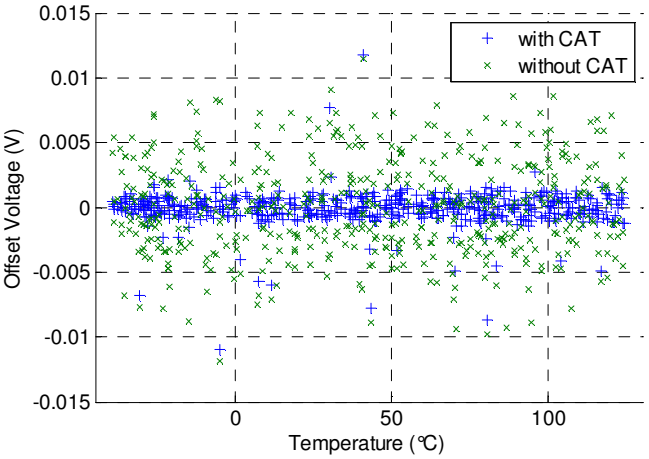
are scattered around this nominal temperature dependence. This scattering represents the magnitude of process parameter variation. In the case of offset voltage, the influence of mismatch variation outweighs the temperature dependence by far. This means that for a particular chip, CAT will be able to reduce the inherent offset voltage well, but online temperature calibration will not be able to improve temperature drift greatly. Conversely, in the case of gain, the temperature dependence is much larger than parameter variation, which means tha CAT will be able to mostly compensate for temperature, but not parameter variation. This can be visualized by the fact that each CAT transistor only has a finite number of possible configurations. When the effects of process parameter or mismatch variation are small compared to the effects of temperature, only a small set of the possible configurations are required for the initial post-fabrication tuning, leaving ample free configurations for online calibration. However, if the effects of process parameter variation outweigh the effects of temperature, the majority of configurations are required for the initial calibration, leaving few or no free configuration states for online temperature calibration.



(a) Gain



(b) Bandwidth



(c) Offset Voltage

Figure 3.15: Temperature dependence of performances with process and mismatch variation.

Condition	Performance			
		Gain	BW	Vos
Nominal	Mean	94.35	$1.99 \cdot 10^6$	$5.55 \cdot 10^{-5}$
Monte Carlo	Mean	92.49	$2.03 \cdot 10^6$	$7.05 \cdot 10^{-5}$
	Standard deviation	6.85	$2.22 \cdot 10^5$	$3.69 \cdot 10^{-3}$
Monte Carlo with CATs sized as per Table 3.8	Mean	94.26	$2.00 \cdot 10^6$	$6.70 \cdot 10^{-5}$
	Standard deviation	2.81	$6.51 \cdot 10^4$	$1.39 \cdot 10^{-3}$
	Standard deviation improvement	59.0%	70.7%	62.3%

Table 3.7: Circuit performance after applying CAT.

		R7	MP0B	MP5B
Nominal size		$11.45\mu m$	$8.75\mu m$	$105\mu m$
CAT size step		$0.52\mu m$	$0.70\mu m$	$2.91\mu m$
CAT device	Main	$9.64\mu m$	$6.33\mu m$	$94.81\mu m$
	1^{st}	$0.52\mu m$	$0.70\mu m$	$2.91\mu m$
	2^{nd}	$1.04\mu m$	$1.39\mu m$	$5.82\mu m$
	3^{rd}	$2.07\mu m$	$2.77\mu m$	$11.65\mu m$

Table 3.8: Optimised sizes of the Configurable Analogue Transistors.

Without calibration, the standard deviation of gain is 6.85, the standard deviation of bandwidth is 222kHz and the standard deviation of the offset voltage is 3.69mV. These results are listed in the upper section of Table 3.7.

For each Monte Carlo parameter set, the optimised ideal adjustment in the critical devices is determined by applying the method described in Section 3.3. The device size optimisation algorithm for CAT is then applied to these ideal adjustments to give optimised sizes for CATs with three adjustment transistors, resulting in eight discrete adjustment steps for each critical device. The resulting sizes of the CAT devices are listed in Table 3.8. For the simulation results after adjustment, each Monte Carlo parameter set, the configuration of each CAT is chosen to be closest to the ideal adjustment.

The performances after CATs have been introduced in the circuit are marked with + in Figure 3.15. Over the entire temperature range, the standard deviations of gain, bandwidth and offset voltage are reduced to 2.81, 65.1kHz and 1.39mV, respectively. These calibrated performances are listed in the lower section of Table 3.7. As can be seen, introduction of CATs again significantly reduces variation in performances over temperature. It is worth noting that a small number of performances after CAT calibration are significantly further from the nominal values than the majority. These points correspond to parameter sets for which no improvement in performance could be achieved within the adjustment constraints. Such instances will also occur in a real set of chips, where there will be a small number that cannot be calibrated at all. Because such circuits are already identified at the post-fabrication adjustment stage, they can be discarded as necessary.

Although the results obtained thus far show a significant reduction in performance standard deviation over the entire temperature range, no statement about parametric reliability has yet been made. For each performance, a pass band can be defined around the mean within which that performance is considered to operate to specification. For both Monte Carlo and CAT calibrated performances, parametric reliability can then be defined as the probability of a certain circuit at a certain temperature being within these bands. Although this definition is equivalent to the definition of yield, there is a practical difference introduced by the inclusion of temperature.

Whilst yield is concerned with the probability of a circuit meeting specifications at static operating conditions, parametric reliability is concerned with the circuit meeting specifications over the entire range of operating conditions. For the purpose of illustration, the pass bands have been defined as ± 5, $\pm 100kHz$ and $\pm 1mV$, for gain, bandwidth and offset voltage, respectively. For the unadjusted Monte Carlo results, the system's parametric reliability is 4.0%. When CAT is applied to the system, parametric reliability increases to 80.8%. This is a significant increase in the parametric reliability of the system. Although this improvement in reliability seems exceptionally large when compared to the improvement in standard deviations, it is not surprising. Firstly, standard deviation is greatly affected by even a few outliers, while they do not contribute as greatly to a decrease in reliability. Secondly, reliability requires all three performances to be within the pass band, which is very improbable in the uncompensated case, leading to a low uncompensated reliability.

Like the example in Section 3.5, the simulations carried out in this section rely heavily on device models. Whilst the demonstration of the method itself does not require realistic models, they are an absolute necessity for practical implementation. The online calibration scheme therefore not only requires accurate variation models, but the device models must also show the correct temperature dependence over the range of interest.

3.7 Implementation

In order to enable the automated optimisation and application of CAT to a circuit, the CDI method of Section 3.2 and the device size optimisation algorithm of Section 3.3 were implemented in software. The software tool was written in Ruby and works in conjunction with the Cadence Virtuoso custom IC design tools either through the OCEAN scripting interface or directly with the Spectre simulator. The architecture of this program and its interaction with the simulation tools are illustrated in Figure 3.16. The three main functions, Critical Device Identification, CAT device size optimisation, and post-fabrication characterisation correspond to the three CAT design tools introduced on Figure 2.26. The inputs to the program are a netlist describing the circuit and the appropriate model libraries, the simulation setup (e.g. analysis types and parameters) along

with a set of expressions to calculate circuit performances from the simulation results. Depending on whether the simulation runs through OCEAN or Spectre directly, these will differ slightly. For example, when using OCEAN, the performance calculations can be formulated using expressions from the Cadence ADE waveform calculator, whereas the expressions and mathematical functions available when using Spectre are somewhat more limited in number. The outputs of the program are, depending on which functions are run, a list of critical devices, optimised slice sizes for these critical devices and the circuit simulation results (e.g. Monte Carlo) with or without application of CAT. Monte Carlo sampling and temperature behaviour can be combined or run separately, such that a circuit is optimised for temperature only, parameter variation only or both at the same time.

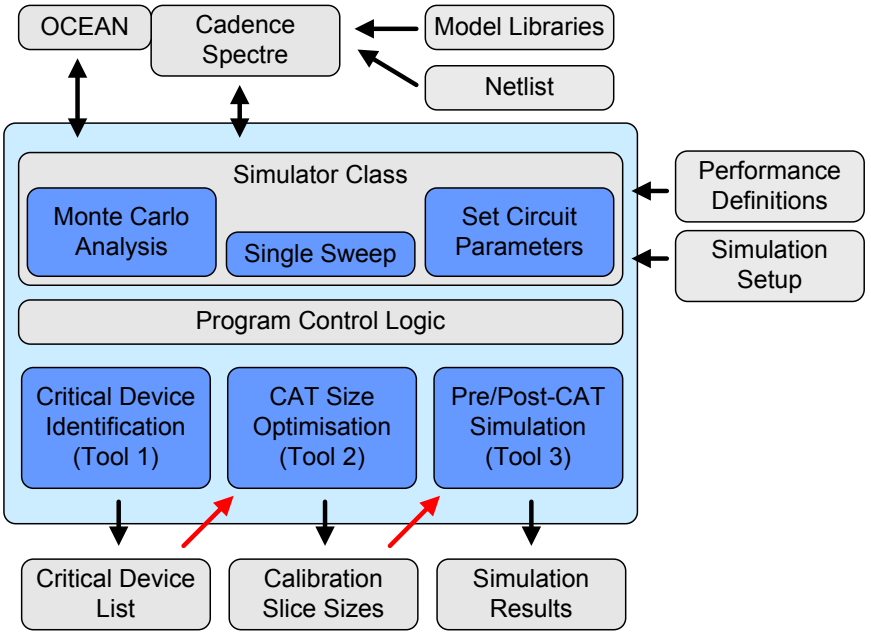


Figure 3.16: Block diagram of the simulation environment.

3.7.1 Simulator Interface

During the course of the CDI and CAT sizing flow, several simulations on the circuit need to be run with different parameters and calculations or decisions based on the simulation results. In order to abstract the actual simulator and performance calculations from the program logic, an interface class was designed that exposes high-level functions such as running a simulation, selecting the simulation type (e.g. single sweep or Monte Carlo) or setting parameters (e.g. temperature or device parameters) to the other parts of the program, whereas interfacing to the simulator is handled internally. Communication with the simulator is through control and results files and, in the case of OCEAN, system pipes to initiate the reading of control files and status indication.

When an instance of the simulator class is created, all transistor widths in the netlist are replaced with parameters and the original values stored internally. This allows changing of device sizes (e.g. for sensitivity analysis or CAT) without the need to read the netlist into the simulator again, which improves overall simulation speed. After this, the file containing simulation setup and performance expressions are read, parsed and stored. If OCEAN is used, two control files are created at this point, which contain the static setup sequence for the model libraries. One of these files links to the nominal parameters model libraries, while the second one links to the Monte Carlo model libraries. The appropriate control file is run whenever a switch from nominal to Monte Carlo simulation or vice versa occurs.

Before a simulation is run, one or more circuit parameters are typically altered by calling the corresponding methods of the simulator class. Any such changes are stored in data structures inside the class. When the method to run a simulation is called, a second simulator control file is generated. This file sets all the design variables to the internally stored values, sets up the simulation type and parameters and calculation of circuit performances. Then, the simulator loads and runs the static and dynamically generated control files in sequence, thereby running the simulation as configured. The results are then written to a results file or returned through the system pipe, checked for validity and stored in a results data structure. The calling part of the program can then retrieve the results from this structure by name and base its calculations or program flow upon them.

3.7.2 Sensitivity Analysis and Critical Device Identification

Sensitivity analysis is performed as described in Section 3.2.3. The width of each transistor is changed by a given percentage in turn, the simulation run and the resulting performances recorded. These results are then combined with the nominal performances according to Equation 3.1 to produce the sensitivity table, which is stored in a file.

When Critical Device Identification is performed, the sensitivity table is read back from the file and the weighted sensitivity table created as per Equations 3.2 and 3.3. There is one critical device for each performance and the resulting critical devices with their corresponding performances are stored in another file. With the list of critical devices established, the process can now proceed to optimally size the CATs.

3.7.3 CAT Device Size Optimisation

As described in Section 3.3.2, device size optimisation of the CATs involves a Monte Carlo simulation where in each iteration the critical devices are sized to best adjust for the MC-induced parameter variation. These ideal adjustments for each device are

recorded over the entire Monte Carlo run and then fed into the sizing algorithm to return the optimised calibration device size for a given number of calibration devices.

For the process of finding the ideally adjusted device sizes in each iteration, the following steps are followed. First, the Monte Carlo iteration is run without any adjustments and the result deviation from the nominal values stored. Then, a limited sensitivity analysis involving only the critical devices is conducted and the ideal adjustment calculated by solving the system of linear Equations 3.4. The critical devices are then adjusted by these calculated values and the Monte Carlo iteration simulated again. If the sensitivities were linear over the required range, the processing in this Monte Carlo iteration would now be complete. However, since the sensitivities are very likely only linear in a small range, the process has to be repeated several times, starting with another limited sensitivity analysis based on the adjusted device sizes. This iterative process is considered finished if the adjusted circuit performances are within an user-defined value (typically 1%) of the nominal performances. The final adjusted device sizes that led to adequately adjusted performances are stored and the process continues with the next Monte Carlo iteration. In the case that the system does not find a solution within a certain number of iteration or diverges beyond reasonable user-specified device size limits, no adjustment information is saved for that Monte Carlo iteration and the process continues with the next Monte Carlo iteration.

3.7.4 Final Simulations

Having optimised calibration device sizes, a final Monte Carlo simulation is run to give an indication of the circuit performance after applying optimally sized CATs. This is done by re-running the same Monte Carlo iterations as in the sizing stage. For each iteration, the previously calculated ideal adjustment is read back from the file and for each critical device the CAT configuration that is closest to the ideal size is chosen. The Monte Carlo iteration is run and the performances saved. The resultant distribution of performances can then be compared to the results without CAT, which concludes the process of Critical Device Identification and device size optimisation.

3.7.5 Parameter Variation

In all of the previous steps, only Monte Carlo sampling was mentioned as a source of parameter variation. However, the software tool also supports to add temperature variation. This can either be done as a pure temperature sweep without any process or mismatch variation or combined with Monte Carlo sampling. In the latter case, a pseudo-random temperature is set in each Monte Carlo iteration, which results in a circuit that is sampled over parameter space and temperature. In any case, all previously described processes for obtaining optimally sized calibration devices remain exactly the same.

3.8 Conclusions

In this chapter, three advancements to the existing work on the CAT were presented. First, Section 3.2 introduced a method for automated critical device identification, which determines which transistors in a circuit are most suitable as calibration devices. In Section 3.3, it was explained how the existing scheme for device size optimisation can be applied to arbitrary devices by basing it on statistical information about adjustments in a Monte Carlo simulation. Section 3.4 conceptually extended the CAT technique from purely static calibration to on-line calibration, where temperature measurements can be used to control optimal CAT configurations during run-time. All of these methods were verified through simulations in Sections 3.5 and 3.6. Finally. Section 3.7 gave a brief overview of the software tool that implements all of the methods in this chapter.

The main aim of the tools introduced in this chapter is a complete automation of the CAT design process, independent of the circuit topology and prior knowledge. The validity of this claim has been demonstrated for two relatively compact linear circuits. However, there are limitations where the proposed processes would not perform well. For example, switching circuits or circuits where devices operate in different modes are not suitable for the proposed method of Critical Device Identification. Furthermore, the required number of simulations, particularly for device size optimisation, means that only relatively small circuits and performances which can be simulated easily are viable candidates for the application of these automated processes. Chapter 5 shows that $\Sigma\Delta$ modulators are a class of circuit which are not suitable for the application of CAT, and that inherently variation-toleration design was in this particular case a more viable route to increase yield and reliability of a system.

These limitations are not unique to the CAT technique, but are common to most of the model-based automated design techniques discussed in Section 2.3.3. General solutions are to reduce complex circuits to simpler building blocks which are treated separately, and to use prior knowledge instead of relying solely on computation. For the CAT, this could mean using machine learning to improve the design space exploration process, or using input from the designer as a starting point. Developing and implementing is suggested in Section 6.3 as possible future work following this thesis.

The practical implementation of CAT has not been considered in depth in this chapter. When describing the circuit of the operational amplifier in Section 3.5, it was pointed out that relative device sizes in analogue circuits are often constrained by the layout strategy. Therefore, arbitrary device sizes, such as the ones of optimised CAT devices, will not be readily applicable to a practical analogue circuit. To incorporate CAT will likely require the designer to round the CAT device sizes to values which are compatible with the existing unit device sizes, and possible to alter the number of unit devices or their size in order to incorporate CAT. A further impact on the layout is the circuitry required to control the calibration devices in a CAT. As will be described in Section

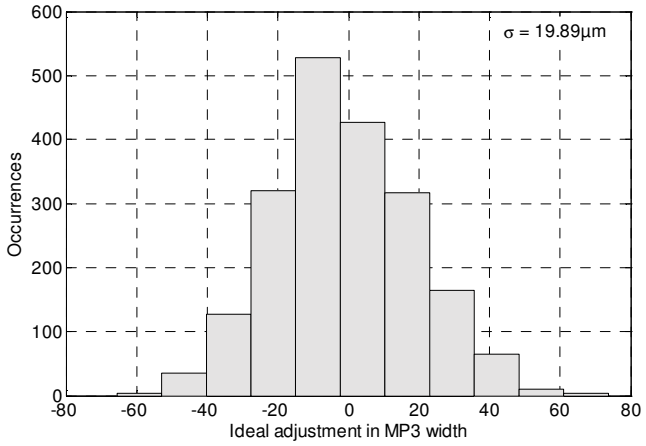
4.4.2, this may consist of a pair of MOSFETs and a shift register. Depending on the circuit, it may be impractical or undesirable to locate these devices at or near one of the critical devices. For example, placing this support circuitry may negatively affect nearby layout patterns (Section 2.3.2), degrade the frequency response of the circuit due to additional capacitive loading, or it may simply not be possible to fit the additional devices in a compact design. All of these points must be taken into considerations for a practical application of CAT and will require a trade-off between circuit performance and efficacy of the CAT.

Aside from design implications, there are also impacts on the fabrication and testing process when using CAT. Like most other offline calibration techniques, the CAT requires circuit performance to be measured after fabrication. This requires a certain amount of time, and may have to be repeated several times to optimise the CAT configuration. Therefore, the time required to find an optimised CAT configuration after fabrication is critical for the economic viability of CAT. It will be shown in Chapter 4 that, depending on certain design decisions, the time required for post-fabrication calibration can approach several minutes for a reasonably complex circuit, which may not be economic for production testing. It is therefore important to consider these aspects during the design stage in order to simplify the post-fabrication measurement process. This is to ensure that CAT remains viable not only during the design stage, but during the entire production cycle of an integrated circuit.

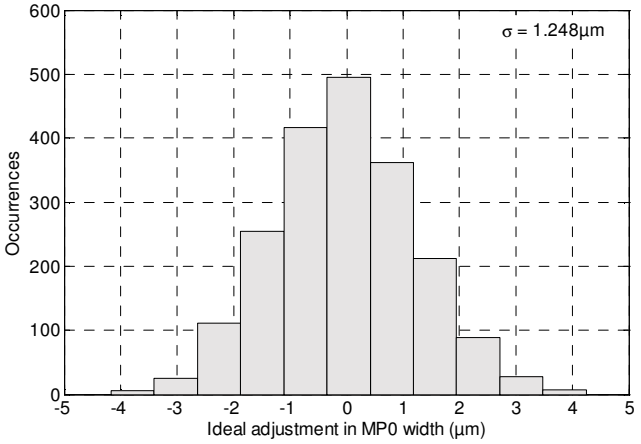
In the following Chapter, CAT is applied to a digital-to-analogue converter to prove its viability and to obtain the first performance measurements of CAT in a realistic circuit. This chapter also discusses some of the effect that including CAT in the layout has one the performance of the analogue parts of the circuit.

Device	Sensitivity (s)			Weighting (w)			Weighted Sensitivity (s')		
	Bandwidth	***Gain***	***CMRR***	***Bandwidth***	***Gain***	***CMRR***	***Bandwidth***	***Gain***	***CMRR***
MP5	0.0280	0.0159	0.0001	0.6361	0.3617	0.0022	0.0178	0.0058	0.0000
MP4	0.0057	0.0045	-0.0006	0.5297	0.4180	0.0523	0.0030	0.0019	0.0000
MP3	-0.0014	**-0.0466**	0.0000	0.0288	**0.9712**	0.0000	0.0000	**-0.0452**	0.0000
MP2	0.0634	-0.0128	0.0005	0.8263	0.1674	0.0063	0.0524	-0.0021	0.0000
MP1	0.0671	-0.0506	0.0003	0.5685	0.4285	0.0029	0.0381	-0.0217	0.0000
MP0	**-0.1287**	0.0996	-0.0007	**0.5620**	0.4351	0.0030	**-0.0723**	0.0433	0.0000
MN9	0.0001	0.0000	0.0000	1.0000	0.0000	0.0000	0.0001	0.0000	0.0000
MN8	0.0021	0.0468	0.0000	0.0422	0.9578	0.0000	-0.0002	0.0448	0.0000
MN7	-0.0048	-0.0019	**0.0970**	0.0467	0.0185	**0.9348**	-0.0007	0.0011	**0.0907**
MN6	-0.0015	0.0019	0.0000	0.4422	0.5577	0.0001	0.0292	-0.0019	0.0000
MN5	0.0370	-0.0095	-0.0004	0.7891	0.2031	0.0078	-0.0345	0.0001	0.0000
MN4	-0.0366	0.0019	0.0003	0.9429	0.0482	0.0089	0.0242	-0.0143	0.0000
MN3	0.0430	-0.0331	0.0002	0.5634	0.4336	0.0030	-0.0350	0.0202	0.0000
MN2	-0.0618	0.0470	-0.0003	0.5663	0.4307	0.0029	0.0005	0.0000	0.0000
MN1	-0.0006	-0.0001	0.0000	0.8699	0.1291	0.0010	0.0001	0.0000	0.0000
MN0	0.0001	-0.0001	0.0000	0.5648	0.4314	0.0038	0.0001	-0.0001	0.0000

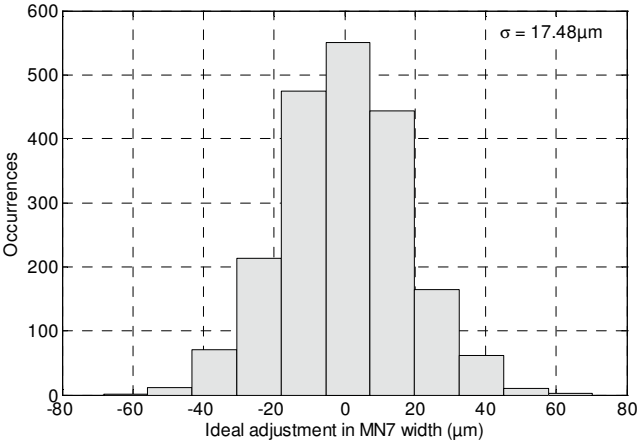
Figure 3.9: Performance sensitivity to variation in transistors.



(a) Transistor MP3



(b) Transistor MP0



(c) Transistor MN7

Figure 3.10: Histograms of ideal relative adjustments.

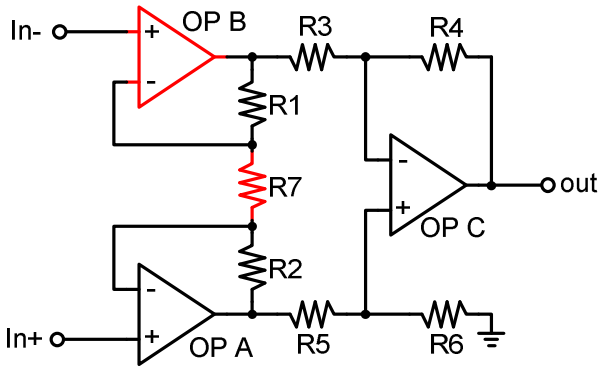


Figure 3.12: Circuit diagram of the instrumentation amplifier.

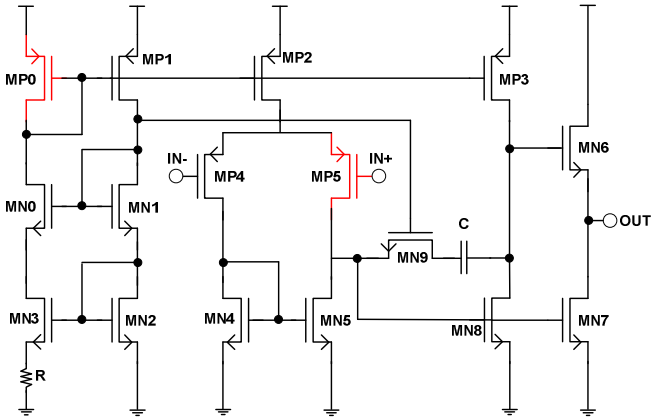


Figure 3.13: Circuit diagram of the operational amplifier.

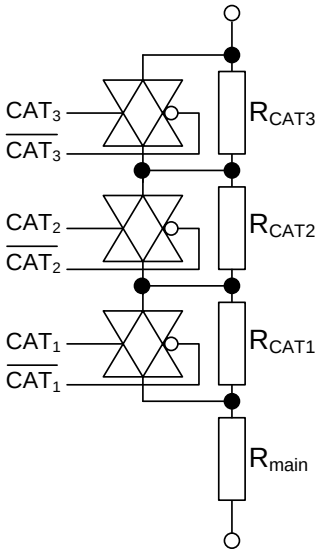


Figure 3.14: CAT-like structure with resistors.

Chapter 4

Application of Configurable Analogue Transistors to Segmented DACs

4.1 Introduction

In the previous chapters, the the theoretical foundations of the CAT technique have been described and tools supporting the automated design flow developed. Furthermore, the possible performance improvements achievable by the CAT have been shown in simulations. In this chapter, CATs are applied to a practical circuit that was later fabricated and tested. This provides insight into the practical aspects of using CAT and proves the concept in reality.

In contrast to the purely analogue sample circuits of Chapter 3, a 14-bit digital-to-analogue converter was chosen as an application circuit for the CAT. This choice was based on several criteria. The most important one is that variation within the unary current cell array in segmented DACs is one of the greatest contributors to performance degradation, which makes it an ideal application for CAT. Furthermore, a DAC is a reasonably complex circuit and it can therefore be proven that CAT is applicable to realistic designs. Finally, segmented DACs are often found in the literature as sample applications for calibration techniques and doing the same for the CAT makes it possible to compare it against other techniques.

This remainder of this chapter is structured as follows. Section 4.2 introduces the basic principle of segmented current-steering DACs. Section 4.3 explains how the CAT can be used to improve INL of such a DAC, which is followed by the conceptual design of the converter building blocks in Section 4.4. Section 4.5 outlines the methodology and equipment used in the measurements of the fabricated chip, which are found in

Section 4.6. The chapter finishes with a discussion of the CAT's ability to improve linearity of current-steering segmented DACs. The work presented in this chapter is the first instance of CAT being applied to a complex fabricated circuit where it is used to improve performance.

4.2 Segmented Current-Steering DACs

4.2.1 Segmented DACs

A common problem when designing high-resolution, high-speed DACs is the matching of binary elements over the full range. For example, a 14-bit DAC based on binary elements would require the MSB element to be accurate to at least 1/16384. Although architectures such as R2R ladder networks remove this matching requirement, they are typically not suitable for high-speed converters. Therefore, high-speed data converters often employ a segmented architecture, which is illustrated in Figure 4.1. Instead of N binary sources (where N is the number of bits), it consists of M binary sources and 2^{N-M} unary sources. The advantage of this arrangement is that the required matching range is reduced to 2^M, which is easier to achieve than a range of 2^N. Common segmentation ratios are in the region of 50% to 70% [83], which is a trade-off between area and matching requirements. This means that a 14-bit DAC is typically divided into a 6 bit binary and a 8 bit unary segment. The unary elements are typically arranged in an array, where rows and columns are driven by thermometer codes from the most significant bits of the input code. For the purpose of illustration, Figure 4.1 shows an 8-bit segmented DAC, where the lower four bits control binary current sources and the higher four bits control the unary current sources. Since the unary array is arranged in a 4x4 matrix, the most significant bits (D6 and D7) selects the row and D4 and D5 select the column in the array. Both row and column decoders are thermometer decoders, such that the output current increases by one unary source value for an increment in D4...D7.

Obvious disadvantages of segmented DACs are the potentially large area of the unary element array and the resulting matching difficulty within the array. Many different approaches exist in the literature that distribute the switching sequence of the unary elements across the array in certain ways to reduce the influence of gradients across the array. Furthermore, the unary elements may be comprised of several smaller elements that are at physically different locations across the array to even out gradients. However, these approaches lead again to further increased area, either in a more complex decoder stage or a larger array. A selection of such techniques was discussed already in Section 2.4.

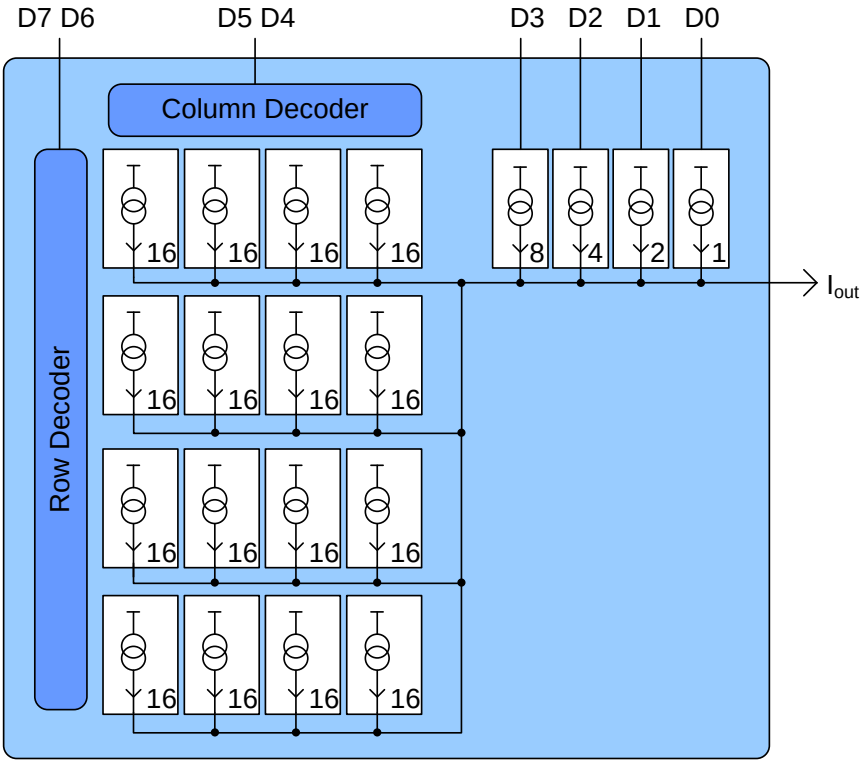


Figure 4.1: Example of a segmented DAC.

4.2.2 Current-Steering DACs

A current-steering DAC is based on the architecture in Figure 4.2. Matched current sources each provide a constant current, which is directed through a switch in one of two output lines. Since the sum of the currents from all sources is constant, the two output lines carry a differential current corresponding to which sources had their switches in the same position. The advantage of this architecture is that the precise current sources ideally carry a constant current, which is merely directed in different paths. This largely eliminates any settling to the end value that would occur if the current sources themselves were turned on and off. Through careful design, the switch can be designed in such a way that the voltage across the current source can be kept constant during the switching process, which leads to an almost ideal switching characteristic and therefore high converter speeds. Further advantages of the current-steering architecture are the inherent differential output and the moderate area requirements compared to, for example, R2R networks.

Due to their simple base circuit, current-steering DACs can be readily implemented in a segmented architecture, which is why this combination is very popular for high-speed, high-resolution converters.

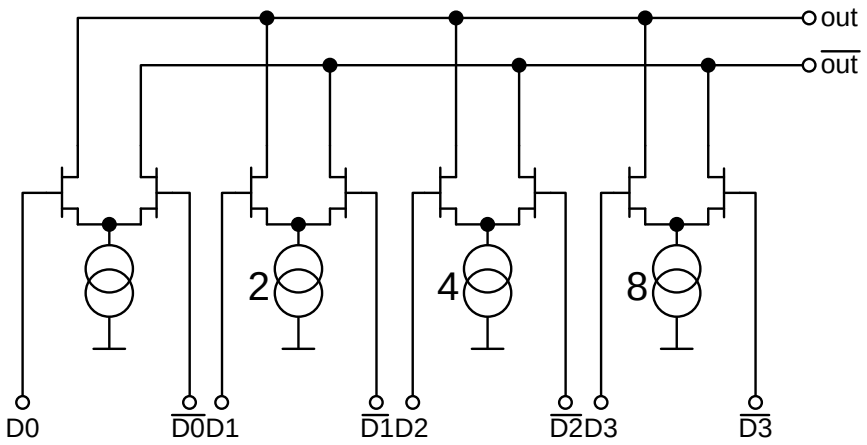


Figure 4.2: Principle of a current-steering DAC.

4.3 Application of CAT to a Segmented Current-Steering DAC

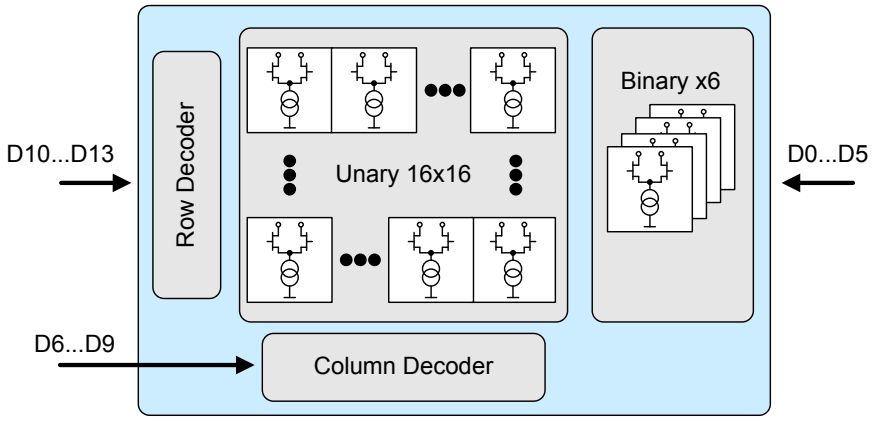


Figure 4.3: Block diagram of the 14 bit segmented DAC.

Figure 4.3 shows the high-level block diagram of the DAC designed in this work. It consists of 256 unary current cells and 6 binary current cells, corresponding to a total of 14 bits. The main source of error are gradients across the unary current cell array which must be kept to less than 1 LSB to achieve 14 bit accuracy. Normally, distributed sources and special switching sequences would be employed to achieve this, as discussed in Section 2.4. However, the aim of this design is to establish whether the CAT alone can improve the linearity of the converter, which is why such techniques are not applied. Figure 4.4 shows a representation of the 16-by-16 current cell array. As the input binary code is incremented, current cells are switched on column by column and row by row, starting at the top left.

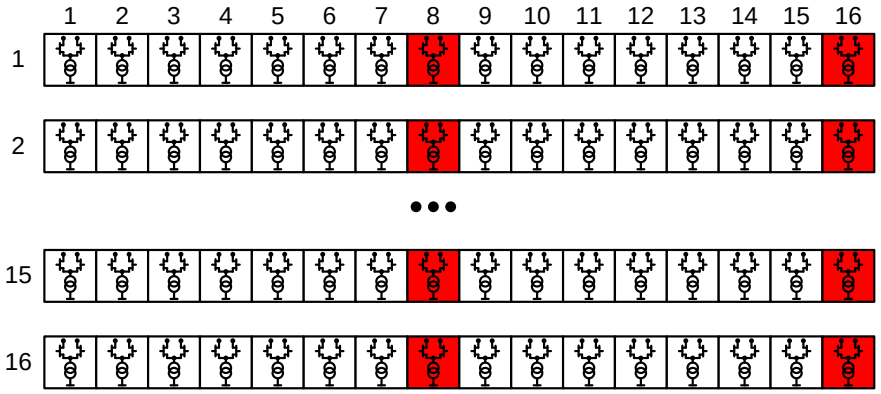


Figure 4.4: Unary current cell array and switching sequence.

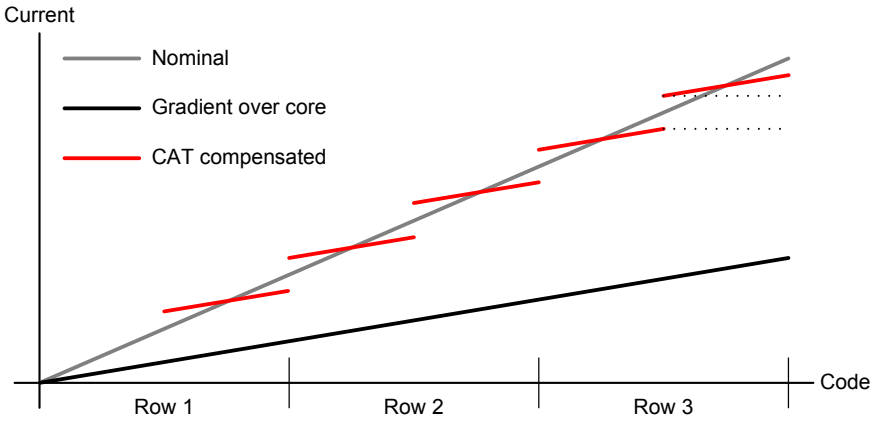


Figure 4.5: CAT-compensated output of DAC.

Suppose the current sources in the array were subject to a gradient, where the current sourced in each cell reduces from left to right and top to bottom, the slope of the output current against the input code would be lower than designed, as depicted in Figure 4.5. By equipping every 8^{th} current cell with CAT, a calibration point is introduced twice in every row. By setting the CATs to appropriate values, the overall slope can be adjusted, as depicted in Figure 4.5. In this case, each CAT-equipped cell adds the indicated amount of current to its nominal value, which is also added to all following currents cells. This has the effect of shifting the entire transfer function to the right of that cell's code up or down by the value introduced by the CAT. However, the gradient from each non-CAT cell to an adjacent non-CAT cell is still the same, which means that the gradient in the transfer function from one point to the next is the same as before, but it is now split into segments that are each centred on the ideal transfer function. Thus, the INL of the current cell array can be improved. In practice, the actual slope of the transfer function is less of a concern than its linearity and CAT would therefore be used to improve linearity rather than the slope. Section 4.6.4 explains how the CAT is used to improve the linearity of the transfer function. It must be noted that using this approach

has in fact a negative effect on DNL. Introducing abrupt changes in output current at CAT-enabled current cells contributes to discontinuities in the transfer function and thus an increased DNL.

The sizing of the CAT in this case must be sufficient to compensate for a current mismatch between to consecutive CAT-enabled current cells. This can achieved by relying on the device mismatch models provided by the process design kit with which the DAC is designed and using the CAT device sizing algorithm on the appropriate sub-circuits.

4.4 Circuit Design

This section gives an overview of the circuit of the DAC and also considers some important details of implementation. A selection of detailed circuit diagrams and cell layouts of the chip can be found in Appendices B and C, respectively.

4.4.1 Current Cell

The most basic building block of the DAC is the current cell. The purpose of the current cell is to sink a constant current from a switchable complementary output. This design uses two different types of current cells for the unary and binary arrays. Apart from the difference in nominal current, these two types only differ in that the unary cells contain a small amount of logic for cell selection within the array which is not needed in the unary current cells. Thus, a current cell of the unary array consists of the following:

- Constant current sink
- Switching transistors
- Data latch and driver
- Cell select logic

Figure 4.6 shows all of these items and external components in a block diagram.

4.4.1.1 Constant Current Sink

The constant current sink used for the current cells is shown in Figure 4.7. It consists of a NMOS transistor (M0), which is supplied with a constant bias voltage at its gate and is therefore sinking a constant current. The second NMOS transistor (M1) acts as a cascode device, which decouples the voltage swing at the output from the drain of M0. This reduces the dependence of the current on the output voltage and therefore allows

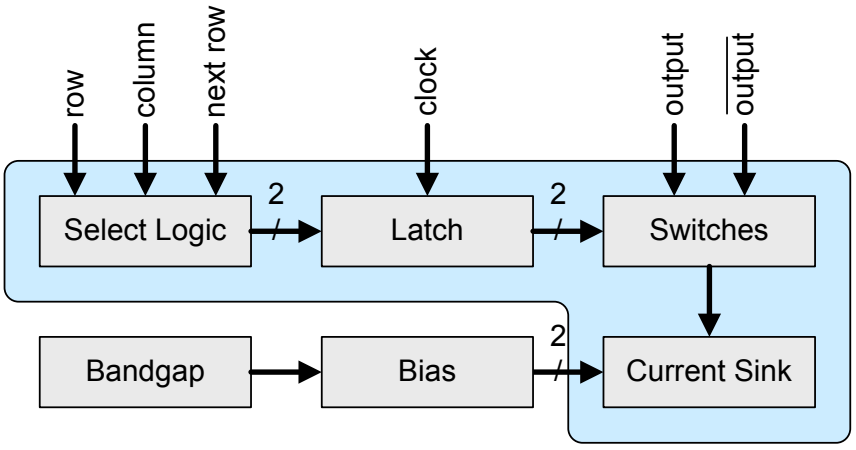


Figure 4.6: Block diagram of a current cell with external bias generation.

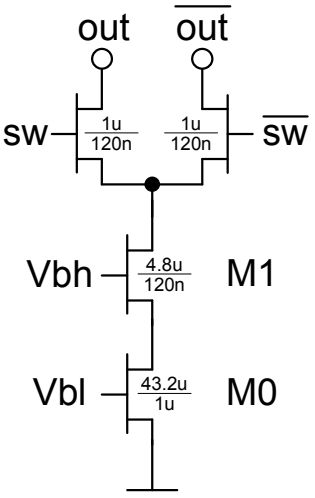


Figure 4.7: Schematic diagram of current sink and switching transistors.

the DAC output to have a voltage swing without affecting the current. The cascode transistor M1 is biased with a different constant gate voltage.

There are two important considerations when sizing the two transistors in the constant current sink. The first is mismatch: The sink current is determined by the W/L ratio of M0, which is ideally the same for all transistors in the unary current cell array. In order to reduce mismatch between individual current cells, transistor M0 should be physically large to reduce effects from fabrication tolerances. It has to be noted that increasing the size of M0 will not compensate for any gradients that can arise over the unary current cell array, which is the purpose of the CAT to compensate. Another consideration is speed and transient performance. The switching transistors are placed at the drain of M1, which means that the drain node of M1 should have as low a capacitance as possible,

which can be achieved by making M1 physically small. This does not have a negative impact on matching since M1 does not greatly contribute to the nominal current.

If the current cell is equipped with a CAT, there are a number of calibration devices in parallel to M0. The gate of each calibration device can either be tied to VSS or the gate of M0 by means of two NMOS pass transistors. These two pass transistors are driven from the complementary outputs of a D-type flip-flop which forms part of the shift register that stores the CAT configurations. The current cell with CAT is described in more detail in Section 4.4.2.

4.4.1.2 Switching Transistors

In order to direct the current from either of the output pins into the constant current sink, a pair of switching transistor is required, as shown in Figure 4.7. The purpose of the switching transistors is to connect one of two complementary current outputs to the current sink. Since only one output should sink current at a time, the gates of the switching transistors are driven with a complementary signal. Although these switching transistors are not critical for current matching and other similar properties, they must be able to switch the current between the outputs fast enough and without causing significant errors in the signal. The latter can be induced if the voltage at the drain node of the cascode transistor changes greatly during switching, which leads to a small change in drain voltage of the main current sink transistor, which in turn leads to a change in current. This can then be seen as glitches in the output current during switching. In order to minimise this change in voltage, the signals at that gates of the switching transistors should be slightly overlapping, as shown in Figure 4.8. In this figure, the solid lines indicate a normal drive signal with no overlap. As can be seen, the sink current changes during the transition, with approximate current and timing values given as an indication. If these complementary signals overlap such that both transistors are on for a brief time during every transition, the current through the current sink will never decrease and therefore the voltage at the drain node of the cascode device will remain constant, as indicated by the dotted signals in Figure 4.8. This largely reduces the glitches in the output current due to voltage changes at any nodes in the current sink. Furthermore, the switching transistors need to be able to switch quickly, even with the parasitic capacitance at their common source node. This requires the transistors to be relatively long, which results in higher transconductance and therefore higher switching speeds.

4.4.1.3 Data Latch

All current cells in the DAC must be synchronised to a common clock signal. In order to achieve this, the signal from the decoders or data inputs to the current cells is latched

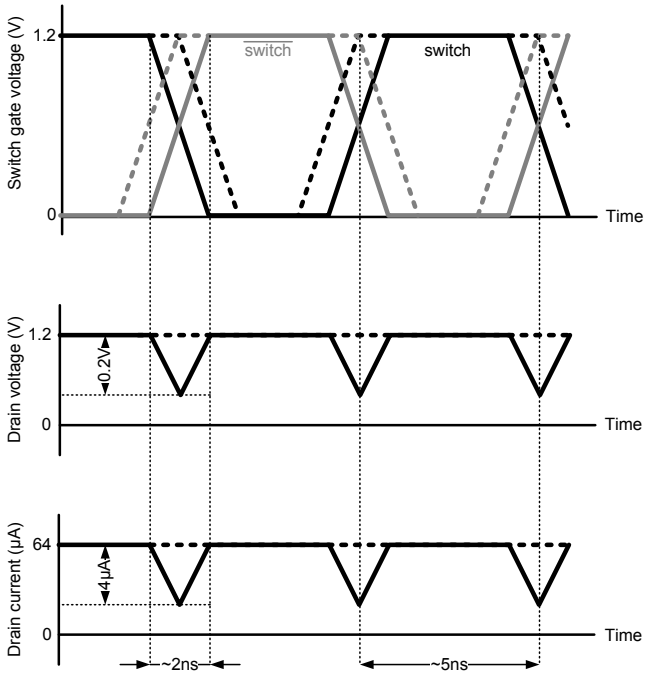


Figure 4.8: Switching waveforms and resultant change in current.

to the switching transistors in each cell. This latch must be sufficiently fast to cope with the required sampling rate and must also be able to drive the relatively high capacitive load of the switching transistor gates. Furthermore, as mentioned previously, the output of the latch must also have a biased crossover point such that both switching transistors are on during a switch-over. The circuit of the core latch is shown in Figure 4.9. It consists of a basic two-inverter structure that stores a state and complementary gated data inputs. If the clock signal is high, the latch is transparent. The data inputs of the latch come either directly from data inputs of the DAC in the case of the binary current cells or from the cell-select logic in the case of the unary current cells. Due to the stacked NMOS transistors, the crossover point of this latch is biased low, which is in opposition of what is required in the current cell. In the DAC, the outputs of the latch are buffered by inverters which also move the crossover point towards the high level. The exact crossover behaviour can then be tuned by sizing the transistors in the inverters appropriately. Furthermore, the inverters help drive the switching transistors at the required speeds.

4.4.1.4 Cell Select Logic

In the unary current cell array, thermometer encoders select the row and column up to which the current cells should be turned on. Ideally, each current cell could be simply enabled when both its row and column are selected. However, as is illustrated in Figure

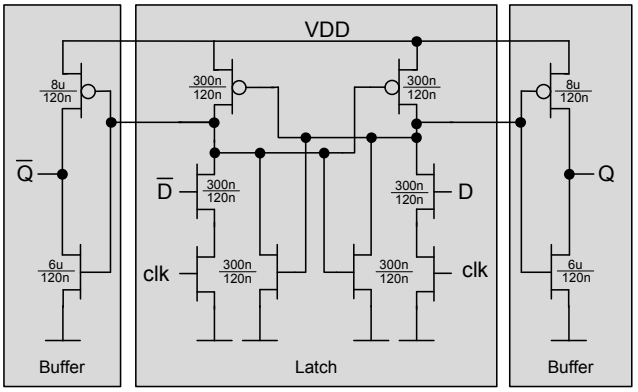


Figure 4.9: Schematic diagram of the data latch.

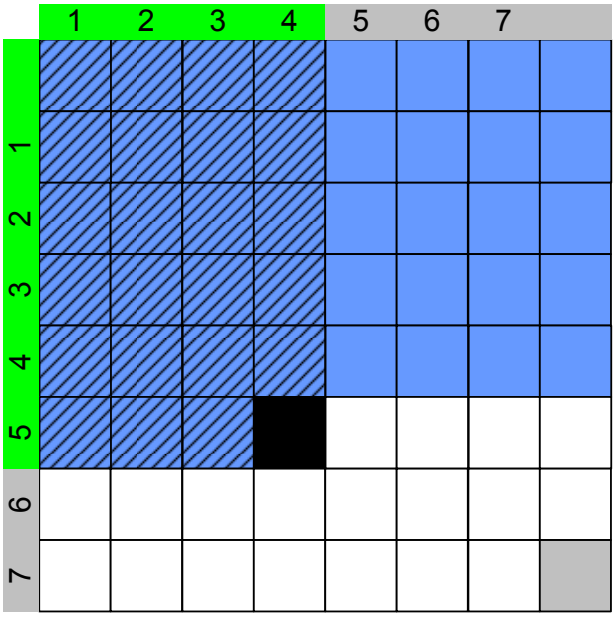


Figure 4.10: Current cell activation from row and column select signals.

4.10, this would only work for cells in the highest selected row. In this example, all cells up to the cell in row 5 and column 4 should be selected, which are shaded blue. If each cell actives when its row and column inputs are asserted, only the cells with a hatched shading would be selected. Therefore, in all but the last row all cells must be turned on even though their column may not be selected. This requires some additional logic in each current cell that detects whether the following row is enabled and, if so, enables the current cell regardless of its own row or column signal. The truth table for this logic in shown in Table 4.1 and the circuit implementation in Figure 4.11. In the last row in the unary current cell array, the next row input is simply tied tied low. Likewise, the row signal for the first row is tied high because the first row is always selected. This allows the use of the same 15-line thermometer decoder for rows and columns. Note that the

row	column	next row	cell enable
0	X	0	0
X	0	0	0
1	1	X	1
X	X	1	1

Table 4.1: Truth table of cell select logic

last cell in each row is never turned on with a column signal, but only through the next row signal: In the 16-by-16 current cell array, there are four input bits for each rows and columns that are decoded into a 15-line thermometer signal. The row lines drive current cell rows 2-16, with row 1 always selected, while the column lines drive columns 1-15. The 16^{th} current cell in each row is turned as soon as the next row is selected, even with no columns selected at all. This means that the 256^{th} current cell in the array is not used at all. Current cells in the binary current cell array do not contain this cell select logic. Instead, the corresponding input data bits are routed directly to the input of the data latch.

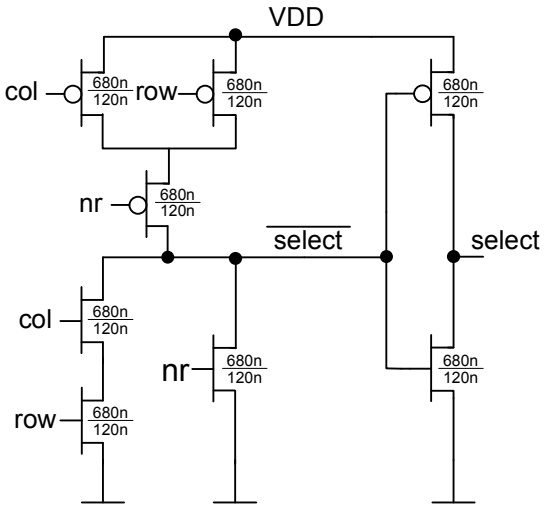


Figure 4.11: Circuit diagram of the cell select logic.

4.4.2 CAT-Enabled Current Cell

As illustrated in Figure 4.4, every 8^{th} current cell in the unary array is equipped with CAT. The design of such a CAT-enabled current cell is shown in Figure 4.12. Fundamentally, it still consists of the constant current sink M0/M1 and the output switches, but there are additional transistors MS_0 - MS_2 in parallel with M0, which are the CAT calibration slices. Each of these calibration slices is controlled by two n-channel transistors, MG_{0A} and MG_{0B}. When MG_{0B} is active, the gate of MS_0 is grounded, thereby disabling it. When MG_{0A} is active, the gate of MS_0 is connected to Vbl, thereby effectively putting

it in parallel with M0. Since M0 and MS_0 have the same channel length, the net result of this is the same as if the width of M0 were increased, leading to a larger sink current in this cell. MG_{0A} and MG_{0B} are controlled by the complementary outputs of a D-type flip-flop, which stores configuration of its associated slice. All CAT controlling flip-flops on the chip are connected in a scan chain such that that only clock, data in and data out signals are needed to configure the CATs.

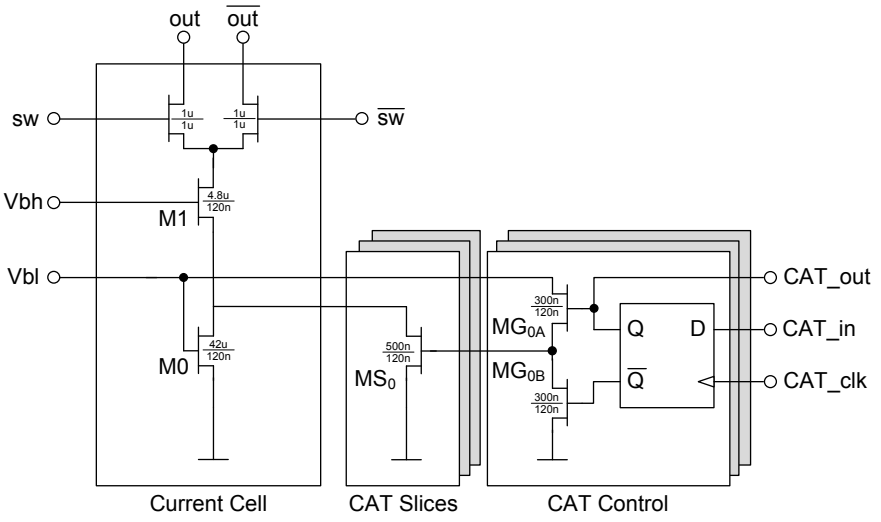


Figure 4.12: Schematic of a current cell with CAT

The calibration slices were sized using the device size optimisation tool of Section 3.3, which in turn used Monte Carlo simulations of the basic current cell (Figure 4.7). For three CAT slices, the optimised slice currents were determined to be $570nA$, $1.14\mu A$ and $2.28\mu A$, for a maximum total CAT current of $4.00\mu A$. In order to allow the CATs to adjust the cell's current around its nominal value, the base current of the cell was lowered from $64\mu A$ to $62\mu A$.

4.4.3 Thermometer Decoders

As mentioned previously, rows and columns in the unary current cell array are selected by means of thermometer decoders. Each decoder converts four input data bits to 15 output lines, as illustrated in Figure 4.13. For an input of 0000, the output is all zeroes, for an input of 0001, the least significant output bit is one, for an input of 0010 the two least significant outputs are one and so on until all output bits are one for an input code of 1111. Although the output of the thermometer decoder is latched in each current cell, it is still desirable to equalise the propagation time for all outputs. To facilitate this, the decoder has been designed as a cascade of 2-to-3 thermometer decoders, whose logic diagram is shown in Figure 4.14. Each of the gates has two additional inputs: assert and inhibit. The purpose of these is to override the normal decoder function and either set

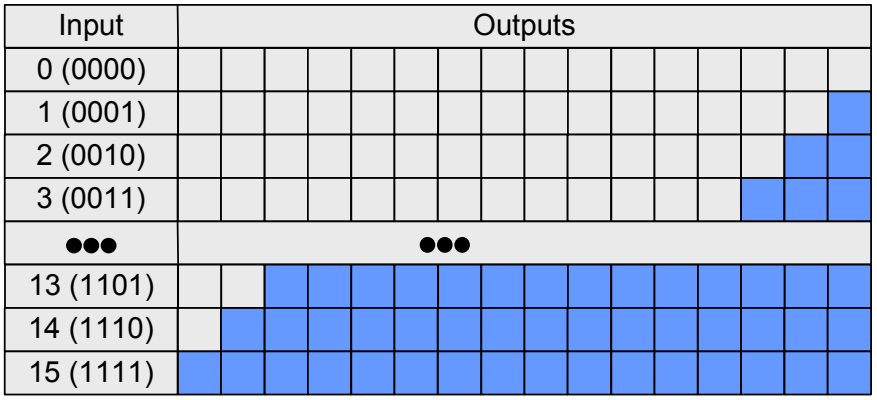


Figure 4.13: Operation of a 4-to-15 thermometer decoder.

or clear all output bits regardless of the input code. If both assert and inhibit inputs are asserted, the assert input has priority. Figure 4.15 shows the block diagram of the 4-to-15 thermometer decoder built from 2-to-3 thermometer decoders. The first decoder on the top left works on the most significant input bits and its output is used to either enable or disable the other four decoders which work on the two least significant input bits. For example, for an input code of 0011 all but the first LSB decoders are disabled and the output of the first LSB decoder will be 111. If the input code increments to 0100, the output of the first LSB decoder would return to 000. However, now the first output of the MSB decoder is high, which keeps all output bits of the first LSB decoder asserted and enables the second LSB decoder. If the input code increments further to 0101, the first output of the second LSB decoder turns on. The output bits that are derived directly from the MSB decoder are passed through a line of inverters to bring their propagation delay in line with the other outputs. Since the outputs of the thermometer decoders are routed across the large unary current cell array and need to drive several inputs each, an additional output stage of is added to increase drive strength.

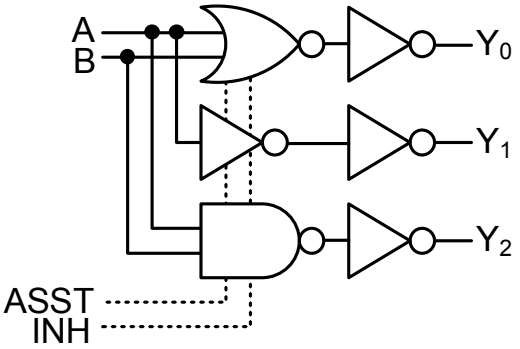


Figure 4.14: Logic diagram of the 2-to-3 thermometer decoder.

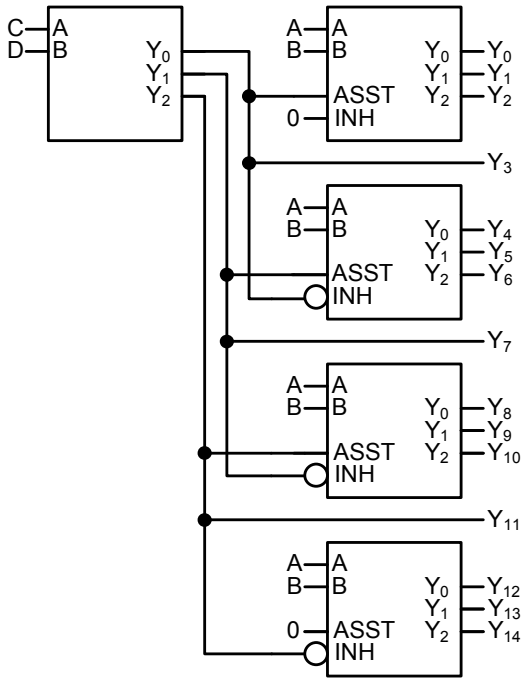


Figure 4.15: Block diagram of the 4-to-15 thermometer decoder.

4.4.4 Biasing Circuit

As was described previously, the current cell needs two reference voltages for the main current sink and cascode transistors. Eight current cells each share a bias generator, which generates these bias voltages from a $10\mu A$ reference current. The schematic diagram for such a bias generator is shown in Figure 4.16. The transistor sizes in the bias generator are designed such to be a ratio of the current sink transistors, which ensures that the sink current will track the reference current over process parameters and temperature. The bias voltage for the sink transistor (Vbl) is generated by M0 and M1 while the bias voltage for the cascode device (Vbh) is generated through the resistor. The advantage of this circuit over a regular cascode current mirror is that Vbh is independent of the gate-source voltage of M1, which allows the drain node of the cascode device in the current sink to go to lower voltages. In the DAC, there is one bias generator for every 8 current cells in the unary current cell array and one bias generator for the 6 binary current cells for a total of 33 bias generators.

The $10\mu A$ reference current for each bias circuit is generated from a bandgap reference. The core of the bandgap reference was designed by Dr Li Ke and current outputs were added as required. The bandgap reference features 33 $10\mu A$ outputs, 32 for the unary current cell array and one for the binary array. Furthermore, it also features a 650mV reference voltage output. This reference voltage is routed directly to a pin on the chip to facilitate measurement of the internal bias condition.

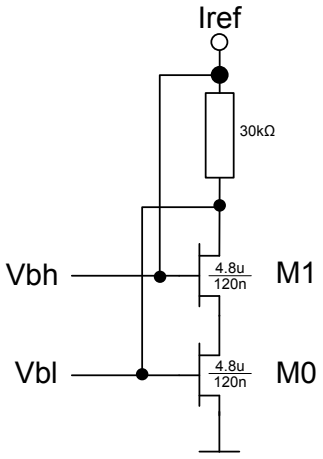


Figure 4.16: Bias generator for the current cells.

4.5 Measurement Methodology

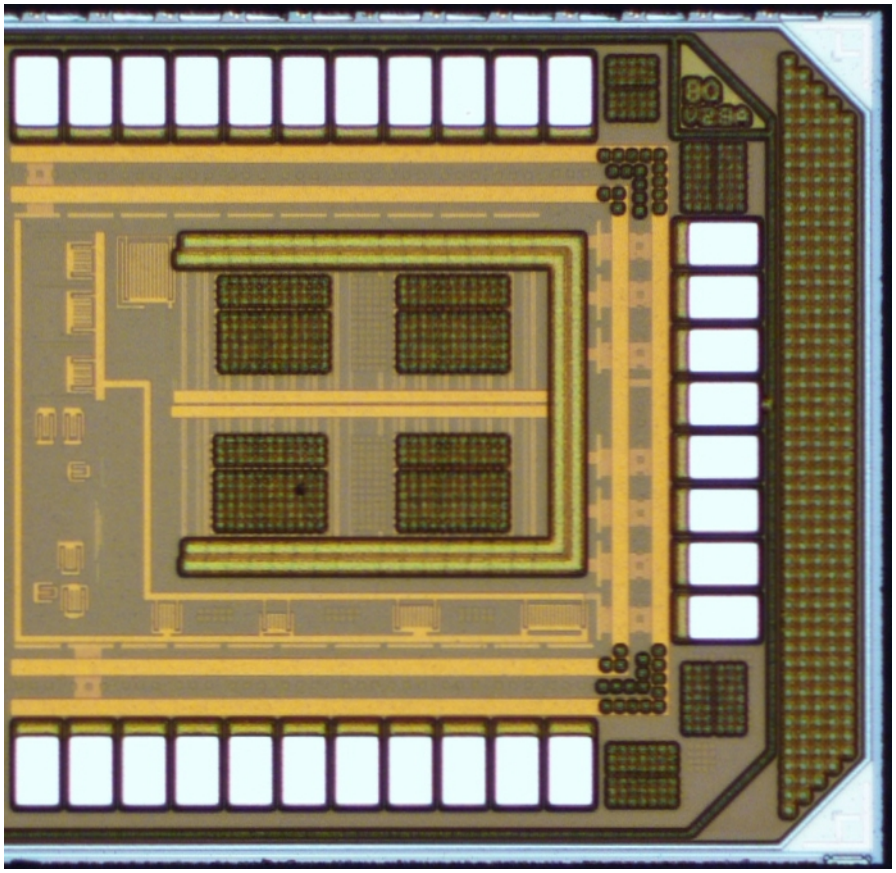


Figure 4.17: Die photograph of the fabricated chip.

The DAC described in Section 4.4 was fabricated on the IBM 8RF 130nm process. A photograph of the fabricated die is shown in Figure 4.17, which corresponds to the top-level

layout view of Figure C.1. The chips were then tested to evaluate the performance with and without the CATs. To facilitate this, a custom test board was designed which allows to evaluate both static and dynamic performance of the DAC and quantify the effects that CATs have on the DAC performance. The primary design goal of this board was to provide a way to control the CAT configuration and apply low-speed data, but the board also provides the option to deal with high-speed signals and temperature-dependent measurements. The board consists of power supplies, a controller and amplifiers, which are described in the following subsection.

4.5.1 Test Board

An overview of the test board is given in Figure 4.18. The controller is an mbed LPC1768 ARM Cortex M3 microcontroller board, which communicates with a PC over USB. The purpose of the controller is to operate the CAT scan chain to load CAT settings to the chip under test, as well as providing static or low-speed data, primarily intended for basic testing and INL measurements. The controller communicates with a PC via a USB virtual serial port using a custom protocol, which is described in Appendix E. Furthermore, the controller can also read the on-chip bandgap reference voltage to facilitate schemes like online calibration. All lines between the controller and the chip under test pass through level shifters to convert between the 1.2V signals required by the chip and the 3.3V signals of the controller. In addition, the parallel data lines and clock are passed through standard 2.54mm headers over a ribbon cable. If the cable is removed, an alternate data source such as an FPGA can be used to provide high-speed signals. From the "incoming" header onwards, data and clock traces are matched length. Each chip under test is mounted on a separate break-out board which plugs into the base board to allow uncomplicated swapping of the chip under test.

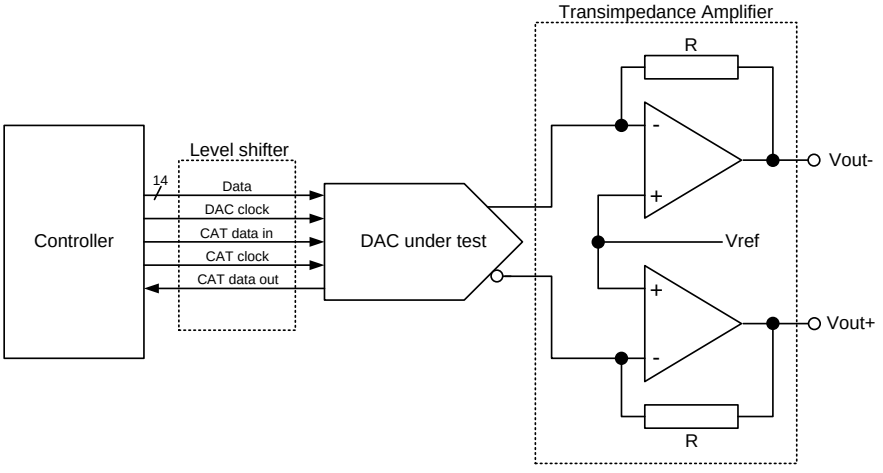


Figure 4.18: Signal path architecture of the test board.

Power is provided to the board through USB or an external plug-in power supply at 5V nominal. Since noise coupled in through USB proved to be a problem, an external power supply was used for all measurements. Three on-board regulators provide 3.3V (for the level shifters), 1.2V (for the chip) and an adjustable reference voltage used in the transimpedance amplifiers. In order to reduce noise on the 1.2V supply and, more importantly, the reference voltage, the boards were later modified to use a 1.5V primary cell as a reference source. This was by far the most economic way to obtain an extremely low-noise reference with reasonable stability. The distribution of power and reference voltages is shown in Figure 4.19.

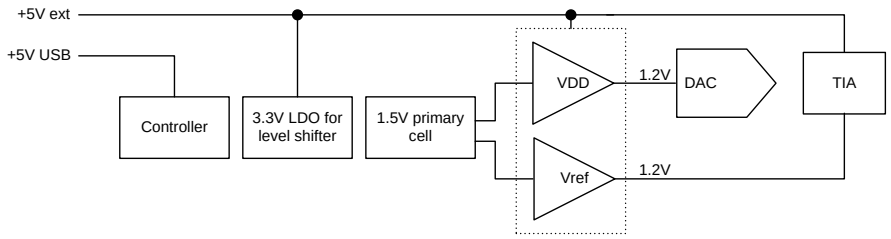


Figure 4.19: Supply and reference distribution.

The two differential output signals from the chip are passed to two separate transimpedance amplifiers, as illustrated again in Figure 4.18. Each of the amplifiers converts the current output of the chip to a proportional voltage with a nominal gain of $47V/A$ for each channel. Simultaneously, the output node of the chip is held at a constant voltage (reference voltage). This is a critical part of the system, because a change in the output node voltage will affect the absolute value of the current provided by the chip. The differential output voltage between the two transimpedence amplifiers is directly proportional to the differential output current of the DAC and can either be measured directly by a non-ground-referenced multimeter or converted to a single-ended signal for measurement with a ground-referenced oscilloscope. Conversion to a single-ended signal is done by a high-speed opamp difference amplifier and its output is provided at a BNC connector.

Since most measurements were carried out using a multimeter with a floating common terminal, both single-ended and differential signals were measured directly after the transimpedance amplifier, skipping the differential to single ended conversion. To minimise disruption created by changing probes between singe-ended and differential configurations, a relay was later added for this purpose. This addition also facilitates computer-controlled switching between single-ended and differential measurements. Another later modification was the addition of a lower-noise transimpedance amplifier, which is more closely integrated with the low noise reference generator.

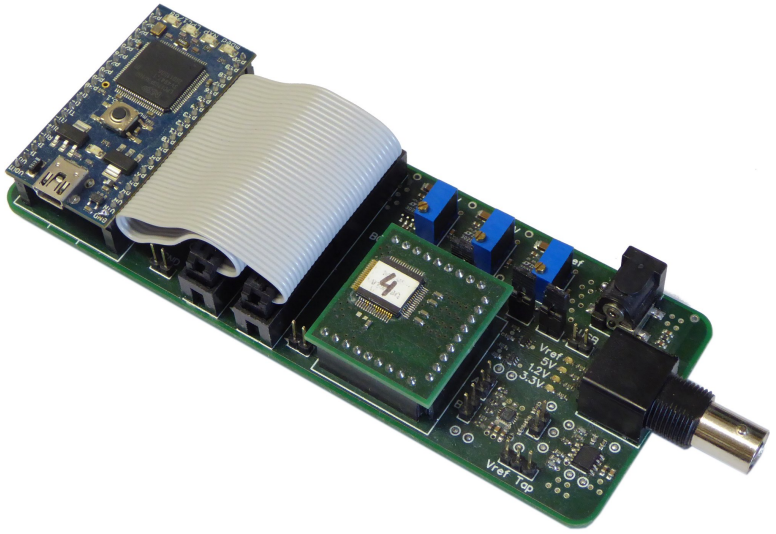


Figure 4.20: Photograph of the test board.

Figure 4.20 shows a photograph of the test board. The PCB contains the power supplies, controller and device under test, while an external breadboard contains the transimpedance amplifiers and reference amplifiers. In addition, the breadboard also contains the relay which is used to switch between singe-ended and differential outputs under computer control to facilitate a greater degree of automation of certain measurements.

4.5.2 Test Equipment and Procedure

In addition to the test board described above, the test setup consists of a high-performance digital multimeter and a PC. The multimeter is an Agilent 34410A 6.5 digit multimeter, which for all measurements in this chapter measures either a single-ended or the differential output voltage of the trans-impedance amplifier. The 34410A is connected to the PC via Ethernet and is controlled by SCPI commands. The purpose of the PC is to control the multimeter and test board, to collect data from the multimeter and perform the necessary calculations. All of these functions are implemented in a number of Matlab routines. There are three fundamental I/O functions around which all following measurements are built: measure voltage through the multimeter, output a given code to the DAC and set CAT configuration.

There are two measurements for which the PC was not used to interface to the multimeter, but which were instead measured manually. The first were some very initial tests, in which the output current of the DAC was measured directly using the multimeter. The second were the noise measurements, which for which a high-speed oscilloscope was used. Apart from that, all measurements in this chapter were done in an automated manner.

As will be explained in Section 4.6.2, due to significant noise present caused by the on-chip bandgap reference, most measurements required averaging over an appropriate number of samples. However, this introduces another potential pitfall: Repeating a measurement often enough can easily lead to a total measurement time of several hours. On such time scales, component drift and temperature changes can become significant enough to affect the measurement. In order to minimise such effects, the sequence of steps within each measurement had to be planned carefully. For example, when obtaining the transfer function of the DAC, the input code is stepped from zero to the maximum value and at each step, the output is measured. In this case, it is necessary to measure one transfer function after another and take the average at the end. This way, each individual transfer function is recorded in a time short enough for drift to not have a significant effect. This ensures that long-term drift does not affect the relative relationship of values within the final averaged transfer function. If instead the measurement at each step were repeated and averaged before going to the next step, thus effectively slowly recording one transfer function, drift over time could have an effect on the measured transfer function as the circuit parameters change while the input code is slowly stepped upwards. Similar considerations must also be made when CATs are used.

4.5.3 Computer software

The test hardware described thus far in this section is complemented by computer software to automate measurements and implement the final step of the automated CAT design flow: Post-fabrication calibration. The software consists of a number of Matlab scripts and functions, whose purpose and general function is outlined in the following subsections and which are available in full in Appendix D.

4.5.3.1 Application of CAT to a Transfer Function

The first function of the software is to modify a measured transfer function by simulating the application of CAT. This is required to estimate the transfer function that results when CAT are applied to the chip and more importantly, it forms the foundation of the post-fabrication optimisation algorithm.

The principle of the tool is the same as illustrated in Figure 4.33, which is discussed in more detail in Section 4.6.4. Given a transfer function and a CAT at a certain code point along the transfer function, the modified transfer function is calculated by adding the CAT's current to all subsequent points of the transfer function. This is a direct implementation of the real-world function of CAT in the unary current source array. Since many of the measurements in Section 4.6 are done on the differential output, an extended version of this function exists for differential transfer functions, where the CAT current is added to all points after the CAT position, and subtracted from all

points before it. An example where this function is used directly is in Figure 4.39, where it was used to calculate the expected transfer function after applying a given CAT configuration. However, the main purpose of this tool is in the process of optimising the CAT configuration for a given transfer function, which is explained below.

4.5.3.2 Optimisation of CAT Configuration

The purpose of the second software component is to optimise the CAT configuration for a given transfer function. This process is schematically illustrated in Figure 4.21. For a given transfer function, a CAT configuration can be applied and the resultant transfer function computed using the tool described in Section 4.5.3.1. Then, the INL of the modified transfer function can be calculated. This INL can then be used as a metric for an optimisation algorithm, which modifies the initial CAT configuration until a minimum in INL is reached. This is an instance of a single-objective multi-variate optimisation problem, which was discussed in Section 2.3.3.3. Although the design space of this particular problem with 32 CATs having 3 configuration bits each might seem small at first glance, a full design space exploration with 2^{96} possible combinations is nowhere near feasible. Therefore, a conventional numeric solver is used. The solution of choice is a Levenberg-Marquardt algorithm [82], as it is reasonably efficient and provides good solutions for this particular problem. Using the implementation of the Levenberg-Marquardt algorithm that is supplied with Matlab, optimising the CAT configurations for a given transfer function takes in the order of one second on an Intel®Core™i7 processor.

Once this software tool is applied to a measured transfer function, the resultant optimised CAT configurations can be transferred to the chip, where the adjusted transfer function can then be measured. It is important to note at this point that there will be a discrepancy between the calculated and measured transfer functions, as will be seen in Section 4.6.4. The main reason for this is that the CATs in the real chip are not perfect and show variation, as shown in Table 4.3, while the software tool assumes uniform and ideal CATs. This behaviour could be improved by measuring all individual CAT slices before optimisation and using these measurements when calculating the adjusted transfer functions.

4.6 Measurement Results

After the chip had been fabricated and the test board assembled, some basic functionality tests were conducted. These included determining whether the on-chip bandgap reference was operational, testing the CAT scan chain and basic DAC operation.

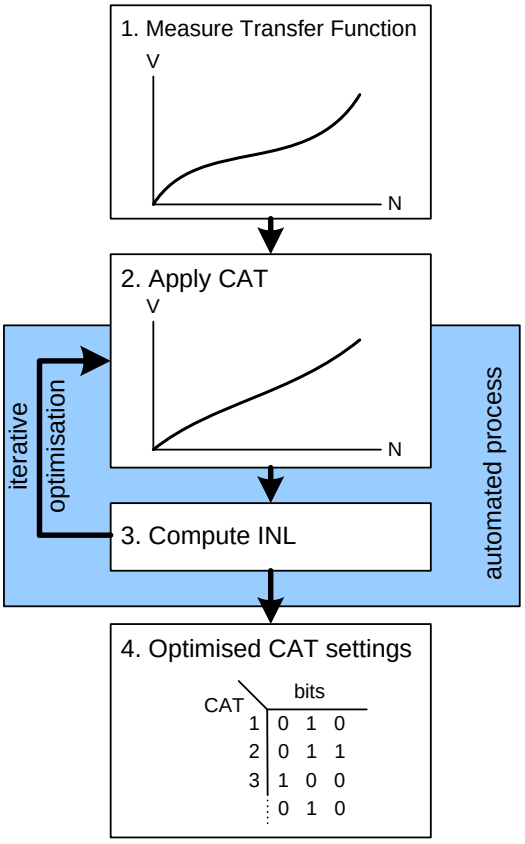


Figure 4.21: Optimisation of CAT configuration

Once basic operation of the DAC and CATs was confirmed, further measurements were carried out to characterise the DAC in more detail, especially with respect to INL. Then, the CATs were characterised and finally the improvement on the converter's INL by applying CATs was measured. All of these tests were carried out on three different chips.

4.6.1 Functional Testing

The very first measurements were to establish whether the DAC is functional at all. For this, the output currents of the DAC were measured for zero, full-range and certain intermediate codes to test the unary and binary sections. Since these measurements were only intended to arrive at a pass-fail result, the currents were measured with the multimeter in current mode. This would of course not be acceptable for measurements requiring any level of accuracy since the relatively high internal resistance of the ampere meter in the lower ranges affects the output characteristics of the DAC. As a result of these initial tests, it was established that the DAC design works in principle. However, a number of chips appear to have faults in the thermometer decoders and the logic within the current cells, as shown by the transfer function in Figure 4.22. It can be seen from

the "sawtooth" pattern repeating every 16 codes that the lower four bits of the binary current cell array work as expected. However, the upper two bits of the binary cell array seem to never activate at all. Furthermore, there is also a very obvious problem with either the thermometer encoders or the logic in some unary cells. For instance, the first unary cell at code 64 activates correctly, but then de-activates at code 128, without the second cell becoming live either. At code 512, a number of cells turn on, but not enough to give the correct current. Since only certain chips are affected by this problem, this is a fabrication fault rather than inherent in the design. Therefore, testing of DAC and CAT performance in the remaining chips is not affected by this problem. It is difficult to say whether the problem lies with the logic in the current cells or the thermometer encoders. However, the cause of these faults is likely due to failure in single contacts in narrow metal traces, as this is a somewhat common problem. The chips with these faults were naturally not usable for further measurements. Table 4.2 gives a breakdown of the final usages of the fabricated chips.

Number of chips	Description
2	Used for prototyping
3	Fully working
3	Defective
2	Spares
10	Fabricated

Table 4.2: Breakdown of fabricated chips

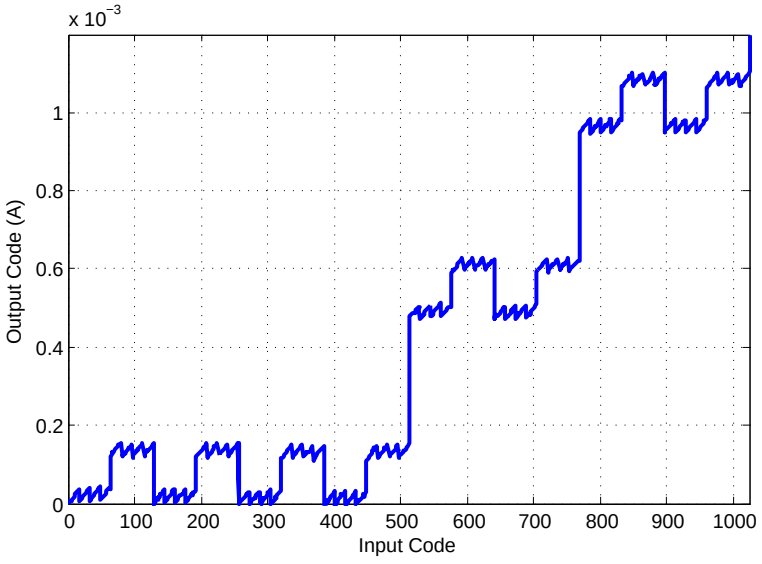


Figure 4.22: Partial transfer function of a defective chip.

Next followed a set of more accurate measurements of the DAC transfer function. For this, the chips to be tested were fitted in the test board and the single-ended output

voltage was measured after the transimpedance amplifier. Figure 4.23 shows the single-ended transfer functions of three different chips. All of these transfer functions are qualitatively consistent with what would be expected from a DAC and the zero and full-scale readings correspond with the designed full-range current of $16.384mA$ for each channel. For this graph, the zero-code offset of each transfer function has been moved to zero current. The differences between the individual transfer functions are due to chip-to-chip variation and can on this scale only be discerned as differences in gain.

While the transfer functions in Figure 4.23 established that the DAC works as expected on a large scale and gives an idea about its linearity, they do not show the DAC characteristics at the lower bit levels. For this, consider Figure 4.24, which is based on the same data as Figure 4.23, but only shows the first 192 codes. This means that in this graph all codes of the binary current cell array are present three times and the first two unary current cells are covered. It is immediately apparent that there is much greater variation between the chips on this scale. For example, chips 1 and 4 both exhibit discontinuities at codes that are multiples of 32, indicating a mismatch between the 6^{th} and the first five binary current cells. In chip 3, discontinuities at these codes are significantly smaller. One common property of all chips are discontinuities at codes that are multiples of 64. Since at each multiple of 64 a new unary current source is enabled, this indicates a general mismatch between the binary current cells and the unary cells.

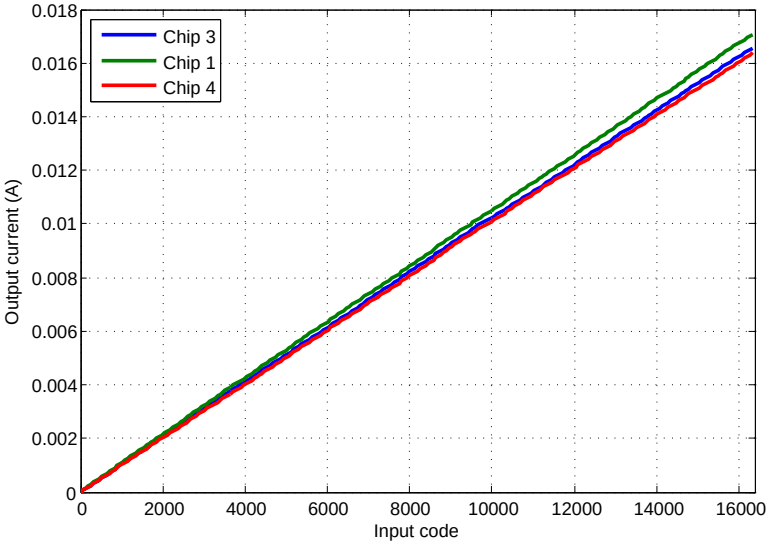


Figure 4.23: Transfer functions for three chips.

Another interesting measurement is to compare the currents of the individual unary current cells. For this, the output current was measured at codes that are multiples of 64. Since the output current is the sum of all current cells up to a particular code, the difference in current between two successive codes is the current of that particular current cell. Figures 4.25, 4.26 and 4.27 show the absolute currents in each cell for three different chips, arranged in a matrix as they are physically on the chip. It is worth noting

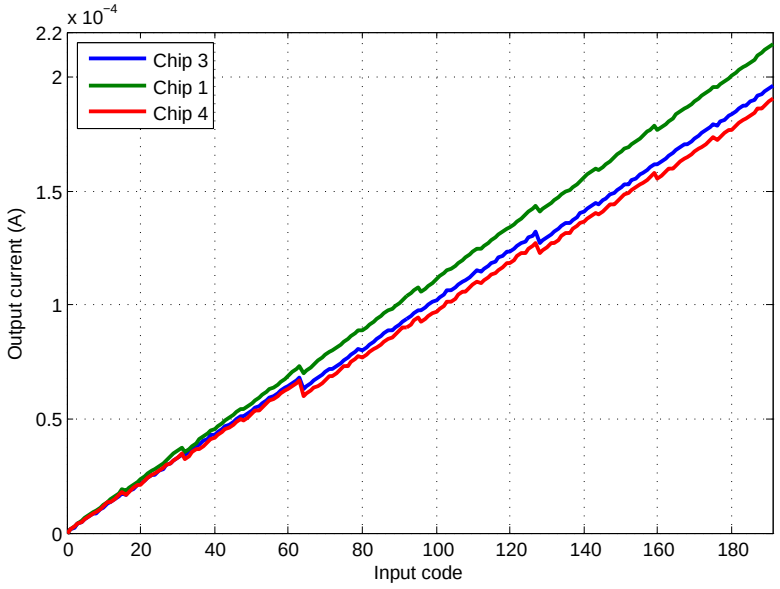


Figure 4.24: Detail of transfer functions for first 192 codes.

that the 256$^{\text{th}}$ cell (column 16, row 16) is never activated on the chip and has been set to the mean current in the graph.

There are several features that become immediately apparent and which can be linked to layout features and CAT. For instance, in all cases the currents of cells in columns 8 and 16 are noticeably lower than neighbouring cells. This is because these cells are equipped with CAT and all CAT slices have been disabled for this measurement, leading to lower overall currents in these cells. There are also horizontal features in rows 9 and 10, which are especially apparent in chips 1 and 4. Also noticeable are horizontal features, where cells in one half of each row show similar deviations from the average current. These are likely due to the fact that each set of consecutive 8 cells shares a bandgap output and bias voltage generator. Thus, any error in the reference current or mismatch in the bias generator affects these 8 cells in a similar manner. The increased currents in the top-left cells of chip 1 cannot be readily explained by layout features or design. However, thermal effects from the adjacent bandgap reference can provide a plausible explanation.

Since the aim of this work is to apply the CAT to reduce the effects of these variations, their cause is only of secondary interest. Thus, there is no further investigation into the exact causes of these variation patterns, apart from the brief suggestions of plausible mechanisms above.

4.6.2 Initial Findings

During these tests, two issues with the chip became apparent, one of which has an effect on subsequent measurements. First, the bandgap reference tends to become unstable

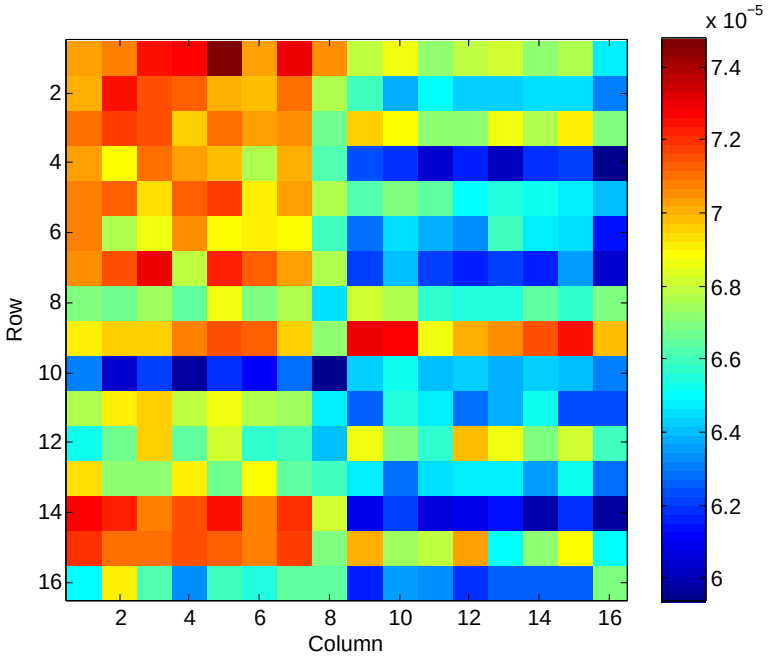


Figure 4.25: Map of unary current cells of chip 1 (in A).

at supply voltages of around 1.1V and higher. This instability causes high-frequency oscillations on the reference outputs and subsequently on the DAC output currents. Since the DAC is designed to operate at 1.2V, this meant that the DAC could not be used at its nominal supply voltage. Instead, the operating voltage for all measurements had to be lowered to 1.0V to guarantee stable operation of the bandgap reference over all chips with an appropriate safety margin. Since the current cells and switches of the DAC have been designed with enough margin to operate without any problems at 1.0V supply voltage, operating the chip at this lowered supply voltage does not affect the measurements in any significant way.

The second problem is noise in general, and noise from the bandgap in particular. By design, the LSB of each single-ended channel is $1\mu A$, which corresponds to $47\mu V$ after the transimpedance amplifier. However, it became apparent each single-ended output showed noise with a standard deviation (equal to its RMS value) of $\sigma = 88.3\mu V$, or approximately two LSB. While an exact calculation of all the noise sources and noise transfer functions within the chip exceeds the scope of this work, the following observations can be made. Measuring the reference voltage output of the bandgap showed noise with $\sigma = 71.1\mu V$, which, through simulations could be translated to approximately $\sigma = 1.1nA$ in each $10\mu A$ reference current output of the bandgap. Again through simulation it was found that noise with $\sigma = 1.1nA$ in the reference current results in noise with $\sigma = 7.1nA$ in the output of a current cell. For the purpose of arriving at a rough estimate, it is assumed that the only noise source in the chip is the bandgap core and that therefore each current cell adds the same amount of perfectly correlated noise to the output current.

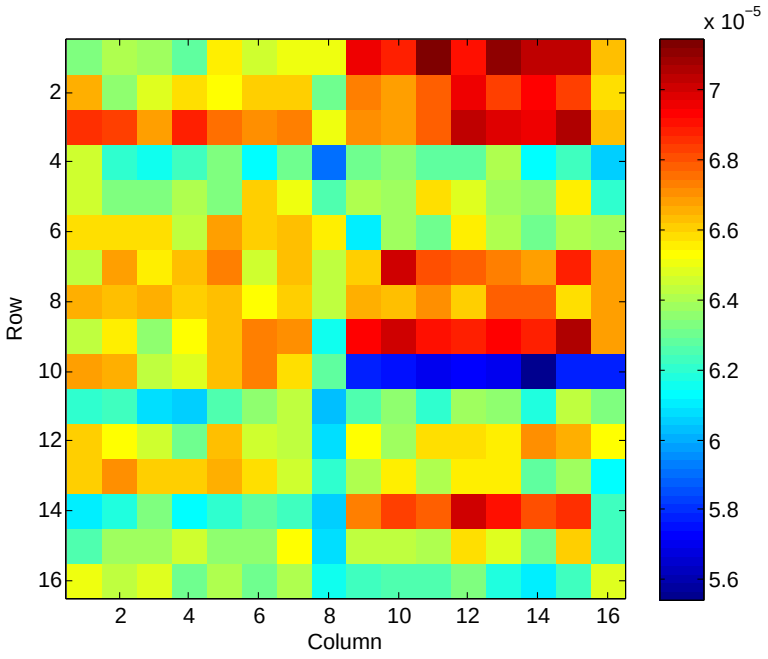


Figure 4.26: Map of unary current cells of chip 3 (in A).

Therefore, the σ from all 255 current cells can simply be added to arrive at the expected noise on the output at full scale, $\sigma = 1.8\mu A$, or $\sigma = 85.1\mu V$ after the transimpedance amplifier. This figure is very close to the measured value, even though some matters have been simplified in these calculations. Nevertheless, this indicates that a significant proportion of the output noise comes directly from the internal reference. Power supply and transimpedance amplifier reference noise have already been minimised by using a primary cell and low-noise buffers, resulting in power supply noise of $\sigma = 5.0\mu V$ and reference noise of $\sigma = 3.3\mu V$. Since no external reference signal can be used instead of the bandgap, there is no feasible way to reduce the output noise. However, since this noise was determined to be gaussian, DC measurements can still be performed but require averaging over a number of measurements to accurately determine the mean. This is absolutely necessary especially for further measurements concerning individual or a small number of CATs, since the designed LSB current of the CAT is $570nA$, which could otherwise not be measured in the presence of noise.

4.6.3 CAT Testing

The basic function of the CATs in the current cell array was tested by the following method. First, the DAC transfer function of a single channel is obtained, with measurements taken at codes that correspond to a CAT-enabled current cell, resulting in a 32-point transfer function. This reduces measurement time and analysis complexity without losing any information significant for testing CATs only. For the first transfer

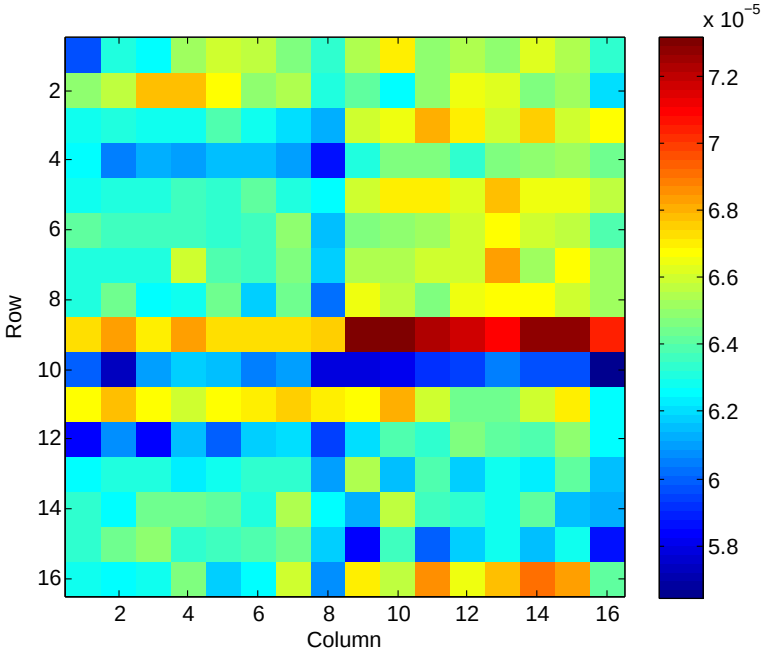
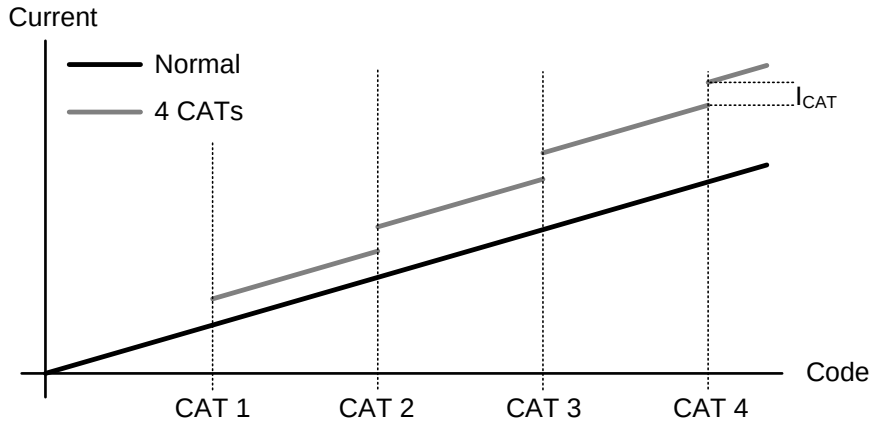


Figure 4.27: Map of unary current cells of chip 4 (in A).

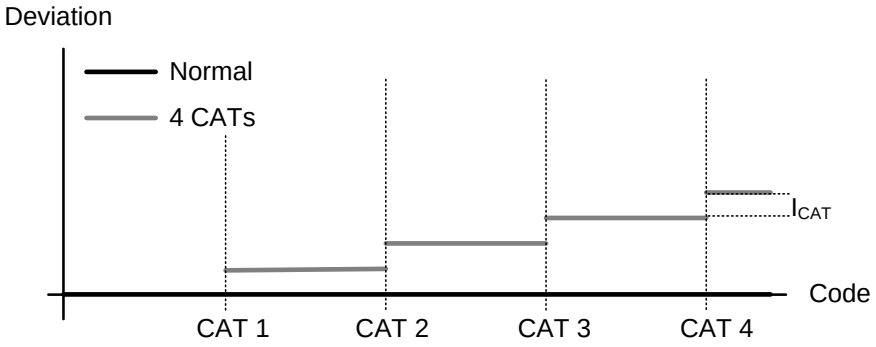
function, all CATs are set to their lowest setting, i.e. all calibration slices off. Then, the CAT calibration slices that are to be measured are turned on and another transfer function is obtained. The resulting two transfer functions are illustrated in Figure 4.28a, where four CATs are activated over the entire range. At each CAT-enabled code, the second transfer function is expected to make a step change corresponding to the number of activated CAT slices (I_{CAT})with the slope between these points unchanged, resulting in an overall higher full-scale current.

In reality, the changes in current caused by CAT are very small compared to the full-scale value, which would make visual distinction on a graph like Figure 4.28a impossible. Furthermore, as will be seen in Section 4.6.4, INL of the DAC is significantly larger than a single CAT's adjustment current, which means that in a direct comparison of the two transfer functions, the CAT current could not be distinguished from the inherent variation between cells. Is is therefore more practical to use the first transfer function as a reference, and subtract it from the second one. This way, any common terms such as slope and INL are cancelled and the effects of the CAT can be measured much more easily, as illustrated in Figure 4.28b.

Several different measurements have been conducted using this method. Figure 4.29 shows the effects of CAT when, starting from the 16^{th} CAT cell, consecutive CATs are turned on. It can be seen that each CAT increases the output current by approximately the same amount ($4\mu A$, the full-scale value of a CAT) and that the other parts of the transfer function remain unchanged, i.e. the slope of the transfer function before the first and after the last CAT cell remain the same as in the reference. The currents of



(a) Transfer function altered by four CATs



(b) Referenced to linear transfer function

Figure 4.28: Effects of CAT on the single-ended transfer function.

each CAT differ slightly as CATs themselves are subject to variation over the current cell array, causing the currents in the sample shown to range from $3.8\mu A$ to $4.3\mu A$. This measurement was taken using a single-ended output and is averaged from 100 samples.

The behaviour of the CATs when measuring one output channel is exactly the one described briefly in the design section. When measuring the differential output, the qualitative effects of the CAT on the transfer function are exactly the same. Consider Figure 4.30a, which shows the transfer functions of both inverting and non-inverting outputs on the same graph. As the input code is incremented, the current in one channel increases from 0 to the full-scale value, while in the other channel the current decreases from the full-scale value to 0, resulting in the same summed value at each point. Therefore, the differential current, shown in Figure 4.30b, ranges between negative and positive full-scale. Consider now the case in Figure 4.30a where a single DAC at code N_{CAT} is activated. The non-inverting channel is affected as seen previously, by a step change in its transfer function at N_{CAT}. Since both channels are fed from complementing sections of the same current cell array, the non-inverting channel starts with the CAT current added and then exhibits a negative step change at N_{CAT}. In the differential output, illustrated in Figure 4.30b, this then results in an overall step change twice the magnitude of the single-ended case, just like the full-range current spans twice the range. Note also in

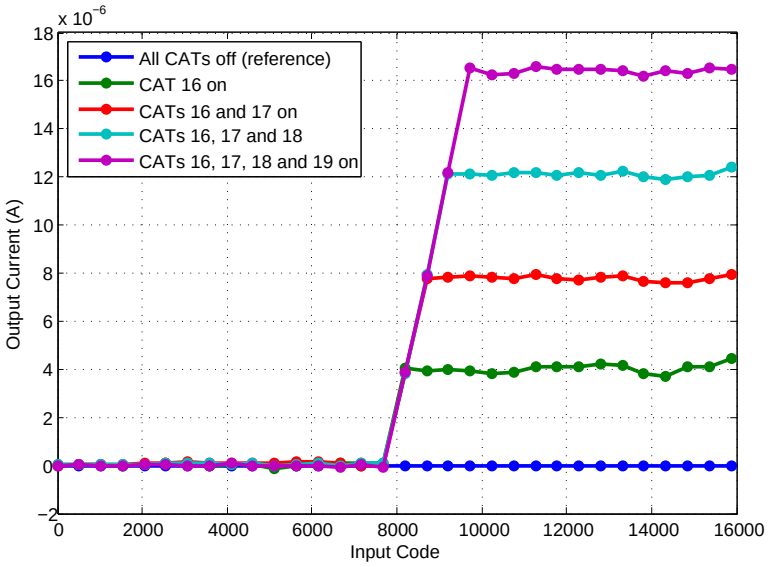
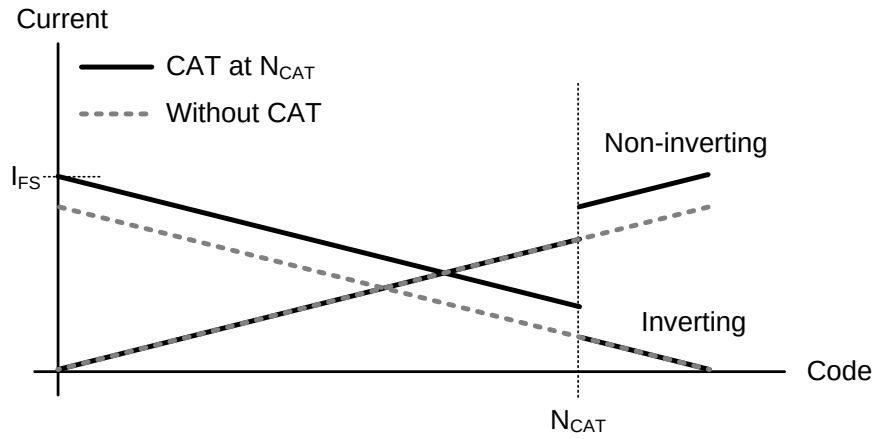


Figure 4.29: Consecutive CATs enabled in single-ended DAC.

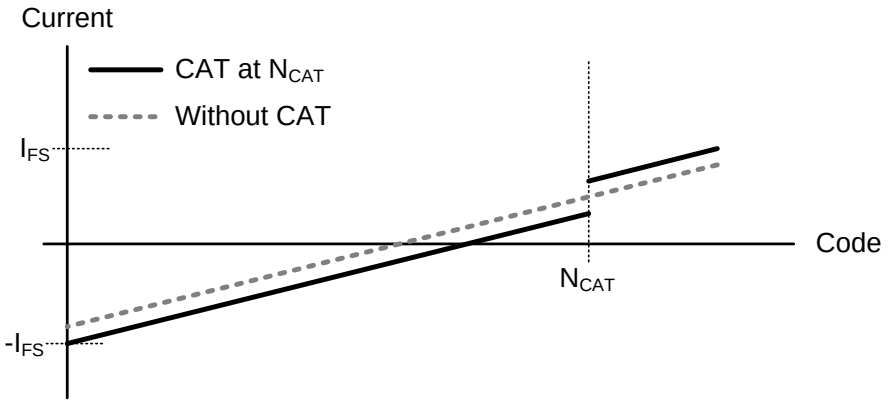
the differential case, the transfer function with CAT is symmetric around the non-CAT transfer function, whereas in the singe-ended case they only started to diverge at the point where the CAT was brought in. Therefore, in effect, the CAT in the differential case causes exactly the same qualitative change on the transfer function and affects it by the same amount relative to full-range, as compared to the single-ended case.

Figure 4.31 shows the same measurement as Figure 4.29 previously, but this time measured from the differential output. The behaviour is exactly as expected, with the change caused by the CAT being twice as large as in the single-ended case and the CAT transfer functions being symmetric around the reference non-CAT transfer function. It it worth noting that compared to Figure 4.29, there is more noise on the measurements for lower codes. This is because the noise in the output current in each channel scales with the current, i.e. in each channel the noise is roughly a fixed fraction of the current. In the single-ended case, this simply means that for very small currents, the absolute noise is also very small and then increases with current. However, the differential output is composed from two complementary channels, where for low codes a large current is provided by the inverting output, adding a large amount of noise

Thus far, CATs have only been used in a binary fashion, where all calibration slices were either turned on or off. In order to test the individual slices, a different measurement is performed. First, all CATs are turned completely off. Then, the input code is set to activate the first CAT-enabled current cell and a reference measurement is taken. Now the CAT of that cell is loaded sequentially with configurations to enable each each slice (0, 1, 2 and 4), and at each step a measurement is taken. The amount of current added by the individual CAT slices can then be determined by subtracting the reference reading



(a) Differential components of transfer function



(b) Differential transfer function

Figure 4.30: Effects of CAT on the differential transfer function.

from each measurement. This process is then repeated for all CAT-enabled current cells, to give a comprehensive overview of the properties of all CATs.

Ideally, all corresponding CAT slices of all current cells would have the same current. However, in reality the CAT slices are subject to variation like any other circuit and therefore each CAT in the circuit has slightly different characteristics. As discussed in Section 4.4, the CATs in the current cells were designed to have a full-range current of $4\mu A$, with slices of $570nA$, $1.14\mu A$ and $2.28\mu A$. Figure 4.32 shows the CAT slice currents for all CATs on chip 3. Table 4.3 summarises the mean and standard deviations of the measured slice sizes and compares them to the designed values. It can be seen that while the means are very close to the designed values, standard deviations are significant, as would be expected for such relatively small transistors. In principle, the CATs do not need to possess a great degree of accuracy since they are only used for calibration. However, in order to optimiser the CAT settings for a given transfer functions, the absolute current supplied by each CAT slice must be known. If the CAT slice currents match the designed values well, the nominal values can be used in that process. Otherwise, each CAT slice

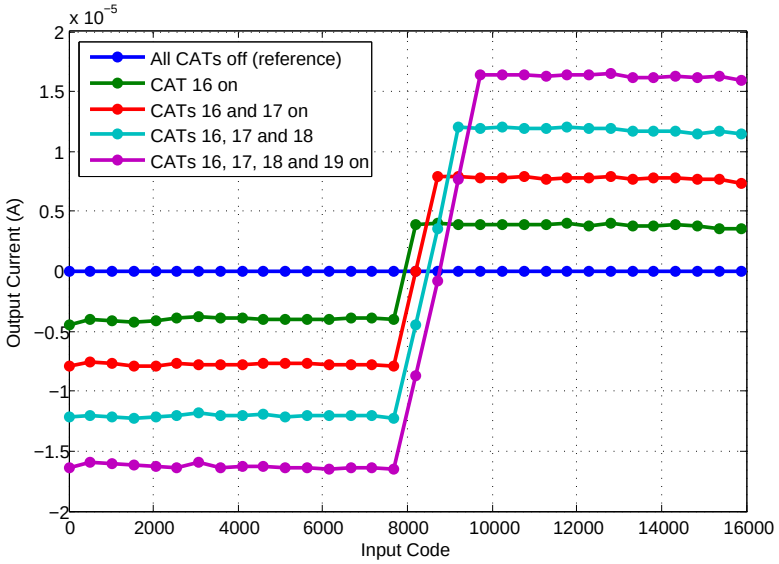


Figure 4.31: Consecutive CATs enabled in differential DAC.

on the chip must be measured individually to provide an accurate model that facilitates optimising CAT settings. This is undesirable, because direct measurement of CAT slices may not be possible in other circuits where the CAT currents cannot be observed directly at the output.

Another factor that may preclude individual characterisation of individual CATs in the time required to examine all CAT devices. When the CATs are configured in a scan chain as in this chip, the time required to load the configuration words is:

$$T_{load} = \frac{N_{CAT} \cdot N_{bit}}{f_{clk}} \tag{4.1}$$

where N_{CAT} is the number of CAT devices, N_{bit} is the number of configuration slices per CAT device and f_{clk} is the clock speed of the scan chain. The time required to characterise all CAT slices by enabling each one in turn and measuring the output is:

$$T_{characterise} = N_{CAT} \cdot N_{bit} \cdot (T_{load} + T_{settle} + T_{measure}) \tag{4.2}$$

where T_{settle} is the settling time of the DAC output and $T_{measure}$ is the time required to take a measurement. For this chip and measurement setup, $N_{CAT} = 255$, $N_{bit} = 3$, $f_clk = 500kHz$, $T_settle \approx 10ns$ and $T_{measure} \approx 250ms$. In this particular case, the relatively long measurement time required to achieve the necessary accuracy dominates this expression, resulting in an overall characterisation time for all CATs of 192s. Such a long time period would not be acceptable in a production environment, making a full characterisation of CATs not feasible for circuits that contain many CATs and have relatively long measurement times, such as in this case. Therefore, CAT accuracy is

an important factor as it may eliminate the need for full characterisation, which may present an obstacle for the practical application of CAT.

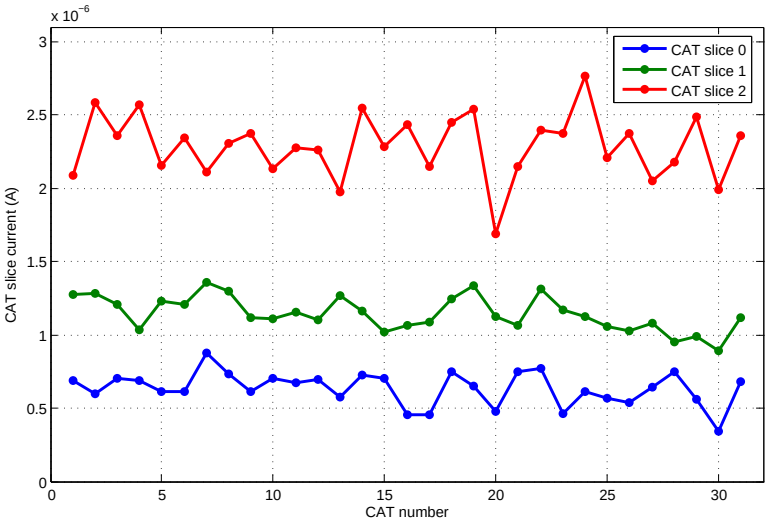


Figure 4.32: Size of CAT slices over current cell array.

Slice	Mean	Std. dev	Std. dev.	Min.	Max.	Designed
1	$635nA$	$114nA$	18.0%	$343nA$	$873nA$	$570nA$
2	$1.14\mu A$	$117nA$	10.2%	$891nA$	$1.36\mu A$	$1.14\mu A$
3	$2.29\mu A$	$218nA$	9.5%	$1.69\mu A$	$1.76\mu A$	$2.28\mu A$

Table 4.3: CAT slice summary

4.6.4 Improving INL

The main purpose of fitting CATs to the DAC is to improve its INL. The concept of INL has already been described in Section 2.4 and from the previous Section 4.6.3 the effects of the CAT on the DAC transfer function are also known. It remains now to apply the CAT to alter the transfer function and thereby reduce INL.

Figure 4.33 illustrates a single-ended DAC transfer function, which is subject to non-linearity. Compared to the linear reference, the uncompensated transfer function has a lower slope up until the mid-point and then sloped up again. Suppose that this hypothetical DAC is equipped with four CATs, each of which has three states: -1, 0 and +1, where 0 means no change to the transfer function and +1 and -1 mean an addition or subtraction of $I_{CAT}$to all following currents. For this particular transfer function, a suitable calibration sequence for the CATs would be +1 +1 0 -1, which will yield the illustrated compensated transfer function. This compensated transfer function shows a much smaller INL ($\mathrm{INL_{CAT}}$) compared to the uncompensated transfer function ($\mathrm{INL_{UC}}$).

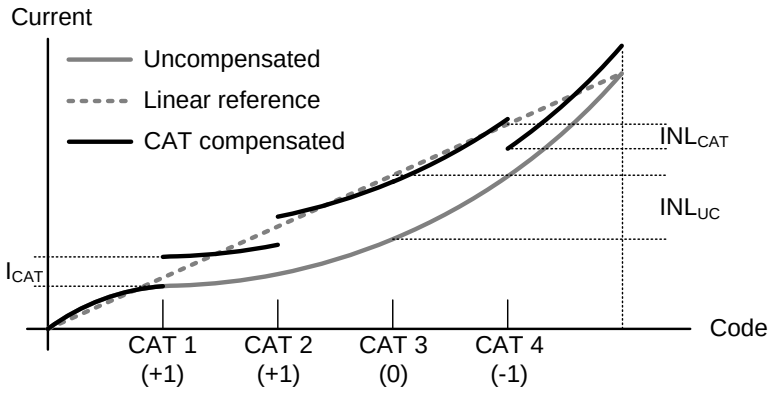


Figure 4.33: Illustration of INL compensation with CAT.

In order to apply this scheme to a DAC, the existing INL must first be measured. Conventionally, this would require a full sweep over all input codes to record the complete transfer function of the DAC. Seeing that due to the noise issues mentioned in Section 4.6.2 several measurements need to be averaged, such a full transfer function would take impractically long time to record. However, for the purpose of measuring INL, a full sweep is in fact not necessary. In this DAC design, INL is only caused by mismatch between the unary current cells. Therefore, it is not necessary to consider the least significant binary current sources, which reduces the number of required codes to be evaluated from 16384 to 256. This is a practical value and is used for measurements throughout this section.

From this transfer function, the INL can be calculated as the difference of each point from the linear reference. Such an idealised linear DAC transfer function is defined by offset and gain or full-scale value. From these, a line can be placed between the 0-code value and the full-range value and then used as the reference. In the case of this DAC, the situation is slightly more complicated by effects of CAT on the transfer function. As shown in Section 4.6.3, CATs not only alter the shape of the transfer function, but also the full-scale value or full-scale values, in the case of differential measurement. It is therefore not possible to find a set of of parameters, such as offset and gain that are valid for each chip. Instead, these parameters are different for each CAT setting on each chip. This requires that each INL measurement uses its own parameters to determine a linear reference that is specific to that particular measurement. Whenever different INL curves of the same chip are compared, they each use their own reference line. Figure 4.34 illustrated this point, where two transfer functions of the same chip use different reference lines because they have different gain due to different CAT settings. In this example, the upper transfer function has all CATs turned on to give a steeper slope, while for the lower transfer functions, all CATs are off. This results only in a change in gain, but does not alter the shape of the transfer function. Furthermore, the ratios of peak-to-peak INL to full-scale value the same in both cases, so that the peak-to-peak INL in the second case is reduced by the same factor as the gain. This is a very important

point to note because it means that simply by applying CATs to a transfer function in a way that alters the full-scale value also changes the magnitude of the INL already, without even considering the effects of individual CATs on the shape of the transfer function. However, for a typical DAC such as the one in this work, the effects of CAT on the full-scale value do not contribute enough to a change in INL that this effect would need to be considered in detail. For example, the maximum change in full-scale value and therefore INL by design in this DAC is 0.8% of the nominal value.

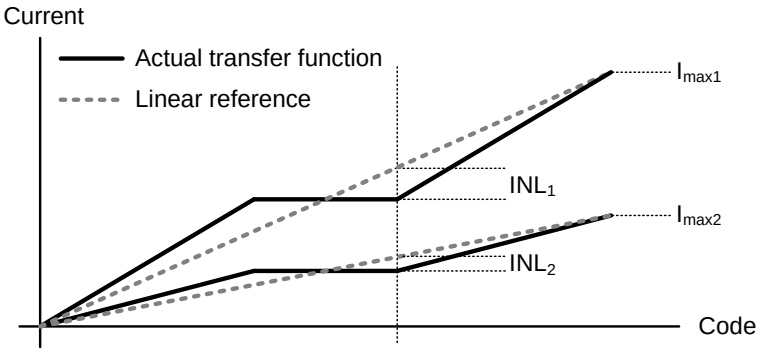


Figure 4.34: Different reference lines for same transfer function with different gain.

Once the transfer function has been measured and the reference line established, the INL can be readily calculated. Figure 4.35 shows the INL for one of the chips. As can be seen, the INL is mostly positive, meaning that the transfer function is above the linear reference for most of the time. It should also be noted that the peak-to-peak INL is $274\mu A$, while a single CAT can only compensate $4\mu A$. The immediate consequence for compensating the given INL is that it is not possible to adjust all the individual deviations, such as the smaller initial and larger middle peak. Instead, several consecutive CATs need to be combined to achieve an overall improvement by appropriately shaping the overall transfer function.

In order to compensate for the given INL, a modified transfer function like the one shown in Figure 4.36 is desirable. Compared to the non-CAT reference, it features a lower slope in the first half and a steeper slope in the second half, resulting in a dip around the centre. When this dip is lined up with the peak of the INL, the INL can be reduced. Fortunately, shaping the transfer function in such a way with CAT is possible and also yields the maximum possible deflection in the centre of the transfer function. For this, the first half of CATs are set to lower their cells' current, while the second half is set to increase the current, leading to exactly the change in transfer function illustrated in Figure 4.36. This optimised CAT configuration was determined by a software tool, which applies CATs of a given size to the uncompensated transfer function and optimised the configuration by means of a numeric algorithm. The simulated CAT INL curve is also obtained during this process as the output of the optimisation process.

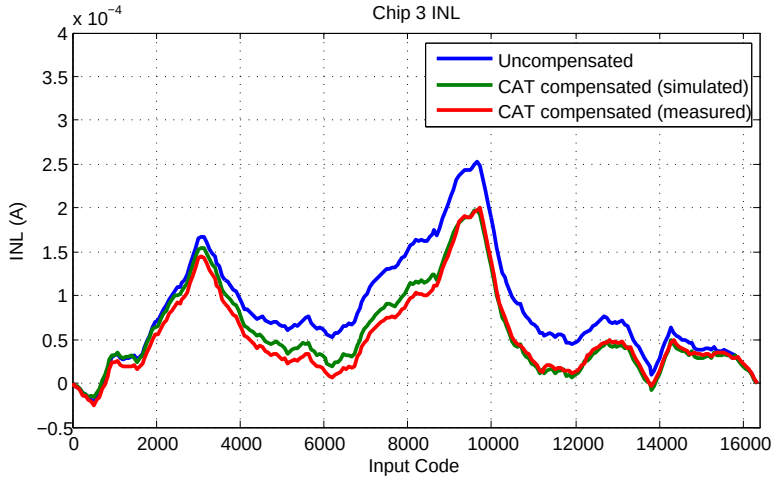


Figure 4.35: INL of Chip 3, uncompensated and compensated by CAT

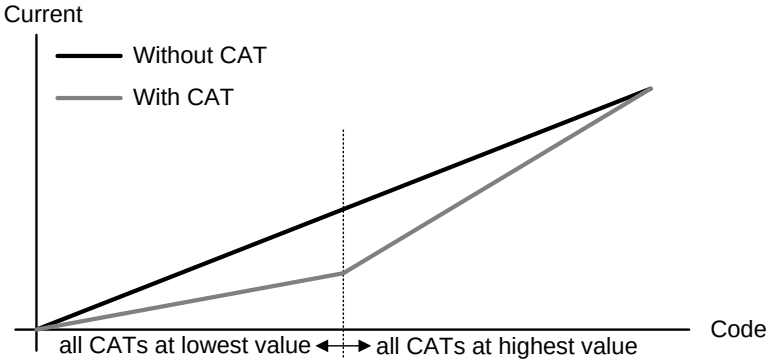


Figure 4.36: Largest possible adjustment through CAT

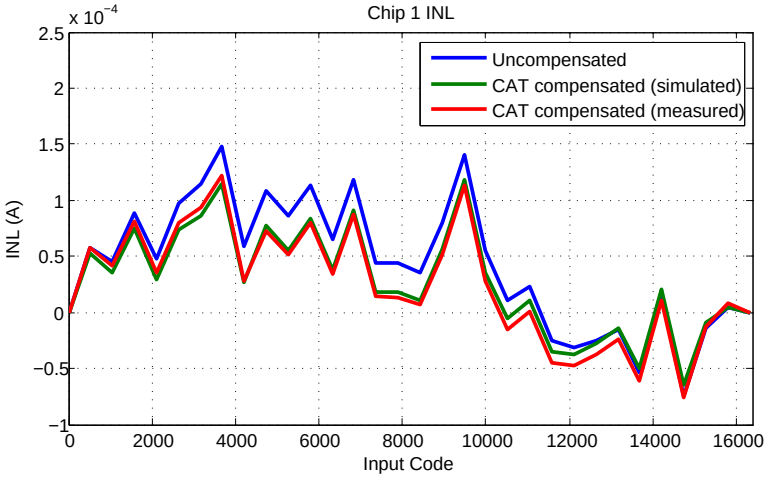


Figure 4.37: INL of Chip 1, uncompensated and compensated by CAT

It can be seen that the peak-to-peak INL with CAT is lower than without. In the uncompensated case, peak-to-peak INL is $274\mu A$, while compensation reduces it to $226\mu A$, an

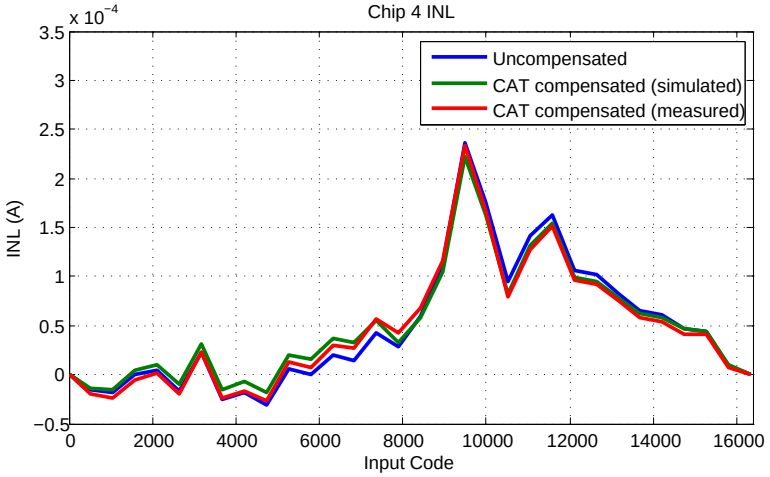


Figure 4.38: INL of Chip 4, uncompensated and compensated by CAT

improvement of 17.4%. In addition, Figure 4.35 also shows the expected compensated INL, based on the uncompensated measurement and $4\mu A$ CATs. The measured and expected graphs are nearly identical, with only small deviations being due to non-ideal CATs and measurement error.

Figures 4.37 and 4.38 show the uncompensated, compensated and expected INL curves for two further chips. For chip 1 in Figure 4.37, the simulation is again reasonably close to the measured result and the achievable improvement of the same order as chip 3 in Figure 4.35. However, chip 4 in Figure 4.38 shows a much smaller improvement and also less agreement between the simulation and measured results. The INL parameters of these three chips are summarised in Table 4.4 in Section 4.6.5.

In order to illustrate how the INL of a DAC could be improved using larger CATs, the software tool described in Section 4.5.3 was run on the uncompensated transfer function of chip 3, but with a CAT full-scale current of $32\mu A$, again split over three slices. Then, the resulting CAT configuration was applied to the transfer function in software to determine the resultant INL curve like in the previous simulations. The result of this can be seen in Figure 4.39. The uncompensated peak-to-peak INL is again $274\mu A$, but in this case, the compensated INL is $125\mu A$, an improvement of 45.7%, compared to 17.4% that were achieved on the real chip with $4\mu A$ CATs. With these CAT sizes, the transfer function shows the distinct dips that are expected from CAT compensation, which are especially visible in the region between codes 6000 and 8000. This illustrates that with appropriately sized CATs, the adjustment of the transfer function can be significantly more effective. With this capability, this software tool completes the fully automated CAT design cycle of Section 2.5.2 in that it provides a means for automated post-fabrication CAT configuration.

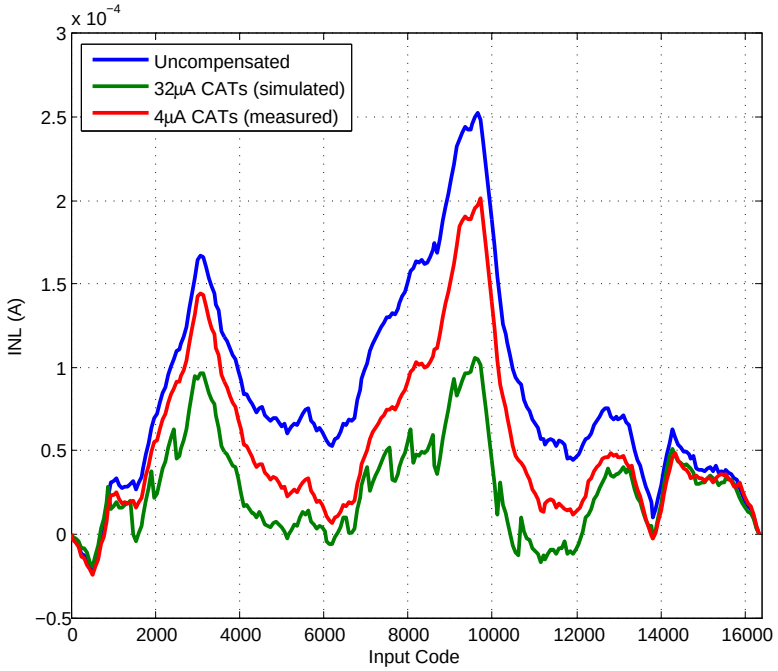


Figure 4.39: INL compensated with bigger CATs.

4.6.5 Summary of Results

As discussed in Section 4.6.1, of the ten chips fabricated, three were successfully tested and their performance measured. The initial INL performances of these chips and the achievable CAT-compensated INL are given in Table 4.4.

Chip	Uncompensated	Simulated	Measured
1 (Figure 4.37)	$222.6\mu A$	$182.7\mu A(-17.9\%)$	$197.6\mu A(-11.2\%)$
3 (Figure 4.35)	$273.6\mu A$	$213.8\mu A(-21.9\%)$	$226.0\mu A(-17.4\%)$
4 (Figure 4.38)	$267.6\mu A$	$240.4\mu A(-10.2\%)$	$259.6\mu A(-2.99\%)$

Table 4.4: INL improvements

In this section, only improvements of INL were considered, whilst DNL was not. This is because the CATs were designed into the chip in a way that only facilitates improvement of INL. The sources of DNL in this design are mismatch between all binary current cells, and the sum of the binary and each unary current cell. Since CAT are placed in every 8^{th} unary current cell, they cannot be used at all to improve the matching between the binary current cell. Furthermore, every unary current cell would have to be equipped with CAT to allow matching of the unary cells to the binary cells. These CATs would then also have to be sized differently to allow for adjustment of both INL and DNL. Therefore, if adjustment of DNL were required, the design of the chip and CAT placement strategy would be significantly different.

4.7 Discussion

In this chapter, the application of CAT to a segmented current-steering DAC was described. This DAC was then fabricated and the concept of CAT proven on silicon.

The design of the DAC was relatively straightforward, with the DAC itself being a standard design item. The current cells to which CATs were applied were chosen manually in regular intervals rather than though the fully-automated CDI algorithm because the large number of simulations required for CDI would be impractical for a large mixed-signal circuit like this DAC. Furthermore, because of the regular structure of the current cell array and knowledge of the variation mechanisms that affect it, manual CAT placement very likely led to a good, if not optimal, CAT placement. Sizing of the CATs was done using the automated tool based on variability simulations of an individual current cell. In this step, the variability from the bias circuit was not considered and variability models of some components were optimistic, which resulted in CAT sizes that were not optimised for the fabricated chip.

Measurements on the chips largely confirmed the expected operation of the CAT. All fundamental CAT functions, such as the way in which the DAC transfer function can be altered, were consistent with the design. Furthermore, the achieved INL improvement, as well as the automatically determined CAT configuration, are qualitatively and quantitatively consistent with calculations from the post-fabrication CAT configuration tool.

As far as chip performance is concerned, the inherent and calibrated INL are significantly larger than that of other works discussed in Section 2.4, which generally achieve INL in the order of a few or less than one LSB, while this design shows INL of around 200 LSB. This is due to several factors, most of which can be ascribed to insufficient experience with chip design to gauge the quantitative effects of layout strategies. For instance, certain layout and design techniques that would inherently improve INL were intentionally not applied in order to show the improvement that CAT alone can achieve. This included techniques such as separate current cell and switch arrays, distributed current cells and an optimised switching sequence. In [69], the authors show that by using such techniques alone, the INL of a 14 bit current-steering DAC could be reduced to 0.3LSB.

Additionally, an oversight at the design stage meant that mismatch between the bias generators was not considered when sizing CATs. This resulted in CATs that were sized too small even for a DAC with good layout strategies, and significantly too small when considering all other effects from the actual layout. To briefly summarise, the best and worst uncompensated INL of the DAC were 223LSB and 274LSB, respectively. The best and worst relative INL improvements through CAT were 17.4% and 2.99%, respectively, where the best CAT-compensated DAC showed an INL of 198LSB.

Considering these inadequately sized CATs and sub-optimal initial performance, it was still possible to achieve significant INL improvement on some of the chips. Although the INL of these adjusted transfer functions are not nearly good enough to compete with other state-of-the-art calibration techniques, they nevertheless show that the CAT can in principle be used for calibration of complex circuits and DACs in particular. Considering the relative improvement that was achieved by these non-optimal CATs, it would seem reasonable to expect performance in line with other current work (Section 2.4) if the circuit were designed again with the knowledge gained from this design.

4.8 Conclusions

Up to this point, the focus of this work has been on post-fabrication calibration techniques. Regardless of the practical performance of this particular example application of CAT, this chapter has shown how an integrated circuit with a calibration scheme can have its performance tuned after fabrication. In the case of CAT, this tuning was with a particular focus on performance variation due to parameter variation during fabrication.

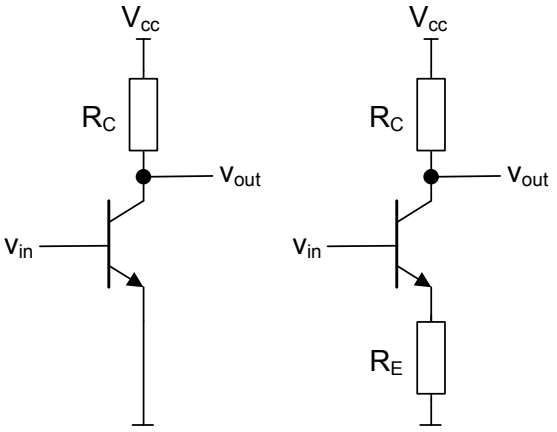
However, there are alternative techniques which may be employed. One category of such techniques is robust design, which adds inherent variation tolerance to a circuit. In some cases, robust design techniques may be used as an alternative to calibration if calibration is not applicable. In most cases, inherent variation tolerance achieved by robust design is desirable regardless of the application of calibration schemes. Therefore, the following chapter will highlight a different approach of minimising performance degradation due to variation effects. It will be shown how a circuit can be made inherently more robust against parameter variation, thus removing the need for calibration and increasing reliability.

Chapter 5

Variation-Tolerant Design

5.1 Introduction

The preceding two chapters of this work have focused on calibration as a method to reduce the effects of parameter variation on circuit performance. However, as discussed in Section 2.3, calibration is not the only way to achieve better circuit performance under the presence of parameter variation. Another approach, for which several examples were given, is robust circuit design. Some examples of the "flavours" of robust design discussed in Section 2.3 were optimising component values to provide better resistance against parameter variation, or adding compensation circuits. These are explicit methods to increase the inherent variation tolerance of a circuit.



(a) Amplifier without feedback. (b) Amplifier with feedback.

Figure 5.1: Bipolar common-emitter amplifiers with and without negative feedback.

However, robust design in a more general sense is a fundamental trait of analogue circuit design. For instance, consider negative feedback. Virtually every analogue circuit uses

negative feedback in some form, and its ability to decouple circuit performance from component values is employed universally. Yet the full breadth of its usefulness as a method of robust design is seldom appreciated. Consider for example the bipolar transistor amplifiers in Figure 5.1. The amplifier in Figure 5.1a does not have negative feedback, whereas the amplifier in Figure 5.1b contains negative voltage feedback through resistor R_E. The small-signal voltage gain for the amplifier without feedback is given by

$$A = \frac{v_{in}}{v_{out}} = -R_C \cdot \frac{\beta}{r_{BE}} \tag{5.1}$$

and for the amplifier with feedback by

$$A = \frac{v_{in}}{v_{out}} = -\frac{\beta \cdot R_C}{r_{BE} + R_E\,(\beta + 1)} \approx -\frac{R_C}{R_E} \text{ for } \beta \ggg \tag{5.2}$$

where β is the small-signal base-collector current gain and r_{BE} is the small signal base-emitter input resistance of the transistor.

Given a desired amplifier gain, a designer will likely choose the circuit of Figure 5.1b, mainly because its gain can be set by the resistors. This is because negative feedback largely removes the dependence of gain on the device parameters of the transistor, β and r_{BE}. However, the consequences of this property extend far beyond convenience for the designer. As the device parameters have a smaller effect on the overall circuit performance, the circuit becomes more tolerant against variation in those parameters. Temperature dependence, ageing, supply voltage dependence and other extrinsic sources of variation no longer show in the gain of the circuit, resulting in greater stability during operation. Likewise, as the sensitivity to intrinsic parameters like β and r_{BE} is reduced, the circuit can be reproduced with different transistors whose parameters do not have to be precisely matched, allowing economic production of the circuit. These effects are even more striking when other circuit performances such as its AC transfer function are considered. In the case without feedback, the AC gain shows strong frequency dependence, whereas negative feedback results in a constant, well-defined gain for a certain bandwidth. Indeed, it were these aspects of robust design that Harold Black focussed on when he formally invented negative feedback [84]. This shows that the scope of robust design is much wider than a number of specific techniques with particular applications. Instead, it is a fundamental concept of electronics design which can be specifically targeted at improving variation tolerance.

The focus of this thesis is not specifically on robust design, but methods for reducing the effects of variation in general. The calibration methods discussed in the previous chapters are valuable and can significantly reduce the effects of variation on a circuit. However, there are applications where calibration is not an ideal choice. For example, most calibration methods require the circuit performance or other parameters to be

measured, either during fabrication or during operation. Furthermore, most calibration techniques only produce results which are valid for a single set of conditions, and require the settings to be changed as parameters change, as shown in Section 3.4. Finally, most calibration schemes result in some kind of overhead, either in the form of additional circuitry or special requirements on the fabrication process or equipment. In some cases, these properties can make a product expensive, may be inconvenient or otherwise make calibration unsuitable for a given circuit or application.

In such cases, robust design on its own can be a viable alternative to calibration. Unlike calibration, robustly designed circuits do not typically need post-fabrication tuning or intervention as parameters change. Robust design can therefore increase both yield and reliability, much like the combined online and offline calibration scheme of Section 3.6. It must also be noted that in many cases calibration and robust design are not mutually exclusive. Circuits can be designed with inherent robustness against variation in some parameters, and additional calibration for other parameters. However, the case study presented in this chapter focuses on robust design as the sole method to achieve variation tolerance.

To illustrate the key concepts of robust design, an existing circuit is replaced entirely by another circuit working on different principles, which shows less sensitivity to parameter variation, and which is therefore more robust. The case study followed in this chapter is based on Electromechanical $\Sigma\Delta$ Modulators used in pick-off circuits for MEMS accelerometers, which are very sensitive to parameter variation. The novel replacement circuit, the Electromechanical Phase-Locked Loop, is significantly more robust while providing similar performance by virtue of its completely different operating principle.

The EM-PLL was developed during the process of researching interesting target circuits for the CAT. Whilst the DAC of Chapter 4 was ultimately chosen as an example circuit for the CAT, Electromechanical $\Sigma\Delta$ pick-off circuits were also considered due to their sensitivity to parameter variation. However, EM-$\Sigma\Delta$ are in fact one example for a circuit which is unsuitable for the application of CAT due to several reasons. First, as will be discussed in Section 5.2.1, the stability and performance of high-order $\Sigma\Delta$ modulators is affected by a large number of parameters, including ones over which the designer has no control, such as the input signal. Furthermore, this dependence is only approximately linear in a very narrow range of parameters, with small changes in parameters leading to abrupt changes in performance. The CAT tools in their current form are not suited for circuits that show such behaviour, therefore precluding the use of the automated CDI and sizing processes. Second, the simulation of Electromechanical $\Sigma\Delta$ modulators is computationally expensive. Even if the CAT tools were applicable to high-order $\Sigma\Delta$ circuits, the resultant simulation times required would be impractical. Therefore, a more heuristic approach was sought to improve the variation tolerance of Electromechanical

ΣΔ systems. The development process started with the idea of replacing an amplitude-based system with a frequency-based system, and is mirrored by the description of the system in Section 5.3.3.

To better understand the issues of the original state-of-the-art EM-ΣΔ circuit, Section 5.2 will begin with a summary of the operating principle of force-feedback MEMS accelerometers. This is followed by a brief review of current interface circuits, with a particular focus on Electromechanical ΣΔ Modulators. Section 5.3 then introduces the novel EM-PLL, describes its operation and highlights the key differences to EM-ΣΔ circuits. Section 5.4 then compares simulation results between the EM-PLL and EM-ΣΔ to highlight the achievable improvements in terms of robustness to parameter variation and general performance characteristics.

5.2 MEMS Accelerometers

The operational principle of modern MEMS accelerometers with force feedback is described in detail in [85], but will be reviewed briefly in this section. The simplified structure of a typical MEMS accelerometer is shown in Figure 5.2. A moveable proof mass is suspended by springs and restricted to move in one axis (vertical in this example). The proof mass is furthermore equipped with combs that form parallel plate capacitors with corresponding stationary combs on the sensor substrate. When the sensor is subject to acceleration, the mass is displaced by the force, which changes the air gap of the parallel plate capacitors and thus their capacitance.

These sense capacitors are commonly made differential, such that displacement of the mass results in an increase of capacitance in one capacitor and equivalent decrease of capacitance in another capacitor. The accelerometer illustrated in Figure 5.2 also features a force-feedback system, which allows the exertion of electrostatic force on the proof mass through application of a voltage. This force is used to counteract the force caused by acceleration, which in turn reduces the displacement of the proof mass for any given acceleration. This reduction of mechanical movement increases the dynamic range of the accelerometer by keeping the proof mass within its mechanical limits for higher accelerations.

From an electronics point of view, sensing acceleration using such an accelerometer effectively means measuring differential capacitance. Likewise, closed-loop operation requires the application of a differential feedback voltage derived from the measured acceleration to the feedback system. The remainder of this section will review different techniques used for sensing acceleration using such MEMS accelerometers.

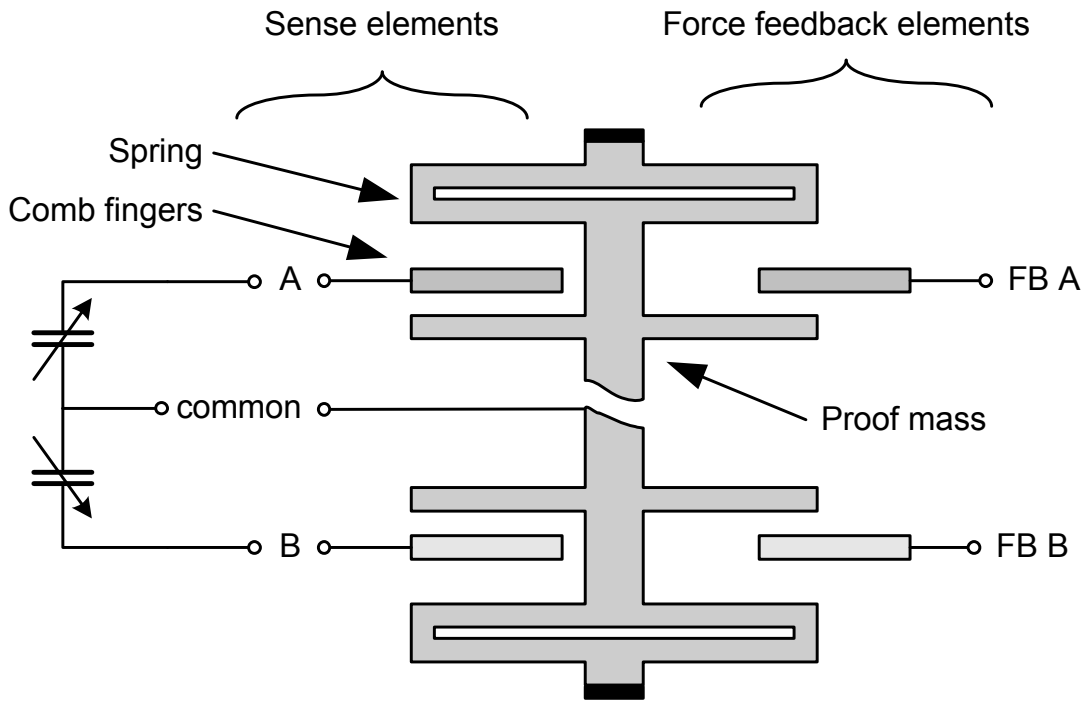


Figure 5.2: Structure of a MEMS accelerometer.

5.2.1 Electromechanical ΣΔ Modulators

5.2.1.1 Oversampling Data Converters

Currently, the most widely used state-of-the-art technique that inherently employs both sensing and feedback are Electromechanical Sigma-Delta Modulators. Before going into a description of its operation and features, the following is a brief recap of ΣΔ modulators and their general properties.

Any data converter suffers from quantisation errors, which stem from the fact that it can only resolve a finite number of discrete input values [86]. Any inputs not coinciding perfectly with a quantisation level therefore lead to an error, which manifests itself as noise in the resultant digitised signal. Notwithstanding other noise sources introduced during the conversion process, the quantisation noise itself is only dependent on the size of the LSB, as follows:

$$e_{RMS}^2 = \frac{q^2}{12} \tag{5.3}$$

where q is the LSB size in volts. From there, the Signal to Quantisation Noise Ratio (SQNR) for a sull-scale sinusoidal input signal can be calculated as

$$SQNR\,[dB] \approx 1.76 + 6.02 \cdot Q \tag{5.4}$$

where Q is the number of bits of the converter.

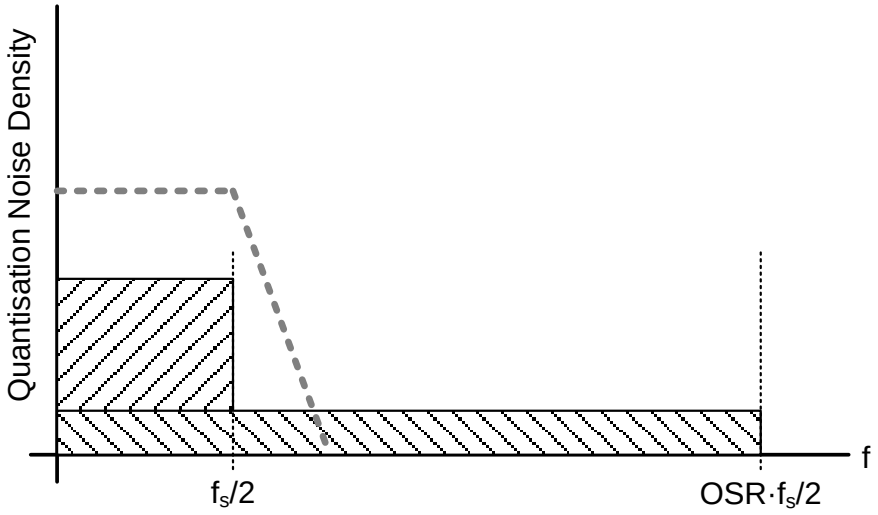


Figure 5.3: Quantisation noise spectra for signal at minimum and oversampled sample rates.

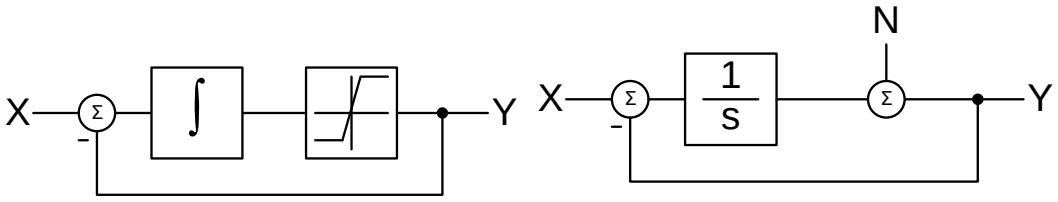
Consider now the spectrum of the digitised signal, as illustrated in Figure 5.3. The bandwidth of the digitised signal is directly dependent on the sample rate as per Nyquist's criterion. However, the total quantisation noise power is constant regardless of sample rate. This means that, for a higher sample rate, quantisation noise is spread over a wider bandwidth and thus leads to a lower quantisation noise density. This property is exploited in a concept known as oversampling. Instead of sampling a signal with the minimal required sample rate $f_s = 2f_0$, it is sampled with a significantly higher sample rate, leading to a lower noise density. Then, the digitised signal is filtered to extract the desired signal bandwidth. Since the noise density within this bandwidth is lower, the resultant noise power in the final digitised signal is also lower. This reduction in quantisation noise power is effectively equal to sampling at Nyquist frequency with a higher-resolution converter. The resultant in-band quantisation noise power scales linearly with

$$e_{RMS}^2 = \frac{q^2}{12} \cdot \frac{2f_0}{f_s} = \frac{q^2}{12} \cdot OSR \tag{5.5}$$

where OSR is the oversampling ratio, the factor by how much larger than the Nyquist frequency $2f_0$ the sample frequency f_s is. Again, the SQNR can be calculated for an oversampled converter:

$$SQNR\,[dB] \approx 1.76 + 6.02 \cdot Q \cdot log\,(OSR) \tag{5.6}$$

from where it can be seen that the oversampling ratio improves the SQNR. However, this relationship is only logarithmic, and therefore increasing the oversampling ratio is not as effective as increasing the resolution to increase the SQNR. It must be noted that this



(a) Block diagram of a ΣΔ modulator

(b) Linearised frequency-domain model of a ΣΔ modulator

improvement only considers quantisation noise. Other noise sources, such as out-of-band noise that has not been filtered by the anti-aliasing filter, cannot be reduced by this method, since its power spectral density is not affected by the sample rate.

5.2.1.2 ΣΔ Converters

ΣΔ converters are based on the concept of oversampling but utilise a further technique called noise shaping. This effectively "pushes" the quantisation noise out of the signal bandwidth and results in a much higher resolution gain from oversampling. In fact, ΣΔ converters most commonly only have a 1-bit AD converter and gain all of their resolution from noise shaping and oversampling. The theoretical and mathematical concepts behind the ΣΔ converters are fairly complex and not absolutely necessary for the understanding the the Electromechanical ΣΔ Modulator. Therefore, the following functional overview will just outline the most fundamental properties, as excellent in-depth technical reviews exist in the literature and are out of the scope of this thesis [87].

Figure 5.4a shows the basic block diagram of a first-order ΣΔ modulator. It consists in the forward path of an integrator and a quantiser, whose output is fed back and subtracted from the input signal. If a 1-bit quantiser is used, which is often the case, the output of the ΣΔ modulator is a binary bitstream. For continuous-time analysis, a linearised frequency-domain model can be used, which is shown in Figure 5.4b. Note that the quantiser has been replaced by an additive noise source to add quantisation noise. From this linearised model, an expression for the output signal $Y(s)$ can be derived:

$$Y(s) = N(s)\frac{s}{1+s} + X(s)\frac{1}{1+s} \tag{5.7}$$

It can be seen that the signal transfer function follows a low-pass function, whereas the noise transfer function follows a high-pass function. This is the basis behind noise shaping, where quantisation noise is removed from the signal band and moved to higher frequencies. For higher system orders, this behaviour becomes more and more pronounced. Figure 5.5 illustrates the noise shaping function for different ΣΔ orders [88]. The normalised frequency 1 equals the sample frequency f_s, and the signal frequency is usually much lower. Consider for instance the case where the signal bandwidth is $0.1 \cdot f_s$: As the order of the ΣΔ converter increases, the in-band noise up to $0.1 \cdot f_s$ is reduced

significantly, whilst the out-of-band noise from $0.1 \cdot f_s$ upwards increases. As the signal is band-limited to the signal bandwidth after conversion, the remaining quantisation noise power is reduced and therefore the SQNR increased.

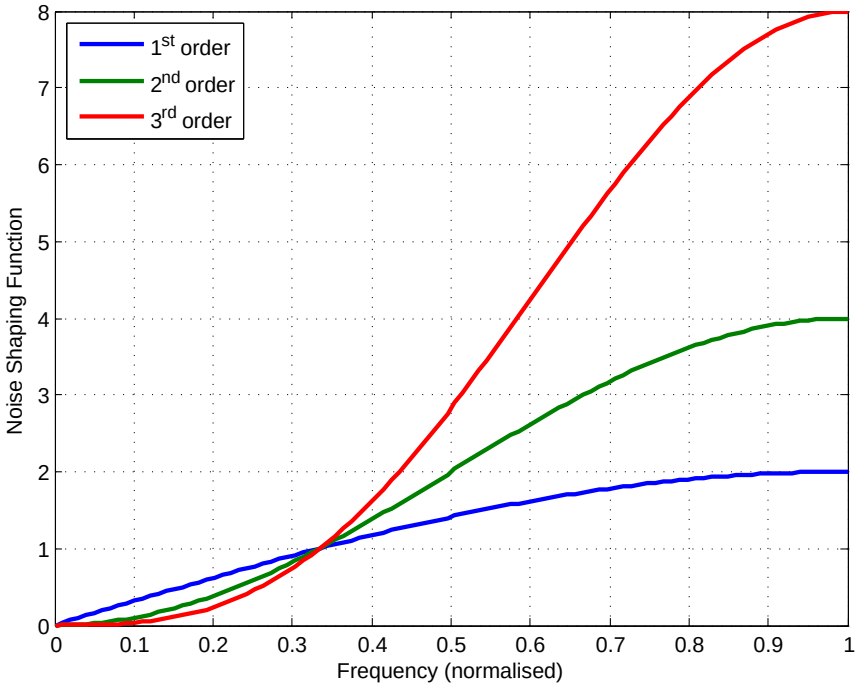


Figure 5.5: Noise shaping functions for ΣΔ modulators.

Higher-order systems can be constructed in a multitude of ways, from simple cascading to complex topologies [87]. To take advantage of noise shaping, a ΣΔ converter must also contain a reconstruction filter at its output, whose purpose it is to pass only the signal band and hence remove the high-frequency out-of-band noise. It is crucial to note at this point that this simplified linear analysis of ΣΔ modulators is only an approximation. Especially with higher system orders, the the linear approximation is only valid for a narrow range of operating parameters, outside of which the modulator may show undesired behaviour, e.g. oscillation [89]. This is important in the context of variation-tolerant design, as it means that the performance of ΣΔ modulators is susceptible to parameter variation, even if they are designed to work well at nominal parameters.

5.2.1.3 Electromechanical ΣΔ Modulators

Having discussed the fundamentals of ΣΔ modulators, their application as sense circuits for force-balanced MEMS accelerometers can now be explored. The block diagram of an EM-ΣΔ system in shown in Figure 5.6. The input acceleration is translated through the proof mass to a force on the spring resulting in a displacement of the mass and ultimately a change in differential capacitance, as described earlier. This differential capacitance is

then measured through a pick-off circuit and fed into a ΣΔ-modulator. At this point, the system would work in open-loop configuration. If closed-loop control is desired, the output bitstream is then returned to the feedback plates, where it controls the position of the proof mass.

The analogue pickoff circuit itself can be implemented in a number of ways, but is typically a charge amplifier, as illustrated in Figure 5.7. A carrier signal (typically several MHz) is injected onto the proof mass at the common point of the differential capacitors, and is then picked off by a differential charge amplifier from the other ends of the capacitors. Any differential change in the value of the capacitors caused by force on the proof mass will manifest itself in a change in output amplitude at the pickoff amplifier, meaning that the signal is effectively amplitude modulated by the force on the accelerometer.

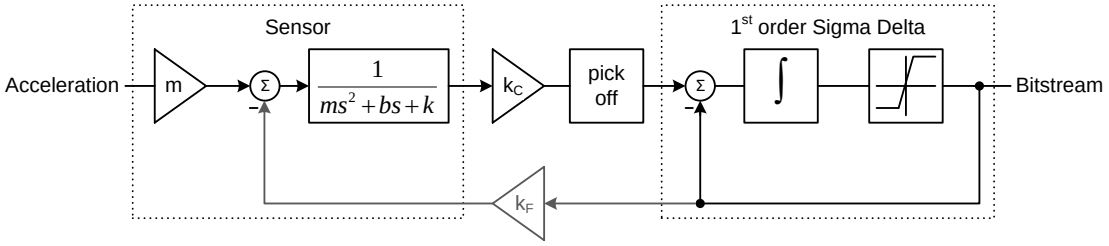


Figure 5.6: Block Diagram of an electro-mechanical sigma-delta modulator.

The AM signal can then be demodulated by normal means, such as a diode demodulator or a double-balanced mixer, resulting in a demodulated baseband signal that is proportional to the acceleration, which then serves as the input to the ΣΔ modulator.

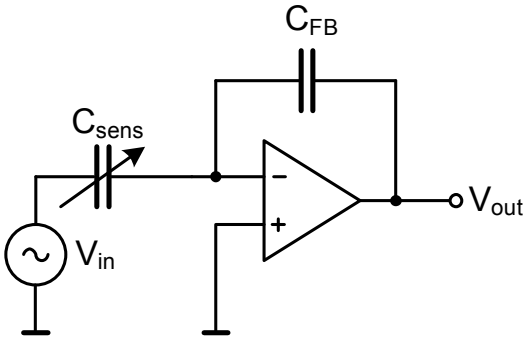


Figure 5.7: Basic capacitive sensing circuit.

There are several properties that make EM-ΣΔ modulators an interesting choice for such MEMS accelerometers. One advantage is that they provide a direct digital output signal for further processing and can incorporate the sensor in the feedback loop simultaneously, which reduces the overall system complexity. In addition, the feedback signal is

digital, eliminating any requirement for linear amplifiers in the feedback path. They also generally offer a good SQNR of the order of 100dB.

However, in spite of all these advantages they pose significant challenges to the designer, most of which are rooted in the aforementioned difficulty of designing stable and robust high-order ΣΔ modulators. Importantly, for Electromechanical ΣΔ systems it is not only intrinsic parameter variation in the ΣΔ modulator (e.g. gains, integrators) that contributes to this difficulty, but also external parameters such as input signal amplitude or mechanical tolerances of the MEMS components. These are, of course, in addition to any other extrinsic sources of variation, such as environmental factors and ageing.

As the sensitivity of ΣΔ modulators to parameter variation increases with order, using a MEMS sensor with a lower order ΣΔ modulator provides systems that are simple to design [90, 91], relatively stable and have reasonable performance. Unfortunately, the inherent disadvantages of lower order ΣΔ modulators in the electronic domain are also well known and also appear in their electromechanical counterparts, resulting in higher quantisation noise due to insufficient noise shaping, dead-zones and the issue of idle tones becoming apparent in the signal bandwidth [92].

With the increasing requirement for sensitivity and low noise, the standard approach has been to develop higher order Electromechanical ΣΔ modulators, with 5^{th} order systems typically providing Signal to Noise ratio in excess of 100dB and excellent overall performance [93, 94]. However, for such higher order systems it has become necessary to provide advanced optimization tools [95] to establish the correct ΣΔ design parameters to ensure stability. While this type of approach can calculate the nominal optimum parameters for a circuit, in practice these can be extremely sensitive to component variation. Even if the parameters are designed to be more robust to variation, as suggested in [95], this involves a complex and time consuming optimization process.

With all of these roadblocks on the way to easy-to-design, stable, high-performance EM-ΣΔ converters, it may seem prudent to explore other topologies that can achieve similar performance, but in addition offer stability and a more analytic design approach. The alternative presented in this work is based on a phase-locked loop (PLL), which is a common building block in communications and signal processing systems. Unlike oversampling systems like ΣΔ modulators, a PLL is to a first approximation a linear system. This means that the system can be designed using largely analytical methods and, more importantly, it is very robust against parameter variation and will show a graceful degradation in performance outside its nominal operating rage. After concluding this review section by briefly exploring work related to the EM-PLL in Section 5.2.2, Section 5.3 will explain the operation and benefits of the EM-PLL in detail.

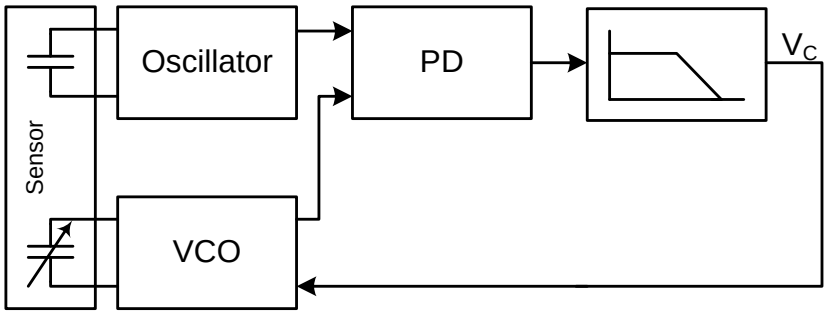


Figure 5.8: Accelerometer pick-off circuit described in [96].

5.2.2 Frequency- or Phase-based Systems

In addition to EM-ΣΔ systems, MEMS sensor interface circuits are an area of ongoing active research. Some of these include systems which use frequency or phase to measure the capacitance of a MEMS accelerometer. As the EM-PLL circuit introduced in this chapter is phase-based, a brief survey of such related techniques is carried out in this section. In addition to interface circuits for MEMS accelerometers, there are other kinds of MEMS sensors whose interface circuits are also relevant. The remainder of this section reviews a small selection of the most relevant sensing techniques, but it is worth noting that thus far there are no systems in the literature which incorporate all key features found in the EM-PLL. These are, namely, the use of differential sensors, where the sensor capacitances are used in a differential oscillator, and a force feedback system whose feedback voltage is derived from the phase difference between the two oscillators.

Matsumoto [96] describes a single ended PLL pickoff system, which is illustrated in Figure 5.8. For all practical purposes, it consists of a conventional PLL, as described in Section 5.3.1. The primary input of the PLL is a stable reference frequency, whilst the VCO inside the PLL uses the sensor capacitance for timing. During steady-state operation with no acceleration, the reference and VCO frequencies are in lock. If acceleration is applied to the sensor, its capacitance and therefore the VCO output frequency changes. The PLL reacts by a change in output voltage, which in turn pulls the VCO frequency back to the reference frequency. The output voltage of the PLL is therefore a measure for the applied acceleration. Since the sensor used does not have provisions for electrostatic control, the PLL does not form a closed control loop over the sensor. However, this work is one of the most fundamental examples of how the sensor capacitance can be used in an oscillator and a PLL structure used to generate a proportional output voltage. Furthermore, by fabricating the reference oscillator capacitor on the MEMS device, this particular paper also addresses some aspects of robust design: Temperature and ageing will affect both sense and reference capacitors in a similar manner, thereby reducing the overall effects on the measured output.

In earlier works, Matsumoto [97] presented a system which is also PLL-based, but does contain a sensor with electromechanical feedback within the control loop. Again, the MEMS substrate contains both sense and reference capacitors. However, the PLL structure is more involved this time: One of its inputs is derived directly from the sense capacitance, whilst the other is derived from the overflow of a binary counter running at a fixed frequency. The top value of this counter is updated periodically by measuring the reference capacitor and is therefore a function of the reference capacitance. These two signals are then input to a phase comparator and a loop filer. After this stage, the voltage output of the loop filter is proportional to the input acceleration. Crucially, this voltage is then fed back to the feedback system of the MEMS sensor, thus closing the electromechanical control loop. The general concept behind this circuit is similar to the EM-PLL, but there are number of key differences: First, the EM-PLL is a fully differential system and therefore does not require an elaborate reference frequency generation circuit to achieve the same resilience against environmental effects. Furthermore, the structure of the MEMS sensor and the feedback system in particular are significantly different: In the EM-PLL, the sensor proof mass is grounded, and the feedback plates can operate at any reasonable potential. In Matsumoto's work, the sensor proof mass was fixed to $V_{DD}/2$, severely restricting the range of the feedback voltage. Despite all these differences, this work is by far the closest to the EM-PLL that exists in literature.

Whilst the two aforementioned works are the most relevant with regards to the EM-PLL, interfacing MEMS sensor in general is an active area of research, which has seen countless different approaches to more or less related problems. For example, a pick-off and control system for resonant gyroscopes presented in [98]. There are two control loops, one to drive the resonator and another one for compensation. However, although the resonator part looks like a PLL, this approach does not utilize a frequency that is generated from a variable oscillator. There are numerous other pickoff systems for gyroscopes that work in a very similar fashion such as [99].

An interesting view from a completely opposite angle in presented in [100], which deals with the effects of vibration on frequency references and the resultant consequences in the behaviour of PLLs. Similarly, [101] looks at random vibration noise in resonant systems and how it could be compensated, also by using a dedicated accelerometer. Neither of these works are directly concerned with MEMS pickoff circuits, but they are nevertheless relevant for noise considerations. Finally, the design and stability problems with state-of-the-art EM-$\Sigma\Delta$ discussed in Section 5.2.1 are widely recognised and have resulted in research founded on the same principle as the EM-PLL, namely that linear analogue systems may provide more inherently robust closed-loop sensor systems [102].

In this section, a number of relevant interface circuits for MEMS accelerometers and other related techniques were discussed. Whilst some aspects of the EM-PLL can be found in previous work, there are generally still significant gaps in terms of overall system topology. Furthermore, it is worth noting that few works focus solely on the development of sensor

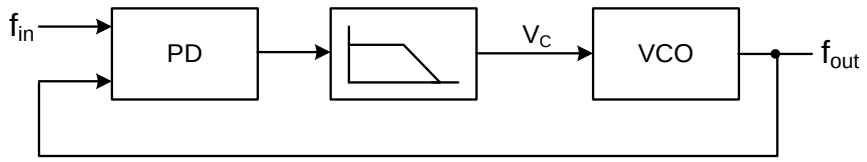


Figure 5.9: Block diagram of a basic PLL.

interface circuits, whilst the majority concentrate on the development of the sensors themselves instead. With this overview concluded, the following section will review the fundamentals of oscillators and phase-locked loops and introduce the EM-PLL.

5.3 The EM-PLL Technique

In the previous section, a number of interface circuits for force feedback MEMS accelerometers were reviewed, with a focus on Electromechanical $\Sigma\Delta$ converters. Whilst they offer good performance and an elegant system design, they are prone to performance degradation when under the influence of parameter variation. The Electromechanical PLL proposed in this thesis is intended as a replacement for EM-$\Sigma\Delta$ systems as it provides similar performance but is inherently more robust against parameter variation. Before the EM-PLL is explained in detail in Section 5.3.3, Sections 5.3.1 and 5.3.2 will briefly review the fundamentals of Phase-Locked Loops and variable-frequency electronic oscillators, which are both fundamental to understanding the EM-PLL.

5.3.1 Electronic Phase-Locked Loops

A phase-locked loop (PLL) is a versatile electronic circuit in which an output signal is controlled to have a constant phase relationship to an input signal. PLLs are used in a wide range of applications, such as frequency synthesis, demodulation or signal reconstruction. The minimal block diagram of a PLL is shown in Figure 5.9. It consists of a phase comparator, a low-pass filter and a voltage-controlled oscillator (VCO). The phase comparator takes two input signals and produces an output signal that corresponds to the phase relationship between these two signals. The VCO generates a signal whose frequency is linearly dependent on the input voltage of the VCO. The purpose of the low-pass filter is to determine the characteristics of the control loop and hence the overall performance of the PLL. In many PLL applications, such as frequency synthesis, frequency dividers are incorporated in the input or feedback paths, but they a are not necessary for the description of fundamental PLL operation.

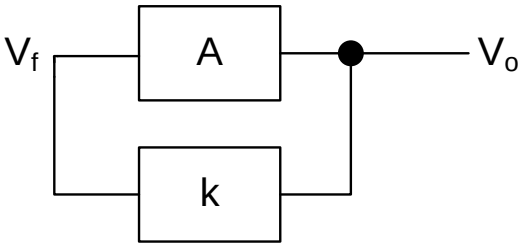


Figure 5.10: Block diagram of an oscillator

The operation of a PLL in controlled closed-loop operation can be explained intuitively. If the VCO and input signal are at exactly the same frequency and phase, the output of the phase detector is zero, thus leaving the VCO input at its quiescent voltage which corresponds to an output signal of the same phase and frequency as the input signal. If the phase of the input signal changes, the phase detector will produce an output that briefly changes the frequency of the VCO until both signals are in phase again, at which point the VCO continues to operate at its nominal frequency. Similarly, a difference in frequency also results in an ever increasing difference in phase, which results in a permanently changed output frequency of the VCO. In this case, the control voltage of the VCO tracks with the input frequency and can therefore be used to measure the input frequency.

5.3.2 Electronic Oscillators

As has been touched upon already in the preceding description of PLLs, an integral part of a phase-locked loop is a voltage-controlled oscillator. An electronic oscillator in general consists of an amplifier and a feedback network, as illustrated in Figure 5.10. A necessary condition for this arrangement to function is the Barkhausen criterion, which states that the open-loop gain through amplifier and feedback network must be unity, and the phase shift must be an integer multiple of 2π [103]. In other words, such a configuration can sustain oscillation if the input signal (V_f) to the amplifier is exactly its output signal V_o.

There are uncountable different types of electronic oscillators, ranging from well-known electronic circuits, such as the Pierce oscillator, to highly specialised arrangements, such as Reflex Klystrons. For the majority of mainstream electronic applications, the available oscillators can be categorised by their feedback networks. One category is resonant oscillators, which includes LC and crystal oscillators, while the second major category is RC oscillators. They all operate on the same principle of oscillation occurring due to specific phase and amplitude conditions, but differ in how this condition is attained. In a LC or crystal oscillator, the oscillation criterion is determined by the resonant element, e.g. the impedance of a parallel LC tank circuit at resonance. In contrast, there is

no physical resonance that determines the oscillation frequency of RC oscillator, but instead a frequency at which oscillation conditions in terms of phase and amplitude are met. From a systems perspective, a resonant oscillator is a second-order system, whereas a RC oscillator is one or more cascaded first-order systems. This fundamental physical difference is expressed in the frequency of oscillation. Regardless of the circuit used, in a LC oscillator the oscillation frequency is always proportional to $\frac{1}{\sqrt{LC}}$, where L and C are the inductance and capacitance of the tank circuit. For an RC oscillator, the frequency is always proportional to $\frac{1}{RC}$, where R is the resistance and C the capacitance of the feedback network. It is important to note that in most practical circuits, these expressions will contain additional terms, e.g. sums of capacitance values or parasitics. However, the fundamental relationships of inverse square root and inverse remain unchanged.

Analogue voltage-controlled oscillators are fundamentally no different from oscillators discussed above and often use the same circuits. Frequency control is usually achieved by making one of the frequency-determining components tunable, which is typically a capacitor or resistor. In discrete circuits, common ways to tune capacitance are varactor diodes or other junction devices, whilst resistance is commonly tuned by FET channel resistance. In integrated circuits, VCOs are often based around constant current sources. For instance, the classic LM331 uses a constant current source and a capacitor as the time base, and the control voltage as a threshold on the capacitor voltage [104]. Modern CMOS VCOs are often based on ring oscillators, where a the bias current of a ring oscillator is adjusted, causing a change in switching speeds and therefore a variable oscillation frequency [105].

5.3.3 EM-PLL Description and Analysis

In most conventional electromechanical accelerometer circuits, the change of capacitance of the accelerometer due to acceleration is sensed through signal amplitude. For example, in the EM-ΣΔ system described in Section 5.2.1, a high-frequency carrier is injected onto the common point of the sense capacitors and senses using a charge amplifier, which results in an amplitude-modulated voltage. This voltage is then demodulated and processed further and eventually used to keep the feedback loop under control.

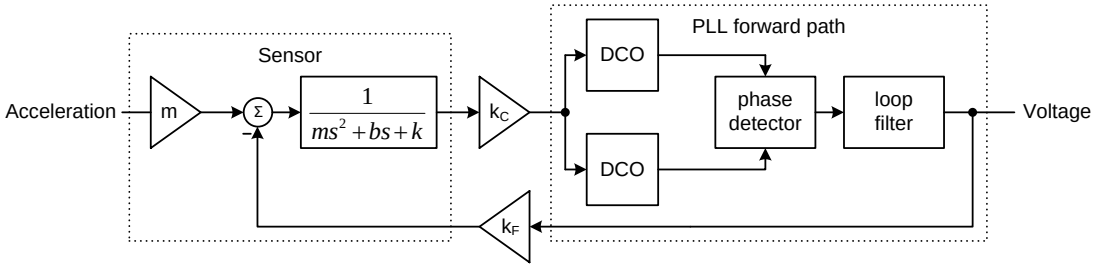


Figure 5.11: Architecture of the EM-PLL.

In contrast, the EM-PLL concept uses frequency modulation to sense the accelerometer capacitance. A block diagram of the EM-PLL is shown in Figure 5.11. Conceptually, it works as follows: Like before, acceleration results in a force on the proof mass, which is displaced by a proportional distance in the spring-mass-damper system of the MEMS accelerometer. This displacement causes a change in the two differential sensor capacitors. The two differential sensor capacitances determine the frequency of two oscillators, which are labelled DCO (displacement-controlled oscillator) in Figure 5.11. The name displacement-controlled oscillator is in reference to a voltage-controlled oscillator (VCO), but indicates that the oscillation frequency is determined by the displacement of the accelerometer's proof mass rather than a control voltage. Since the capacitors change differentially when the sensor is subject to acceleration, the oscillator frequencies change likewise. The difference in frequency between the two oscillators is thus a measure for the input acceleration. This principle is illustrated in Figure 5.12. The outputs of the oscillators are input to a phase detector and then to a loop filter which suppresses the carrier frequency and adjusts overall loop performance. At this stage, the output voltage is proportional to the force exerted in the accelerometer proof mass and the system could be used in an open-loop configuration, much like the EM-ΣΔ. For closed-loop PLL operation, the output signal is appropriately conditioned and applied to the feedback plates of the MEMS sensor, thus achieving closed-loop proportional control of the proof mass. The advantage of closed-loop operation is the same as for an EM-ΣΔ system, namely an increased input acceleration range.

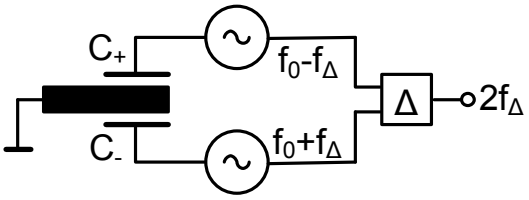


Figure 5.12: Frequency difference as a measure for acceleration.

Effectively, this entire system behaves like a phase-locked loop, where one of the sensor oscillators represents the reference oscillator and the second oscillator is equal to the feedback controlled VCO. In this configuration, the EM-PLL shows several advantages over ΣΔ modulators or conventional pick-off circuits. First, no linear low-noise charge amplifier is required, which simplifies the circuit configuration in one of the most critical parts. In contrast to the ΣΔ System, the order does not need to be increased to gain SNR, and this has the secondary effect of making the system stability inherently more robust against parameter variation. Finally, there is no digital switching of feedback voltage, which has a positive impact on the noise in the system. However, the most important fact about the EM-PLL is that is is linear to a large extent, which makes linear analysis and design valid for a wide range of conditions, unlike the way EM-ΣΔ

systems behave. This further contributes to the inherent robustness of the EM-PLL, as it can maintain its general operation and performance over a wider range of parameters.

Having just described the EM-PLL conceptually, it can now be analysed in more detail to better understand its operation. The force (F) experienced by the proof mass (m) in the accelerometer under acceleration (a) is:

$$F = ma \tag{5.8}$$

and the resulting displacement of the proof mass is that of a damped second-order mass-spring-damper system:

$$d = \frac{1}{ms^2 + bs + k} F = k_{sens}(s)a \tag{5.9}$$

with

$$k_{sens} = \frac{m}{ms^2 + bs + k} \tag{5.10}$$

where m is the proof mass, b is the damping factor, k the spring constant of the mass-spring system that models the accelerometer and s is a complex number in the Laplace space. Recalling the generic capacitance of a parallel plate capacitor as an approximation for the sensing capacitance:

$$C = \epsilon_0 \epsilon_r \frac{A}{d} = k_{cap} \frac{1}{d} \tag{5.11}$$

with

$$k_{cap} = \epsilon_0 \epsilon r A \tag{5.12}$$

where A is the sensor capacitor plate area and d the spacing between the plates. Substituting the instantaneous displacement (d) from equation 5.9 around the nominal capacitor plate spacing (d_{nom}), the sensor capacitance can be written as a function of acceleration (a):

$$C = \epsilon_0 \epsilon_r \frac{A}{d_{nom} \pm d} = k_{cap} \frac{1}{d_{nom} \pm k_{sens} a} \tag{5.13}$$

It is desirable to have a linear relationship between frequency and acceleration, which requires the oscillator frequency to be inversely proportional to capacitance. As reviewed in Section 5.3.2, RC oscillators behave in such a manner and therefore, if an RC oscillator is used for the EM-PLL's DCO, the desired behaviour can be achieved. An LC oscillator would generally be a better choice from the point of view of low noise design,

but as its frequency depends on the inverse square root of the capacitance, the resultant relationship between proof mass displacement and frequency would be non-linear. Thus, the frequency of a single DCO is:

$$f_{osc} = k_{osc}\frac{1}{C} \tag{5.14}$$

The parameter k_{osc} relates the frequency to capacitance, depending on resistance values and switching thresholds. As can be see in equation 5.13, the capacitance is not exclusively dependent on the inverse of the acceleration, but also on the nominal capacitor plate spacing d_{nom}. However, this constant term, which may well be larger in magnitude than the change in capacitance due to acceleration, cancels when two oscillators are driven from a pair of differential capacitors (in the symmetrical structure shown in 5.11), which are subject to the same acceleration. The frequency difference of two such oscillators is:

$$\Delta f = f_{+d} - f_{-d} = \frac{k_{osc}}{k_{cap}}\left(d_{nom} + d - d_{nom} + d\right)$$
$$\Delta f = 2\frac{k_{osc}}{k_{cap}}d \tag{5.15}$$

As was shown in Equation 5.9, the displacement can be replaced with a function of acceleration and as a result the frequency difference can be seen in 5.16 to be linear with respect to acceleration:

$$\Delta f = 2\frac{k_{osc}}{k_{cap}}k_{sens}(s)a \tag{5.16}$$

If it is furthermore assumed that the phase detector provides a voltage output that is linear with frequency difference at its input, the resultant output voltage in the forward path will hence be linearly dependent on acceleration.

Unlike the bitstream of an EM-ΣΔ system, this output voltage cannot be applied directly to the feedback plates to achieve closed loop operation. The reason for this is that the electrostatic force between the two feedback plates is only attractive, regardless of the polarity of the feedback voltage. It is therefore necessary to apply a certain bias voltage to both feedback plate pairs, causing oppositely directed forces on the proof mass resulting in no net force when under quiescent conditions. This technique is known as force-balancing. The feedback voltage is then superimposed on the bias voltage, resulting in an increase of one and decrease of the other force, resulting in a net feedback force on the proof mass, as depicted in Figure 5.13.

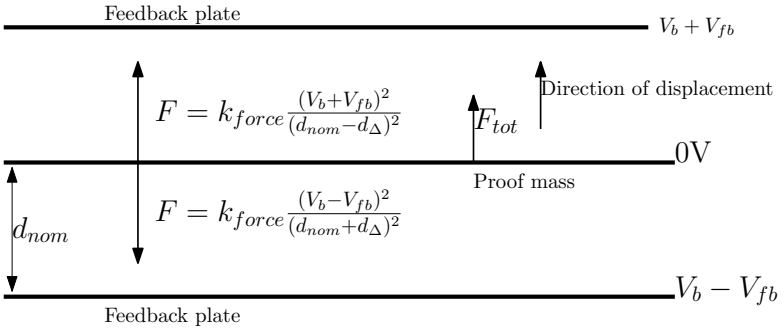


Figure 5.13: Forces in the force-feedback system.

Like the forward path, the feedback path can also be analysed as follows. The electrostatic force between two parallel plates is given by Equation 5.17:

$$F = k_{force}\frac{V^2}{d^2} \tag{5.17}$$

where:

$$k_{force} = \frac{1}{2}\epsilon_0\epsilon_r A \tag{5.18}$$

The total force on the proof mass is the sum of two oppositely directed electrostatic forces, each depending on the bias voltage and the differential feedback voltage as given in Equation 5.19:

$$F = k_{force}\left(\frac{(V_b + V_{fb})^2}{(d_{nom} - d_\Delta)^2} - \frac{(V_b - V_{fb})^2}{(d_{nom} + d_\Delta)^2}\right) \tag{5.19}$$

$$F = k_{force}\left(\frac{V_b^2 + 2V_bV_{fb} + V_{fb}^2}{(d_{nom} - d_\Delta)^2} - \frac{V_b^2 - 2V_bV_{fb} + V_{fb}^2}{(d_{nom} + d_\Delta)^2}\right) \tag{5.20}$$

Due to the dependence of the force on the instantaneous spacing of the plates, the relationship between feedback force and feedback voltage is not linear. However, at this point it is important to note that normally, the displacement of the proof mass is actively controlled by the feedback system to be as small as possible. Assuming that the loop is closed and under control, the displacement d_Δ is much smaller than the plate spacing d_{nom}, which means that Equation 5.19 can be reduced to Equation 5.21, where the V_{fb}^2 terms have cancelled:

$$F \approx k_{force}\frac{4V_bV_{fb}}{d_{nom}^2} \tag{5.21}$$

This means that the entire closed-loop system of the EM-PLL is approximately linear, which is a very significant result. Unlike electromechanical ΣΔ modulators which are inherently non-linear and therefore require difficult analysis methods, initial design calculations and simulations on the EM-PLL can be achieved using a linear approximation. Once the system properties and parameters have been determined roughly using the linear approximation, final tuning can be conducted using the full non-linear models, which also give the most accurate performance metrics.

An example for these non-linearities featured in the full EM-PLL model can be seen in Equation 5.19: In addition to non-linearity introduced by the instantaneous displacement, the feedback voltage also appears as a squared term in the numerator. This causes distortion of the signal, which causes harmonics in the signal spectrum, as will be shown in Section 5.4. ΣΔ systems also show the same harmonic, but there an adaptive feedback voltage is usually used to reduce this harmonic component. In the EM-PLL, reduction of this component can be achieved more easily and without any additional circuitry, as can be seen from Equation 5.19: If the feedback plate bias voltage is increased to be relatively large with respect to the feedback voltage, the feedback force is dominated by the constant V_b^2 and linear $2V_bV_{fb}$ terms. Therefore, increasing the feedback plate bias voltage reduces the non-linear contribution and thus the contribution of the harmonic components in the output signal spectrum.

5.4 Simulation Analysis

5.4.1 Introduction

In order to evaluate the performance of the EM-PLL and its resilience against parameter variation, three standard tests have been used to compare the new EM-PLL circuit with a reference EM-ΣΔ system. The first test is to calculate Power Spectral Density (PSD) of both the EM-PLL and EM-ΣΔ circuits. The main criteria at this stage is to establish whether the basic performance is comparable between the two systems. Furthermore, this test will also quantify the effects of non-linearities in both systems. The second test is to compare the response of the two circuits to a wide range of accelerations. This primarily serves to establish the operating range, but will also illustrate each system's response when the input acceleration moves away from the nominal design value. The final test is to evaluate the impact of parameter variations on the two circuit configurations, using equivalent basic parameter tolerances, to determine the resultant variation in circuit performance.

All simulation in this sections were carried out using Matlab/Simulink. The models are direct implementations of the block diagrams of Figure 5.9 and 5.6 and can be found in Figures F.1 and F.2 in Appendix F. The system model of the EM-ΣΔ was adapted from

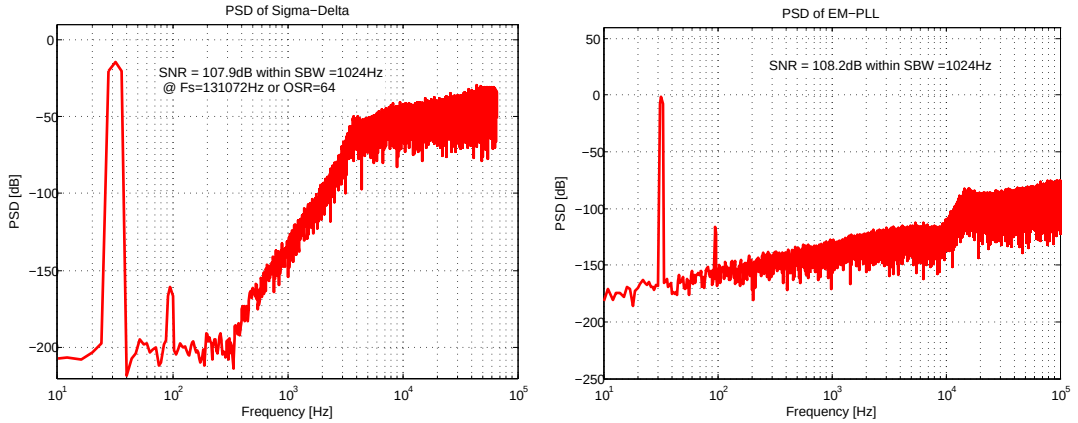
[95]. Tables F.1, F.2 and F.3 in Appendix F show the system parameters used for the EM-PLL and EM-ΣΔ simulations. The optimised parameters were obtained using the Cheetah GA system [95], with a nominal sinusoidal input of 5g at 32Hz. This process of finding an optimised parameter set is an example of automated model-based design, as described in Section 2.3.3. For both the EM-PLL and the EM-ΣΔ, the optimiser was provided with the system models, a fitness function and appropriate parameter ranges. From there, the genetic algorithm was used to find a number of viable parameter sets, which where then subjected to variation analysis and the most robust chosen, as described in Section 2.3.3.2.

The system models themselves are straightforward: For the EM-PLL, the DCOs and phase detector operate in the phase domain, reducing them to integrators and a summing junction, respectively. Noise is modelled as additive phase noise at the output of the phase comparator, the magnitude of which has been determined experimentally by measuring the phase noise of a discrete RC CMOS oscillator. In the EM-ΣΔ model, which contains a 3^{rd} order ΣΔ modulator, the pickoff amplifier and all other analogue components are assumed noiseless, with the only noise being quantisation noise. None of the models implement full resolution transient behaviour, i.e. they do not model the carrier or base frequencies. Hence, in the EM-PLL the combination of DCO, phase detector and loop filter is assumed to provide a perfectly filtered signal, whilst in the EM-ΣΔ a similar assumption is made for the pick-off circuit and demodulator.

5.4.2 Output Power Spectrum

The first simulation is to compare the output power spectrum of the EM-PLL and the EM-ΣΔ. This is useful for several reasons. First, the spectrum will immediately show any harmonic components which are due to non-linearities in the system, which will give an indication of the overall system performance and signal fidelity. Furthermore, the resultant signal-to-noise ratio is a direct measure for system performance and achievable resolution. Finally, qualitative observations can be made about the behaviour at frequencies outside the signal bandwidth, which is of particular interest in a noise shaping system, such as the EM-ΣΔ. All of these simulations were again carried out with the nominal input signal of 5g at 32Hz.

The output power spectra of the EM-PLL and the EM-ΣΔ were obtained in the same manner. With the input signal applied, 32 cycles of the output signal were captured after a settling period of 3 cycles. Then, the output signal was windowed and its FFT taken, which results in the output spectrum. The SNR was calculated by relating the power in the signal bin to the sum of the remaining bins within the nominal signal bandwidth of 1024Hz.



(a) PSD of EM-ΣΔ (b) PSD of EM-PLL

Figure 5.14: Comparison of PSD between EM-ΣΔ and EM-PLL

Figure 5.14 shows a side-by-side comparison of the output power spectra of the EM-PLL in Figure 5.14b and the EM-ΣΔ in Figure 5.14a. The EM-ΣΔ system achieves as SNR of 108.5dB, which is a typical value for such a system [106, 107]. In comparison, the EM-PLL achieves a SNR of 108.2dB, which is practically identical. It must again be noted at this point that in the case of the EM-ΣΔ, the only noise is quantisation noise. If real amplifier or integrator noise was considered in the EM-ΣΔ, the resultant SNR would be lower.

At this point, it should be noted that in each case, the dominant contribution to non-signal components within the signal bandwidth is the harmonic at 96Hz. This component is due to non-linearities in the feedback system, as was explained in Equation 5.19. Since the EM-ΣΔ features an adaptive adjustment of feedback voltage to linearise this function, the relative magnitude of this component in the EM-ΣΔ spectrum is lower compared to the EM-PLL. The EM-PLL does not have such a mechanism and therefore the harmonic component will increase with signal level, reducing the SNR at higher signal amplitudes, as will be shown in Section 5.4.3. However, this harmonic can still be adjusted in the EM-PLL through the feedback plate bias voltage, as was described along with Equation 5.21.

A further comparison can be made between the general shape of the spectra. In the case of the EM-ΣΔ, the spectrum follows the general shape of a ΣΔ modulator in that it clearly shows noise shaping, as discussed in Section 5.2.1.2. Within the signal bandwidth, the noise floor is almost flat, but then rises sharply when approaching the upper end of the signal bandwidth, which is the expected behaviour of a ΣΔ modulator. The EM-PLL, on the other hand, has a noise floor that slowly rises with frequency, indicating a low-frequency zero in its noise transfer function. There is also a sharp rise in the noise floor at higher frequencies, but this is not due to deliberate quantisation noise shaping. Instead, it is merely the result of the interaction between the loop filter and sensor transfer functions.

5.4.3 Input Amplitude Sweep

As was pointed out when introducing ΣΔ modulators in Section 5.2.1.2, one significant problem of higher-order ΣΔ systems is their sensitivity to parameter variation. This does not only cover intrinsic parameters, such as gains, but also external ones, such as the input signal amplitude. In order to illustrate this property, an input amplitude sweep was performed for both the EM-ΣΔ and the EM-PLL and the SNR within the signal bandwidth measured. The results are shown in Figure 5.15.

As the input signal amplitude increases, the output SNR increases in both EM-ΣΔ and EM-PLL. This is because there is a constant noise floor in both systems, and increasing the amplitude raises the output signal with respect to the noise floor, leading to an improved SNR. However, as the amplitude increases further, the behaviour of EM-ΣΔ and EM-PLL differs: The EM-ΣΔ becomes abruptly unstable for accelerations above approximately 7g. This is because the amplitude has moved out of the relatively small operating region and the control loop has become unstable. There is no signal component left in the output spectrum, which is reflected in the SNR. The EM-PLL, on the other hand, continues to gain SNR as the input amplitude increases. At some point, the SNR reaches a maximum and then begins to decrease again slowly. This is due to the harmonic component becoming larger as input amplitude increases, which reduces the SNR. In simpler terms, the non-linear components of the system distort the signal more at larger signal amplitudes. This decrease of SNR continues until approximately 45g, at which point the EM-PLL also stops operating.

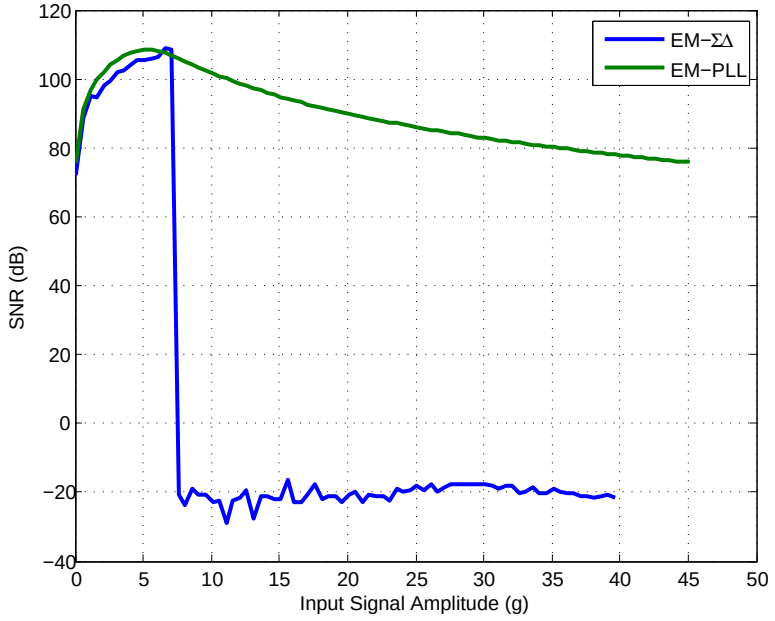


Figure 5.15: Input amplitude sweep for EM-ΣΔ and EM-PLL

This "hard" limit for the EM-PLL has different, better defined reasons than the one of the EM-ΣΔ: In closed-loop control, the accelerometer can be subject to much larger accelerations than without feedback because the feedback force on the proof mass counteracts the force due to acceleration. However, the feedback force is limited in amplitude by the plate bias voltage. If the acceleration results in a force that is larger than what can be compensated by the feedback force, the proof mass can no longer be controlled and the EM-PLL ceases to function. Unlike ΣΔ modulators, this limit is only due to the bias voltage and can be calculated through the previously established equations. Rearranging equations 5.17 and 5.8 for acceleration under the condition where $V_{fb} = V_{bias}$ and zero instantaneous displacement results in Equation 5.22:

$$a_{max} = \frac{k_{force}}{m} \frac{(2V_{bias})^2}{d_{nom}^2} \tag{5.22}$$

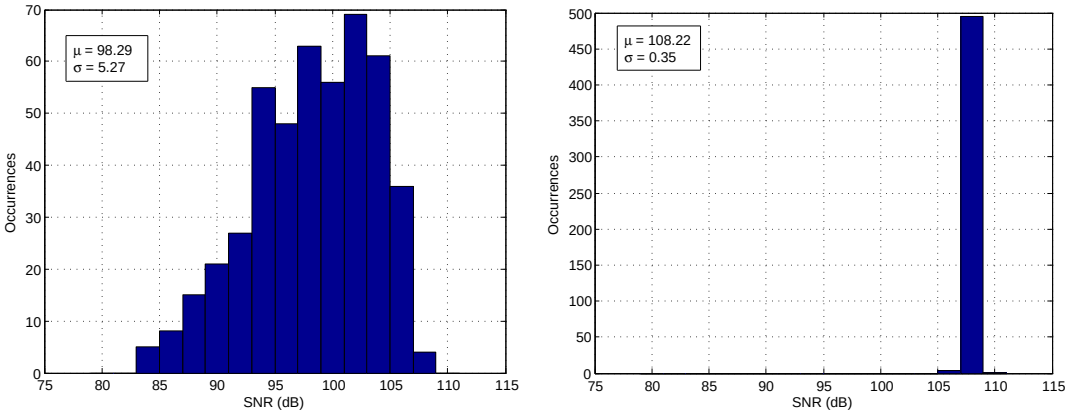
Evaluating equation 5.22 for nominal sensor parameters and a bias voltage of 35.2V results in a maximum acceleration of $447.0ms^{-2}$, or 45.6g. This is in very close agreement with the simulations, where the maximum input amplitude is 45g. Note that since the bias voltage influences this term quadratically, this figure drops significantly for lower bias voltages. For example, at 12V bias the maximum acceleration the system can handle is reduced 5.3g.

This calculation also illustrates once more the key advantage of the EM-PLL: Since it is approximately a linear system, algebraic expressions for system properties such as the maximum acceleration can be found very easily. This is very significant, as basic system characterisation and design (e.g. determining the required plate bias voltage for a given acceleration) can be done analytically. This is in stark contrast to EM-ΣΔ systems, where practically all system parameters and properties have to be obtained through simulations.

5.4.4 Parameter Variation

The final simulation is perhaps the most interesting one from the standpoint of this thesis. Up to now, it has been established that the EM-PLL can provide similar performance to EM-ΣΔ systems, while hinting at its resilience against parameter variation. In order to ascertain whether this is indeed the case, a simulation similar to Monte Carlo analysis is performed. Comparing the resultant distributions of system performance and yields will then allow quantification of the EM-PLL's robustness.

In order to carry out these simulations, each system is simulated 500 times and the resultant SNR recorded. In each iteration, the system parameters are varied by a random amount, as indicated in the parameter tables in Appendix F.1. This is supposed to simulate the real-life variations due to fabrication tolerances or environmental effects



(a) EM-ΣΔ (b) EM-PLL

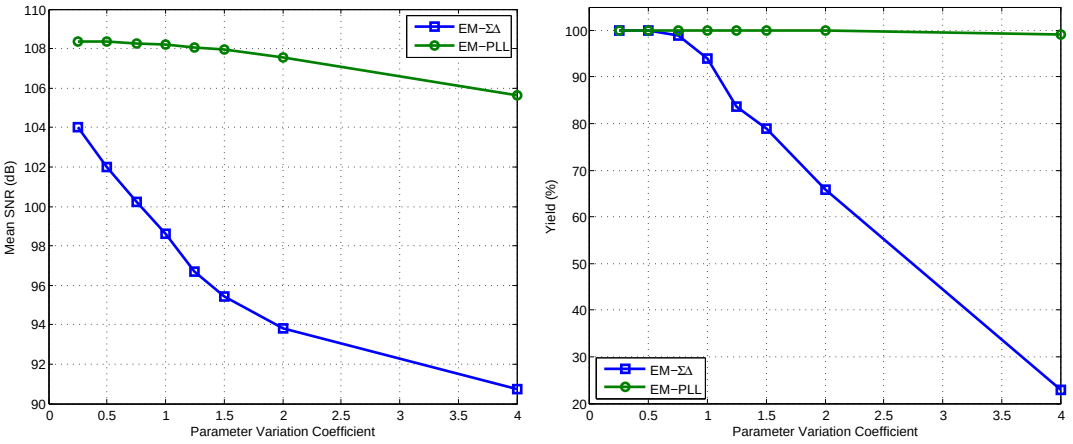
Figure 5.16: SNR histograms of EM-ΣΔ and EM-PLL under parameter variation.

that the system could experience. The values for variations are based on reasonable assumptions for tolerances of electronic components and mechanical tolerances in the MEMS fabrication process. When processing the results, a pass SNR threshold of 80dB has been chosen. Any systems that achieved less than this were excluded from the final histogram and counted as "not working". This excludes both unstable EM-PLL and EM-ΣΔ systems. This 80dB threshold is the same as was used during the GA optimisation process and it was ensured that no working systems with a low SNR were excluded by the threshold. The histograms of the final SNRs are shown in Figure 5.16.

It can be seen immediately that the histogram of the EM-PLL in Figure 5.16b is much narrower than the one of the EM-ΣΔ in Figure 5.16a. The standard deviation of the SNR for the EM-PLL is 0.35dB with a mean of 108.22dB, corresponding to 0.32% of the mean. For the EM-ΣΔ, the standard deviation is 5.27dB with a mean of 98.29dB, corresponding to 5.36% of the mean. From these numbers alone it can be seen that the EM-PLL is significantly more robust against parameter variation, while at the same time providing slightly better performance.

Even more significant is the number of systems that passed the SNR threshold. Out of 500 systems, all EM-PLL systems and 468 EM-ΣΔ systems passed the SNR threshold, corresponding to yields of 100% and 93.6%, respectively.

To further demonstrate the robustness against parameter variation of the EM-PLL, the Monte Carlo simulation has been repeated with varying levels of error on the parameters. Parameter variation coefficients (p) between 0.25 and 4 were used to scale the original tolerances on the parameters given in Appendix F. For example, the nominal variation on the proof mass is 2%. When a parameter variation coefficient of 2 is used, the Monte Carlo parameter set is generated with variation on the proof mass of 4%.



(a) Mean SNR (b) Yield

Figure 5.17: Mean SNR and yield for different levels of parameter variation.

Figure 5.17a shows the behaviour of the mean output SNR for both EM-PLL and EM-ΣΔ as parameter variation increases. As expected, the mean drops in both cases as performance degrades due increased parameter variation. However, the performance of the EM-ΣΔ degrades much more quickly compared to the EM-PLL. Also note that the EM-PLL is able to keep its nominal SNR of approximately 108dB for small values of parameter variation, whereas the EM-ΣΔ degrades significantly from its nominal performance for even the smallest errors. As the SNR drops, so does the yield, as shown in Figure 5.17b. Both the EM-PLL and the EM-ΣΔ keep a high yield initially. At larger parameter variations, the yield of the EM-ΣΔ drops dramatically to a minimum of 22.8%, whereas the yield of the EM-PLL only drops to 99.0% for the highest parameter variation.

The combination of the low yield and low mean SNR of the EM-ΣΔ has an interesting consequence on the SNR standard deviation of the EM-ΣΔ. Recall that the pass threshold in terms of SNR for a given system was set to 80dB, which meant that with $p = 1$, no working systems were falsely excluded from contributing to the yield and final standard deviation. However, for EM-ΣΔ systems at higher p, the mean SNR moves closer to the 80dB threshold whilst the standard deviation increases. This means that as parameter variation increases, systems are more likely to be below the 80dB threshold although they might be functional with lower performance. Due to the low yield, even a small number of incorrectly categorised systems can significantly skew the standard deviation of the systems above the threshold, which can be seen in Figure 5.18. The SNR standard deviation of the EM-PLL increases monotonically with parameter variation, as expected. However, the SNR standard deviation of the EM-ΣΔ starts to drop slightly for larger parameter variation. This might give the false impression that the variation tolerance decreases, whereas it is in fact only an artefact due to the distribution of working systems being cut off, resulting in a lower apparent standard deviation. When the threshold is lowered, more working systems are included in the standard deviation measurement

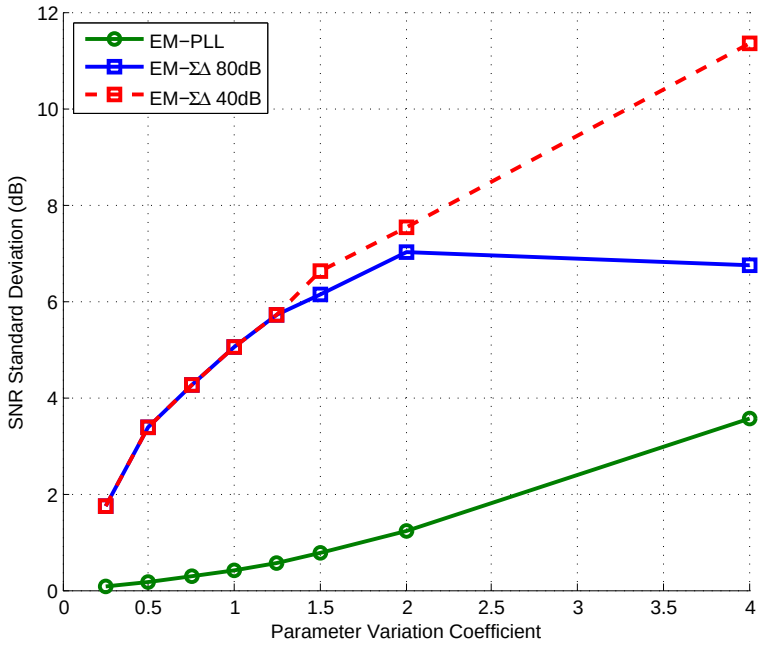
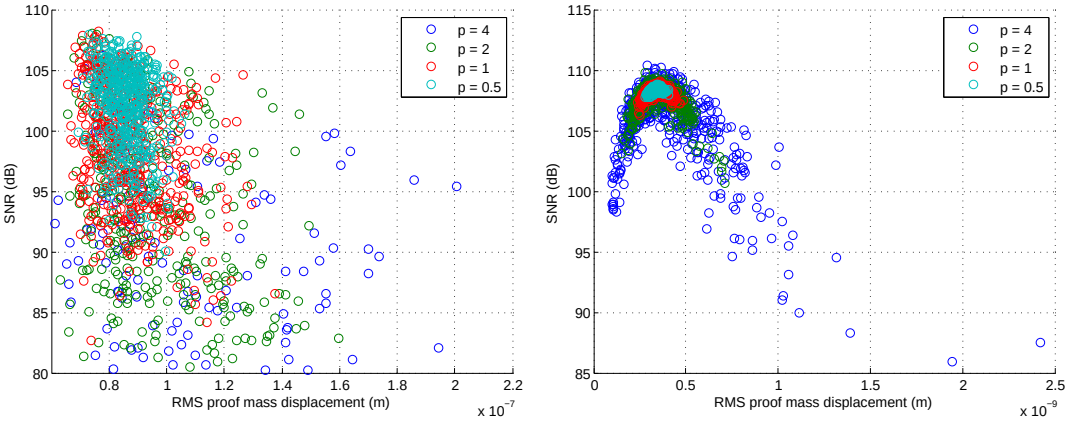


Figure 5.18: SNR standard deviation for different levels of parameter variation.

and the standard deviation increases with parameter variation, as expected. Whilst the change in standard deviation by lowering the threshold is significant, the yield only increases marginally, e.g. to 33.2% for $p = 4$, the reasons for which will be explained later. Lowering the threshold is of course only permissible for the purpose of illustration since the pass threshold would typically be fixed in a production system.

A further performance metric of force-feedback accelerometers that has not been considered so far is the proof mass displacement. Smaller displacement for a given acceleration indicates better loop servo characteristics and therefore more overall robustness. The RMS of the proof mass displacement has been recorded for every system during the Monte Carlo simulation. Figure 5.19b shows a scatter plot of SNR against RMS displacement for the EM-PLL for different parameter variation coefficients. The most striking observation that can be made is that the correlation between SNR and displacement follows an arc, which extends as parameter variation increases. For small variation, most systems cluster around the peak of the arc, which is the nominal design point. As parameter variation increases, there are some systems which have a lower proof mass displacement, and those systems also have a tendency for lower SNR. Similarly, systems with higher than nominal proof mass displacement also tend to have a lower SNR. The latter can be explained by the general degradation of system performance with parameter variation, which can result in both lower SNR and poor loop control, resulting in high proof mass displacement. The relationship for small displacement is most likely due to the fixed phase noise of the oscillators. If parameter variation causes a system to have a small proof mass displacement, the phase difference between the differential oscillators is also



(a) EM-ΣΔ (b) EM-PLL

Figure 5.19: Scatter plots of SNR and RMS displacement of the Monte Carlo simulation.

small. However, the phase noise of the oscillators is a fixed quantity, and therefore becomes relatively large compared to the signal in these cases. Therefore, the overall SNR is lower, despite the lower proof mass displacement and ostensibly better loop control. The interaction between these two effects results in the arc-shaped relationship. This property is in fact very beneficial to SNR robustness against parameter variation: As the slope of the apex of an ideal arc tends towards zero, there would be very little sensitivity to variation at that point. In principle this is similar to the temperature relationship of a bandgap reference discussed in Section 2.3.2, which also has zero temperature coefficient at its nominal operating point.

Figure 5.19a shows the same scatter plot, but for the EM-ΣΔ. Note that this time no clear visual correlation exists between the SNR and the displacement. The system metrics merely spread out in both dimensions as parameter variation increases. It is also apparent that the proof mass RMS displacement of the EM-ΣΔ is in general two orders of magnitude larger than the case of the EM-PLL. This indicates that the pick-off circuit of the EM-ΣΔ is less sensitive and therefore required larger displacements to achieve the same output SNR. This plot also illustrates the aforementioned issue of the distribution being cut off by the 80dB threshold: For $p = 0.5$ and $p = 1$ all systems are well away from the threshold and it is intuitively very unlikely that there are any working systems below the 80dB threshold. However, for $p = 2$ and $p = 4$, there are several systems that are very close to the threshold and it is more likely that some systems with an SNR of just below 80dB were discarded. Also note that the absolute number of systems for higher variation decreases, which is directly represented by the reduced yield. This also illustrates why lowering the threshold significantly increased the SNR standard deviation, but has little effect on the yield: Lowering the threshold means that a few more systems are counted as working, but the absolute number is still comparatively low. Therefore, the yield is not greatly improved. However, because there were so few systems above

the threshold before and all additional systems are under the threshold, the standard deviation is affected greatly.

The simulations and comparisons carried out in this section highlight the EM-PLL's general robustness against parameter variation. In particular, the near perfect yield and performance even under high parameter variation would make the EM-PLL in principle suitable for practical implementation. In addition, the comparisons also further complete the picture of the variation behaviour of EM-PLL and EM-ΣΔ systems. Although all EM-PLL systems in this simulation were stable, it is conceivable that certain parameter combinations would lead to an unstable or very low performance system. This is almost the opposite for the EM-ΣΔ: Only parameters sets that are close to nominal result in a stable system and even then, variations on the parameters lead to large variation in performance. It is also worth pointing out that for the EM-ΣΔ, variation only causes a reduction in SNR, as can be seen from the histograms and the lower than nominal mean SNR. This again indicated that the EM-ΣΔ's nominal operating parameters are very narrow and sensitive, and any change results in worse performance. Conversely, the mean SNR of the EM-PLL is exactly its nominal SNR, meaning that variation can make the individual systems better or worse.

Once again, this difference in behaviour can fundamentally be attributed to the more linear nature of the EM-PLL, which does not result in erratic performance changes for small parameter changes for the most part. Finally, it must be noted again that, as stated in Appendix F.1, the variation on the fundamental device parameters is the same between the EM-PLL and the EM-ΣΔ, yet the EM-PLL is barely affected by these variations at all, which is reflected in the excellent standard deviation of performance and yield.

5.5 Discussion

In this chapter, the EM-PLL was introduced as a novel frequency-domain technique for closed-loop control of MEMS accelerometers. The key benefit of the EM-PLL over established techniques, such as EM-ΣΔ systems, is that it is by design significantly more resilient against parameter variation. This is achieved through a system topology which is largely linear and therefore behaves gracefully when subjected to parameter variation. Performance-wise, the EM-PLL is comparable to EM-ΣΔ systems, but accommodates a much wider input amplitude range.

Simulations have shown that for the same sensor, the EM-PLL and a 5^{th} order EM-ΣΔ circuit both provide an SNR in the order of 108dB for the same signal bandwidth, which suggests similar fundamental noise performance. However, an important difference is the ability of the EM-PLL circuit to tolerate much greater levels of acceleration up to 45g, indicating a much higher tolerance than the equivalent EM-ΣΔ circuit, which is

only capable of operating reliably up to 7g. Furthermore, the EM-ΣΔ system shows significant sensitivity to parameter variation, resulting in a large standard deviation of SNR. Compared to the EM-ΣΔ, the standard deviation of SNR of the EM-PLL is reduced by over 85%. Furthermore, the EM-PLL shows a perfect yield of 100% under parameter variation, whereas the EM-ΣΔ circuit achieves a yield of only 93.6% for the same relative variation on parameters.

Taking all of this into consideration, the EM-PLL is an excellent example of an alternative to a well-established method where susceptibility to parameter variation is a key issue. It shows how a fundamentally different design which is inherently more robust against parameter variation can not only lead to increased yield and tighter performance specifications, but ultimately also improve performance in other areas, such as dynamic range. The previous chapters of this work have been concerned with calibration techniques, which have the ultimate goal of adjusting a given system after fabrication. However, the EM-PLL shows that depending on the situation, a completely different circuit can vastly improve the variation tolerance of a system, completely removing the need for calibration altogether.

Chapter 6

Conclusions

6.1 Contributions

This thesis was concerned with select methods to reduce the effects of parameter variations on analogue and mixed-signal integrated circuits. Two main areas of such methods were explored: Calibration and inherently robust design. Calibration works by equipping a circuit with means to adjust its performance, and performing such adjustments after fabrication to remove or compensate for the effects of parameter variation. The aim of robust design is to design the circuit in way that reduces its sensitivity to variation in parameters. Thus, these two approaches aim to arrive at the same goal - a circuit where the effects of parameter variation have been reduced - through different routes.

In the area of calibration, this work contributed to the Configurable Analogue Transistor (CAT). The CAT is a system-level approach to calibration, which relies heavily on computer simulations and software tools to optimise performance. Whilst the fundamentals of the CAT had been developed already, it lacked a number of key components which were developed in this thesis. Furthermore, this thesis showed the first use of CAT in a practical circuit to prove its viability.

This work on the CAT is complemented by a contribution in the area of robust design. A novel closed-loop pickoff circuit for force-balanced MEMS accelerometers based on the principle of a Phase-Locked Loop has been developed. This novel circuit shows how an existing circuit topology which is fraught with design and stability issues stemming from its sensitivity to parameters can be replaced with an entirely different circuit which is much more robust, while at the same time not compromising performance.

The following is an explicit list of the contributions made in the thesis:

- An algorithm for Critical Device Identification for Configurable Analogue Transistors. This algorithm determines the transistors in a circuit which are most suitable

for calibration of circuit performance. It differs from conventional sensitivity analysis in that it seeks to find a set of devices which permit relatively independent adjustment of circuit performances instead of only considering the absolute magnitude of sensitivity.

- A description of how the existing device size optimisation for CAT can be used to size calibration devices in arbitrary circuits. Before this work, optimising the size of CATs was restricted to CATs that are used under constant current conditions. It was shown how the CAT sizes can be optimised using information gathered from a Monte Carlo simulation and a numeric solver.

- A method for using CATs for online temperature compensation. In addition to inherent parameter variation, CATs can also be used to compensate for the effects of temperature. Because the temperature behaviour of transistors is usually well modelled, temperature compensation can be achieved online through a lookup table, without the need for run-time performance measurements.

- The first application of CAT to a practical circuit. A segmented current-steering DAC was designed and equipped with CATs that allow the adjustment of the DAC transfer function after fabrication. Measurements on the fabricated chip proved that the CAT can be used for this purpose as intended.

- A novel interface circuit for force-balanced MEMS accelerometers. This circuit, named the Electromechanical PLL (EM-PLL) combines the functions of pick-off and force feedback circuits in a closed loop system. Unlike comparable high-performance electromechanical $\Sigma\Delta$ circuits which serve the same purpose, the EM-PLL is an approximately linear system and is therefore more inherently robust against parameter variation and easier to design.

6.2 Final Conclusions

Having summarised the contributions made in this thesis, the following conclusions can be found from the work carried out in this thesis.

First, all of the design-time methods developed for the Configurable Analogue Transistor were proven in simulation. Whilst only two example circuits were used, the methods for Critical Device Identification, device size optimisation, and online calibration performed well and will likely work in other circuits. However, there is still scope for improvement, particularly in terms of required computing time, which will be discussed in Section 6.3.

The application of CAT to a practical Digital-to-Analogue converter was also moderately successful. The CATs and the supporting design and calibration tools worked as expected

and the CATs could successfully improve the INL of the DAC. The maximum achievable improvement in INL was 17.4%, with the best absolute INL still being 198LSB. These number stem from the fact variation and particularly mismatch in the physical circuit were underestimated during the design stage, which resulted in CATs that were inadequately sized. These CATs where then not optimised for the fabricated circuit and therefore the INL improvement was not as good as expected. Nevertheless, the improvement that was achieved was consistent with expectations once the measured variation values were used. Using an automated method for optimising the CAT configuration in the DAC also completes the set of tools required in the fully automated CAT design flow, the development of which was another goal of this thesis.

Lastly, the initial work done on the EM-PLL is very promising, as it shows considerable advantages over $\Sigma\Delta$-based circuits. Simulations and analysis suggest that whilst providing similar baseline performance of approximately 108dB SNR, the EM-PLL is vastly more tolerant to parameter variation and also easier to design. For instance, in the simulations conducted in this thesis, the EM-PLL showed 100% yield and a reduction in standard deviation of SNR of 85% compared to state-of-the-art EM-$\Sigma\Delta$ systems. This makes the EM-PLL an excellent example of an inherently robust circuit that could ideally work as a drop-in replacement for existing systems.

6.3 Future Work

Based on this work, there are several specific areas in which direct further research and development can be undertaken. These areas include the following:

1. The automated critical device identification process presented in Section 3.2 could be extended. In particular, the current algorithm requires significant computing time in order to cover a sufficient portion of the design space through Monte Carlo simulations. One approach could be to apply machine learning algorithms and non-random sampling of device parameters to more efficiently cover the design space and gather more detailed information about regions of interest.

2. Whilst the online calibration scheme of Section 3.4 has been simulated, it has yet to be proven in silicon. For this, a chip could be designed that includes a small analogue circuit equipped with CAT and temperature measurement facility. The ability of CAT to compensate for temperature variation could then be evaluated by subjecting the chip to extreme temperatures and applying the online calibration algorithm.

3. Although the DAC of Chapter 4 proved the ability to improve circuit performance through CAT, the overall performance after calibration did not match other published work. In order to determine whether the CAT can indeed compete with

other schemes that improve DAC linearity, another chip should be designed which incorporates all the knowledge gained from designing this chip. It would also be possible to choose another example circuit altogether. If a smaller analogue circuit were chosen, the CDI tool could also be used on the whole circuit during the design process and hence the entire CAT design flow demonstrated.

4. So far, the EM-PLL of chapter 5 has been characterised in simulation, but it has not yet been tested in practice. In order to prove its viability, an EM-PLL circuit should be built, characterised an compared against other accelerometer pick-off circuits. Additionally, the current EM-PLL loop topology may not be ideal. Further research and analysis in this are could be conducted.

Appendix A

Publications

The following pages contain the paper "Critical Device Identification for Configurable Analogue Transistors", which was published at DATE 2012. It is based on the work presented in Sections 3.2 and 3.3.

Automated Critical Device Identification for Configurable Analogue Transistors

Robert Rudolf, Pouya Taatizadeh, Reuben Wilcock and Peter Wilson
Electronics and Electrical Engineering Group, School of Electronics and Computer Science
University of Southampton, UK
{rr1v07, pt4v07, rw3, prw}@ecs.soton.ac.uk

***Abstract*—A novel approach is proposed for analogue circuits that identifies which devices should be replaced with configurable analogue transistors (CATs) to maximise post fabrication yield. Both performance sensitivity and adjustment independence are considered when identifying these critical devices, giving a combined weighted sensitivity. The results from an operational amplifier case study are presented where it is demonstrated that variation in key circuit performances can be reduced by an average of 78.8% with the use of only three CATs. These results confirm that the proposed critical device selection method with optimal performance driven CAT sizing can lead to significant improvement in overall performance and yield.**

Keywords-configurable analogue transistor; optimal sizing; device variability; sensitivity analysis, post fabrication calibration

I. INTRODUCTION

A. CMOS Scaling

Maintaining production yield at smaller process nodes raises significant challenges due to device variability [1][2]. In analogue and mixed-signal circuits, the resulting performance degradation can be so severe that some form of post-silicon adjustment is necessary [3]. Early approaches concentrate solely on the adjustment of a single device to improve circuit performance [4]. This is often impractical in more complex systems that require multiple adjustment points or a higher level of integration [5]. In contrast, electronic trimming methods such as the use of floating gates [6][7] or substrate biasing [8] to alter the transfer characteristics of MOS transistors allow higher integration at lower cost. Furthermore, a wide range of simple digital trimming techniques exist, e.g. configurable arrays of MOS devices [9] or capacitors [10].

System-level digital methods have also been proposed, where circuit errors are corrected in software [11] or in the analogue domain by reconfiguration, e.g. switching-sequence post adjustment for data converters (SSPA) [12].

B. The Configurable Analogue Transistor (CAT)

The calibration methods described in the previous section tend to be targeted at either fairly specific circuit applications or effects (e.g. floating gates). The choice of which devices to make adjustable is conventionally made early in the design stage and is based on the type of circuit, the technology available and the anticipated main sources of variability. The configurable analogue transistor (CAT) introduced in [13] provides a calibration platform that is independent of the target circuit and the specific mechanisms of variation. This is achieved by considering system-level performance and replacing specific transistors in the circuit with CATs. The number of CATs represents a trade-off between increased circuit complexity and yield improvement. The structure of a CAT is shown in Figure 1. There is a main device M_0 and n calibration devices M_1 to M_n, selected through n digital control lines, B_1 to B_n, resulting in a total of 2^n discrete widths. In contrast to previous digitally adjustable analogue circuits, the CAT methodology includes a unique optimal sizing process [13]. The CAT configuration can be altered at any time after fabrication either as a one-time post-fabrication process, or to dynamically calibrate circuits to compensate for environmental effects or ageing [14].

C. Critical Device Identification

In principle, any transistor in a given circuit could be replaced by a CAT. A designer could manually identify devices that would benefit from calibration based on experience and their understanding of the operation of the circuit, however this becomes more difficult as the circuit complexity increases.

In order to facilitate optimal performance gain from the application of CATs, an automated method of Critical Device Identification (CDI) is an integral part of the approach. The task of CDI is to identify a number of transistors, which when adjusted allow the performances to be tuned after fabrication. This paper proposes a novel approach to critical device identification which is fully automated and independent of circuit type. In addition, the proposed method also optimally sizes the CATs to minimise performance variation.

D. Paper Structure

The remainder of this paper is structured as follows: Section II describes the proposed technique for critical device identification. Section III presents the results from applying the approach to an operational amplifier case study. Concluding remarks are given in Section IV.

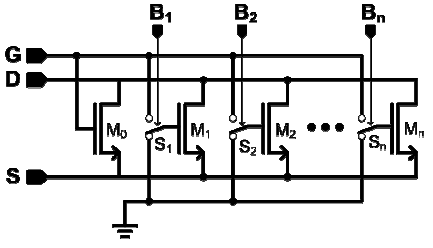


Figure 1: Structure of the configurable analogue transistor.

This work was funded by EPSRC Grant number EP/H014608/1
978-3-9810801-8-6/DATE12/©2012 EDAA

II. PROPOSED METHOD

A. Overview

The proposed process of CDI and its application to a circuit is illustrated in Figure 2 2 and comprises of five steps as follows.

Step 1: A conventional sensitivity analysis is performed in simulation. The dependence of circuit performance to changes in individual devices is recorded in a sensitivity table.

Step 2: The sensitivity information is used to perform CDI, resulting in a list of transistors that are most suitable.

Step 3: A Monte Carlo (MC) simulation of the circuit is used to adjust the critical devices in an ideal manner to minimise performance variability.

Step 4: The critical devices are replaced by CATs with a finite number of calibration transistors (optimal sizing from step 3).

Step 5: In the last step, a MC simulation is performed and the required adjustment for each critical device is calculated.

To illustrate the proposed process, Figure 3 shows the circuit of an operational amplifier. The circuit consists of a differential input stage, MP4 and MP5, gain stage, MN8, and an output buffer, MN6. MN9 and the capacitor form the internal compensation network and MP0-MP1, MN0-MN3 and the resistor form the bias circuit.

The performance characteristics that are considered are DC voltage gain, open-loop bandwidth and common-mode rejection ratio. The operational amplifier is designed in a standard 0.35µm CMOS process at 3.3V, with active and passive component values as listed in Table I.

B. Sensitivity Analysis and Critical Device Identification

The sensitivity of each performance to a change in each transistor's width is normalised as a relative change to its nominal value, as defined in Equation 1:

$$s_{A,n} = \frac{A|_{w_n=nom+5\%} - A|_{w_n=nom-5\%}}{A|_{w_n=nom}} \quad (1)$$

where A is a circuit performance (e.g. gain) evaluated with the

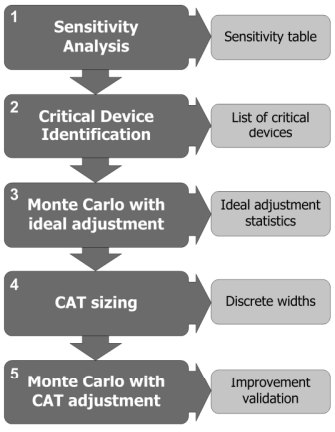


Figure 3: Process flow of critical device identification and CAT replacement

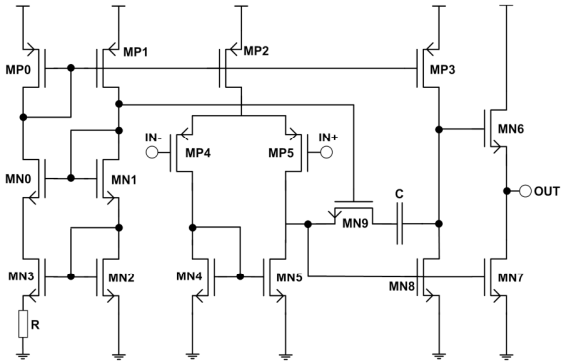


Figure 2: Circuit diagram of the operational amplifier

width of transistor n, w_n, at different values, resulting in the sensitivity of performance A on transistor n, $s_{A,n}$. The sensitivities obtained for the operational amplifier when applying the above sensitivity analysis for all performance-transistor combinations are shown on the left of Table II. The assumption in Table II is that there is a linear relationship between the performance and transistor width within the range of interest and that the circuit is a linear system within this range and therefore superposition applies. In practice these assumptions have been found to be valid and their implications are discussed later in this paper. In the proposed method for CDI, the two goals of high sensitivity and high independence are combined. First, a normalised measure for independence is derived by calculating the ratio of a particular sensitivity to all the sensitivities associated with that particular transistor, as per Equation 2:

$$w_{A,n} = \frac{|s_{A,n}|}{\sum |s_n|} \quad (2)$$

Where A represents a certain performance, n a certain transistor and $w_{A,n}$ the relative amount of transistor n's total impact on performance. The weighting table that was obtained for the operational amplified case study is shown in the centre section of Table II. Secondly, element-wise multiplication of the original sensitivity table and the weighting table is performed as per Equation 3:

$$s'_{A,n} = s_{A,n} \cdot w_{A,n} \quad (3)$$

which results in a sensitivity table, $s'_{A,n}$, that considers both absolute sensitivity and independence. The weighted sensitivity table for the operational amplifier is shown on the right of

TABLE I. DEVICE SIZES AND COMPONENT VALUES

Device	Dimensions	Device	Dimensions
MP0	8.75µm / 0.35µm	MN2	35µm / 0.35µm
MP1	8.75µm / 0.35µm	MN3	35µm / 0.35µm
MP2	105µm / 0.35µm	MN4	52.5µm / 0.35µm
MP3	105µm / 0.35µm	MN5	52.5µm / 0.35µm
MP4	105µm / 0.35µm	MN6	175µm / 0.35µm
MP5	105µm / 0.35µm	MN7	175µm / 0.35µm
MN0	8.75µm / 0.35µm	MN8	98µm / 0.35µm
MN1	8.75µm / 0.35µm	MN9	50µm / 0.35µm
Component	**Value**	**Component**	**Value**
R	823kΩ	C	623fF

Table II. For the purpose of this work, one critical device is chosen for each performance by selecting the transistors with greatest weighted sensitivity. In this case, the critical devices are therefore MP0 for bandwidth, MP3 for gain and MN7 for CMRR. Although MP0 significantly affects both gain and bandwidth it is still chosen as a critical device for bandwidth due to its very high absolute sensitivity to this performance.

C. Calculation of Ideal Transistor Adjustment

From the sensitivity information obtained in the previous section, the resulting change in circuit performance from adjusting the widths of the critical devices can be described by a system of linear equations:

$$\begin{aligned} \Delta A &= s_{A1}\Delta w_1 + s_{A2}\Delta w_2 + \cdots + s_{AN}\Delta w_N \\ \Delta B &= s_{B1}\Delta w_1 + s_{B2}\Delta w_2 + \cdots + s_{BN}\Delta w_N \\ \Delta C &= s_{C1}\Delta w_1 + s_{C2}\Delta w_2 + \cdots + s_{CN}\Delta w_N \end{aligned} \quad (4)$$

where ΔA is the change in performance A (e.g. bandwidth) from adjusting the width of transistor 1 (e.g. MP0) by Δw_1, defined by the sensitivity of parameter A to a change in transistor 1, s_{A1}. Note that this sensitivity is not the same numerical value as found in Table II, which is a dimensionless value normalised to the nominal performance and to a ±5% change in transistor width. In order to evaluate the above equations, the sensitivity has to be de-normalised in both dimensions, resulting in a gradient with units, e.g. Hz/µm.

Before the CAT technique can be applied to a circuit, the necessary ideal adjustment is computed first from a MC simulation of device parameter variation. The deviation of performances from their nominal values is used to solve Equation 4 for the necessary adjustment in transistor widths. The simulation is then repeated with the MC variables unaltered but the critical devices adjusted.

D. Application of Optimally Sized CATs

While the adjustment technique of Section II.C shows a significant improvement in performance, it is unrealistic because it assumes infinite granularity in the adjustment devices. In reality, a CAT consists of a finite number of calibration transistors. The optimal sizing algorithm for CAT [15] operates on statistical information of transistor width.

TABLE III - OPTIMAL SIZING OF THE CONFIGURABLE ANALOGUE TRANSISTORS

		MP3	**MP0**	**MN7**
Nominal width		105.0µm	8.75µm	175.0µm
CAT width step		11.64 µm	0.743 µm	10.20 µm
CAT device	***Main***	64.26 µm	6.150 µm	139.3 µm
	1st	11.64 µm	0.743 µm	10.20 µm
	2nd	23.28 µm	1.486 µm	20.40 µm
	3rd	46.56 µm	2.972 µm	40.80 µm

Each CAT is optimally sized by considering the distribution of required width adjustments shown in Figure 4. This results in an optimised final performance distribution because the distribution of width adjustment is directly related to circuit performance. Using the optimal sizing algorithm for CAT and assuming three calibration transistors in each critical device, the transistor sizes in Table IV have been calculated. To allow adjustment in both positive and negative direction, the nominal width of the main transistor M_0 is reduced and the three calibration transistors are sized to give eight evenly spaced selectable values centred on the original nominal width. After obtaining these transistor sizes, a further MC simulation is performed. Instead of adjusting devices with infinite granularity, the CAT configuration is used to constrain the adjustment to the eight selectable widths. This is equivalent to tuning the CATs on a chip after fabrication. The results of this therefore represent the performance obtained following post manufacture adjustment with optimally sized CATs.

III. RESULTS

In the case-study circuit, which is shown in Figure 3, MP0, MP3 and MN7 have been replaced with CATs, each with three calibration transistors sized according to Table IV. Figure 5 shows the histograms of the performances before and after application of the CATs. Clearly, the spread in all three performances is reduced significantly, resulting in a lower standard deviation and greater yield. Table V compares the standard deviations of the performances before and after application of the three CAT devices. The standard deviations are improved by 76.2%, 80.2% and 79.9% for gain, bandwidth and CMRR, respectively.

Table III-PERFORMANCE SENSITIVITY TO VARIATION IN TRANSISTORS

Device	**Sensitivity (s)**			**Weighting (w)**			**Weighted Sensitivity (s')**		
	Bandwidth	***Gain***	***CMRR***	***Bandwidth***	***Gain***	***CMRR***	***Bandwidth***	***Gain***	***CMRR***
MP5	0.0280	0.0159	0.0001	0.6361	0.3617	0.0022	0.0178	0.0058	0.0000
MP4	0.0057	0.0045	-0.0006	0.5297	0.4180	0.0523	0.0030	0.0019	0.0000
MP3	-0.0014	**-0.0466**	0.0000	0.0288	**0.9712**	0.0000	0.0000	**-0.0452**	0.0000
MP2	0.0634	-0.0128	0.0005	0.8263	0.1674	0.0063	0.0524	-0.0021	0.0000
MP1	0.0671	-0.0506	0.0003	0.5685	0.4285	0.0029	0.0381	-0.0217	0.0000
MP0	**-0.1287**	0.0996	-0.0007	**0.5620**	0.4351	0.0030	**-0.0723**	0.0433	0.0000
MN9	0.0001	0.0000	0.0000	1.0000	0.0000	0.0000	0.0001	0.0000	0.0000
MN8	0.0021	0.0468	0.0000	0.0422	0.9578	0.0000	-0.0002	0.0448	0.0000
MN7	-0.0048	-0.0019	**0.0970**	0.0467	0.0185	**0.9348**	-0.0007	0.0011	**0.0907**
MN6	-0.0015	0.0019	0.0000	0.4422	0.5577	0.0001	0.0292	-0.0019	0.0000
MN5	0.0370	-0.0095	-0.0004	0.7891	0.2031	0.0078	-0.0345	0.0001	0.0000
MN4	-0.0366	0.0019	0.0003	0.9429	0.0482	0.0089	0.0242	-0.0143	0.0000
MN3	0.0430	-0.0331	0.0002	0.5634	0.4336	0.0030	-0.0350	0.0202	0.0000
MN2	-0.0618	0.0470	-0.0003	0.5663	0.4307	0.0029	0.0005	0.0000	0.0000
MN1	-0.0006	-0.0001	0.0000	0.8699	0.1291	0.0010	0.0001	0.0000	0.0000
MN0	0.0001	-0.0001	0.0000	0.5648	0.4314	0.0038	0.0001	-0.0001	0.0000

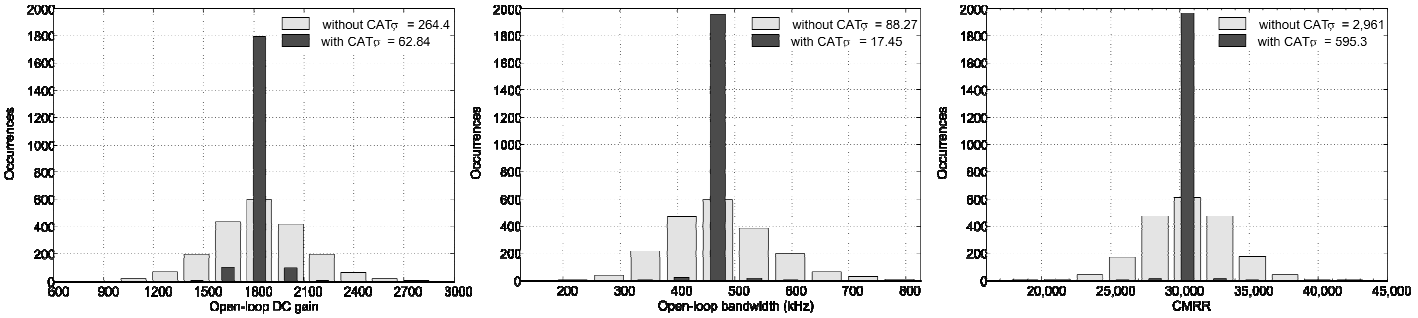


Figure 1. Histograms of circuit perfromances.

Table IV also shows the best theoretical improvement by applying the optimal CAT sizing algorithm in [15]. The theoretical maximum is almost impossible to achieve in practice because the relationship between CAT width and the performance it adjusts is unlikely to be perfectly linear. The results using CAT are remarkably close to the theoretical maximum improvement, indicating that the assumption of linear sensitivity is adequate for this particular combination of circuit, parameters and spread.

IV. CONCLUSIONS

The CAT technique provides a calibration platform that is independent of the target circuit and the mechanisms of performance variation. In this paper a novel automated method is proposed that determines which transistors in a circuit should be replaced with CAT devices in order to achieve maximum post fabrication yield improvement. It is demonstrated that both the performance sensitivity and the adjustment independence should be taken into account, giving a combined weighted sensitivity. In the case study, three critical devices were identified and replaced which led to an average of 78.8% improvement in the variability of key circuit performances. These results demonstrate that the proposed CDI methodology and performance driven CAT sizing can form a successful approach to improve analogue circuit yield.

TABLE IV- CIRCUIT PERFORMANCE AFTER APPLYING CAT

Condition	**Performance**			
		gain	***BW***	***CMRR***
Nominal	Mean	1.83×10^3	471×10^3	30.5×10^3
Monte Carlo	Mean	1.83×10^3	474×10^3	30.5×10^3
	Standard deviation	264	88.3×10^3	2.96×10^3
Adjustment with CAT	Mean	1.83×10^3	471×10^3	30.5×10^3
	Standard deviation	62.8	17.5×10^3	595
	Standard deviation improvement	76.2%	80.2%	79.9%
	Maximum standard deviation improvement	80.7%	80.8%	80.3%

V. REFERENCES

[1] S Saxena, C Hess, H Karbasi, A Rossoni, S Tonello, P McNamara, S Lucherini, S Minehane, C Dolainsky, and M Quarantelli. "Variation in Transistor Performance and Leakage in Nanometer-Scale Technologies." IEEE Transactions on Electron Devices, 55(1):131–144, 2008.

[2] A.-J. Annema, B Nauta, R van Langevelde, and H Tuinhout. "Analog circuits in ultra-deep-submicron CMOS." IEEE Journal of Solid-State Circuits, 40(1):132–143, 2005.

[3] G Gielen, P De Wit, E Maricau, J Loeckx, J Martin-Martinez, B Kaczer, G Groeseneken, R Rodriguez, and M Nafria. "Emerging Yield and Reliability Challenges in Nanometer CMOS Technologies." In Design, Automation and Test in Europe, 2008. DATE '08, pp 1322–1327, 2008.

[4] J.J. Price. "A passive laser-trimming technique to improve the linearity of a 10-bit D/A converter." IEEE Journal of Solid-State Circuits, 11(6):789 – 794, Dec 1976.

[5] K. Kato, T. Ono, and Y. Amemiya. "A monolithic 14 bit D/A converter fabricated with a new trimming technique (DOT)". IEEE Journal of Solid-State Circuits, 19(5):802 – 807, Oct 1984.

[6] J. Sweeney and R. Geiger. "Very high precision analog trimming using floating gate mosfets." 1989., European Conference on Circuit Theory and Design, pages 652–655, Sep 1989.

[7] S.A. Jackson, J.C. Killens, and B.J. Blalock. "A programmable current mirror for analog trimming using single poly floating-gate devices in standard CMOS technology." Circuits and Systems II: IEEE Trans. on Analog and Digital Signal Processing, 48(1):100 –102, 2001.

[8] Z Wang, H S Savci, J D Griggs, N S Dogan, and E Arvas. "Coping with process variations in ultra-low power CMOS analog integrated circuits." Proceedings. IEEE SoutheastCon, 2007, pages 54–57, 2007

[9] J Maunu, M Pankaala, J Marku, J Poikonen, M Laiho, and A Paasio. "Current source calibration by combination selection of minimum sized devices." 2006 IEEE International Symposium on Circuits and Systems, 2006. ISCAS 2006. Proceedings., pp. –552, 2006.

[10] Xin Li; Taylor, B.; YuTsun Chien; Pileggi, L.T., "Adaptive post-silicon tuning for analog circuits: concept, analysis and optimization," IEEE/ACM International Conference on Computer-Aided Design, 2007, ICCAD 2007, pp.450-457, 4-8 Nov. 2007.

[11] Wei Jiang and V D Agrawal. "Designing variation-tolerance in mixed-signal components of a system-on-chip" IEEE International Symposium on Circuits and Systems, 2009, pp 2313–2316, 2009.

[12] Tao Chen and G G E Gielen. "A 14-bit 200-MHz Current-Steering DAC With Switching-Sequence Post-Adjustment Calibration." IEEE Journal of Solid-State Circuits, 42(11):2386–2394, 2007.

[13] P R Wilson and R Wilcock. "Yield improvement using configurable analogue transistors" Electronics Letters, 44(19):1132–1134, 2008

[14] Wilson, P.R.; Wilcock, R.; , "Operational calibration of mixed-signal integrated circuits in hostile environments," 2009 IEEE Aerospace conference, pp.1-7, 7-14 March 2009

[15] P Wilson and R Wilcock. "Optimal sizing of configurable devices to reduce variability in integrated circuits" In Design, Automation Test in Europe Conference Exhibition, 2009. DATE '09., pages 1385–1390

The following pages contain the paper "Reliability Improvement and Online Calibration of ICs Using Configurable Analogue Transistors ", which was published at the 2012 IEEE Aerospace Conference. It is based on the work presented in Section 3.4.

Reliability Improvement and Online Calibration of ICs Using Configurable Analogue Transistors

Robert Rudolf, Reuben Wilcock and Peter R Wilson
Electronics and Electrical Engineering Group
School of Electronics and Computer Science
University of Southampton, UK
+44 23 8059 4162
{rr1v07, rw3, prw}@ecs.soton.ac.uk

***Abstract*—Reliability of electronic circuits over an extended temperature range is a critical consideration in demanding applications such as aerospace and the military. Achieving this reliability on modern deep submicron process nodes is a significant challenge especially for analogue circuits due to the high level of device variability. A novel approach is proposed in this paper that employs online adjustment of configurable analogue transistors (CATs) to address this challenge, significantly improving the consistency of circuit performance over temperature. The proposed method involves optimally sizing configurable devices for temperature and process variation and then employing a calibration lookup table during normal operation to compensate for temperature shifts. In the presented case study of an instrumentation amplifier, the CAT approach is shown to successfully mitigate temperature induced performance loss, demonstrating significant calibration potential and reliability improvement. These advantages are enjoyed at minimal cost in terms of area and complexity overhead, and the process of implementing the circuit changes is highly automated. The promising results detailed in this work demonstrate that the CAT technique has useful applications in the area of reliability improvement for demanding environments.**

TABLE OF CONTENTS

1. INTRODUCTION

The application of modern integrated circuits in hostile environments raises several design challenges. Extreme operating temperature and elevated levels of radiation are typical characteristics of hostile environments which significantly affect circuit performance or may lead to premature ageing and potential failure [1-2]. In most cases, such extreme environment applications also place exceptional demands on the reliability of electronic circuits. Space missions and defense are typical examples, where circuit failure may cost millions or in the worst case, human lives. The majority of research in current high-reliability electronics for extreme environments focuses on two areas. The first area is concerned with devices and processes that enable electronics to operate at extreme temperatures or under high levels of radiation. Examples for this research include silicon carbide (SiC) semiconductors [3], solid-state vacuum devices [4] and packaging and interconnect [5]. The second area is concerned with fault-tolerant circuits, which can resume normal operation despite faults by employing dynamic reconfiguration. Research in fault-tolerant circuits has been carried out for both digital [6] and analogue [7] circuits.

An area of research that has seen extensive exploration in the context of manufacturing yield improvements, but comparatively little in the context of electronics for hostile environments is calibration for device variation and temporal effects. A wide range of approaches on all levels of design exist to improve manufacturing yield or reliability by calibrating circuits after fabrication or during operation, e.g. [8, 9]. Although most of these techniques are optimized for, but not limited to, calibration for yield improvement, they can also be employed for online calibration in extreme environments. However, thus far no successful attempts have been demonstrated in applying existing post-fabrication calibration techniques to enable circuits for extreme environments.

Reliability is a measure of how well a system can perform its functions to specification over a certain period and certain conditions. Traditionally, reliability is viewed in the context of hard faults, meaning that the system fails due to individual device faults or irreversible deterioration of device performances. In this case, the exact system performance is less relevant as long as it is within specification because the decision of whether or not a system has failed is a binary yes/no outcome. However, reliability can also be considered from a parametric point of view, referred to as parametric reliability, and considers parametric faults instead of hard faults. A parametric fault is a temporary condition where system performance is moved out of specifications, but returns to its normal value once the cause has been removed. A classic example for a parametric fault mechanism is temperature drift.

As discussed previously, reliability can be optimized by choosing more robust devices and fault-tolerant circuit and system architectures. On the other hand, measures to improve parametric reliability are ideally taken at the circuit level. Examples include variation-tolerant circuit design and online calibration, as will be described in this work.

This work was funded by EPSRC Grant number EP/H014608/1
978-1-4577-0557-1/12/$26.00 ©2012 IEEE

The Configurable Analogue Transistor (CAT)

In this work, the Configurable Analogue Transistor (CAT) [10] is proposed as a circuit-level calibration technique that can significantly improve reliability and performance over the operating temperature range of circuits in hostile environments. The principle of CAT is to replace certain devices (critical devices) with digitally adjustable width devices, thus allowing circuit performance to be controlled. The original application of the CAT is to compensate the effects of process parameter or mismatch variation, which can facilitate the application of devices or processes with high inherent variability, such as analog devices on small process nodes or novel processes.

Since the CAT technique relies on the availability of standard CMOS devices and does not extend the device operating temperature range, the absolute maximum and minimum operating temperature of a circuit are still limited by the underlying fabrication process. However, the CAT technique improves the variation of circuit performance within this range and thus extends the useable operating range of a circuit, that is, the range of temperatures over which it operates to specification. The CAT technique has previously been proposed as a means of improving reliability in hostile environments [11]. However, the discussion of this matter did not consider a specific application and environment. In this paper, temperature is suggested as a possible target environmental parameter for the application of CAT. The application of CAT to improve parametric reliability over temperature is described and the concept illustrated by means of a demonstrator circuit.

The structure of the CAT is shown in Figure 1. It consists of a main device M_0 and n calibration devices M_1 to M_n, which can be selected through n digital control lines, B_1 to B_n. Each of these control lines either grounds the gate of a calibration device or connects it to the gate of the main device, resulting in a total of 2n discrete widths. Although similar circuit structures have previously been used in digitally adjustable analogue circuits, the CAT methodology includes a unique optimal sizing process which ensures the highest possible level of calibration [12].

The CAT technique does not only consist of the configurable CMOS device, but also of a set of design tools. These tools are an integral and unique part of the CAT technique. Figure 2 shows the typical IC design flow where CAT is employed. As can be seen, CATs are primarily applied between schematic capture and layout, with a single post-fabrication calibration step. The individual tools of the CAT design flow are briefly described below.

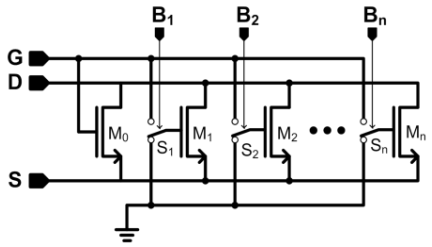


Figure 1. Structure of the configurable analogue transistor.

The task of the first tool is to determine which devices in a circuit should be replaced by CATs, in a process called Critical Device Identification (CDI). In order to perform CDI, the circuit must be embedded in a testbench and the circuit performances such as gain, bandwidth, etc. must be described by simulator expressions. By means of sensitivity analysis, the CDI tool determines which transistors are most suited for adjusting these particular performances. A difference to conventional calibration techniques is that the addition of calibration elements (CATs) is performed after schematic capture. This means that the designer does not need to concern themselves with finding a good calibration solution during the design of the circuit. Automating this process is not only more efficient in terms of design time, but it also allows optimal selection of critical devices according to the given performance specifications.

The second tool in the CAT design process determines the optimal sizes of the calibration transistors (M_1 to M_n) of the CATs. This sizing is based on stochastic information about the performances when the circuit is subject to device parameter variation. An optimal sizing algorithm [12] is then employed to size the CATs such that the overall performance variability of the circuit is minimized. Once the CATs have been sized, the design can proceed to the layout stage, where the CATs are treated like an array of regular CMOS transistors.

Once the circuit has been fabricated, the optimal configuration of CATs is determined for each individual chip. The main focus of this work, is the online reconfiguration of the CATs after fabrication. The description of the CAT design process in this section was with focus on device variability. It will be shown in the next section how this design process and the application of the CAT can also incorporate calibration for temperature variation.

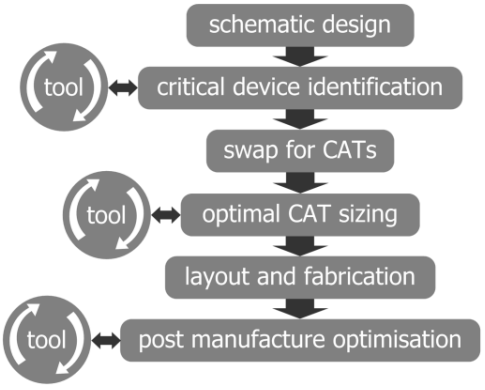


Figure 2. Design flow for CAT

The rest of this paper is structured as follows. Section 2 describes how the CAT can be used for online calibration over temperature. Section 3 applies this online calibration technique to a demonstrator circuit and discusses the obtained results. Section 4 concludes this paper and summarizes the results.

2. APPLICATION

Online Calibration Mechanism

The primary design goal of a CAT is to allow post-fabrication calibration to compensate for errors introduced by process variation. After the CAT design flow described in Section 1, each chip is individually tested and the optimal CAT settings to achieve best performance are determined. This optimal configuration is typically stored in nonvolatile on-chip memory so that it can be restored whenever necessary, e.g. after the chip is powered up.

Since both process and mismatch variation are largely time invariant, a static CAT configuration is sufficient to counteract any errors introduced by these mechanisms to achieve optimal performance. However, in this configuration the circuit is still subject to environmental influences, such as temperature, radiation and ageing. Performance degradation introduced by these means cannot be compensated with a static CAT configuration, which calls for an online calibration approach.

Online calibration of a circuit equipped with CAT is conceptually very simple, and requires the CAT configuration to be altered during run-time according to certain rules. In principle, this involves measuring the current system performance and, if necessary, switching to a different CAT configuration that will improve performance. However, there are at least two complications in this generic case. First, to determine the current performance of the circuit, it may be necessary to suspend normal operation and put the circuit in a test mode. Second, determining the optimal CAT configuration can be an iterative process, during which the circuit is not likely to operate at optimal performance. The result from these issues is that the circuit will not be able to perform its normal operation continuously and that it may operate outside specifications for a certain amount of time. In this work, it will be shown that in the case of temperature, online CAT reconfiguration can be based on a simple lookup table without the need to measure system performance or perform iterative optimization.

Online Temperature Compensation

Online calibration of CATs with respect to temperature is a special case that lends itself well to practical implementation. The dependence of circuit performance on temperature is well described through SPICE models and the temperature of the chip can be easily measured continuously, which allows the system to conduct the appropriate reconfiguration before the performance has dropped below a threshold. Additionally, the temperature behavior of the circuit can be accurately modeled before fabrication, which reduces the reconfiguration process to a simple lookup table. This type of online reconfiguration can be carried out without any interruptions in the operation of the circuit, because the current performance does not need to be measured and the optimal configuration is predetermined. However, signals processed in the system may still be subject to short glitches at the moment when the CAT configuration is changed.

Figure 3 illustrates the required system architecture for online CAT reconfiguration. The temperature of the chip is continuously monitored, and the corresponding optimal CAT configurations obtained from a lookup table. There are several points to note about this concept. First, in most practical applications, temperature does not need to be measured continuously. Instead, it may be sufficient to sample its value at given intervals or only under certain conditions. The latter is especially interesting for applications onboard spacecraft, where the system temperature may only change, for example, after certain navigational maneuvers. Similarly, the temperature of a planetary probe is likely to be known either from the current time of day or the probe's main instruments, which completely removes the need for on-chip temperature measurement. Furthermore, discontinuous sampling of temperature also reduces power consumption, since the temperature sensor and the associated reconfiguration hardware operate only in short bursts. Secondly, the task of digitizing temperature readings and looking up the corresponding configuration words in memory bear very little computational load. It is therefore practical to handle this task in an already existing digital processing system, rather than a dedicated computer for CAT reconfiguration. Again, this is especially beneficial for applications in which energy conservation is a primary requirement.

In summary, the hardware overhead for incorporating online CAT reconfiguration is potentially very low. Apart from the CATs themselves, the only other required on-chip component is a temperature sensor, which may be as simple as an appropriately biased PN junction. All remaining components, such as the ADC, computation, lookup table and configuration memory may be incorporated into an existing signal processing system at little additional cost.

Design of CAT for online temperature calibration

The CAT design process when considering temperature variation is in principle no different to the process

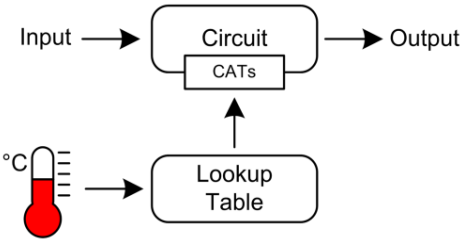


Figure 3. System structure for online temperature calibration using CATs

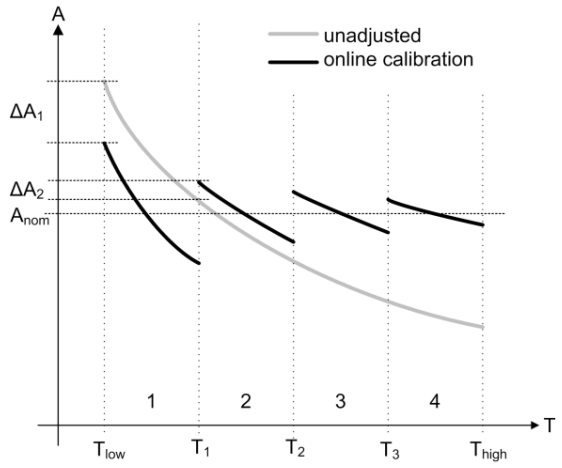


Figure 4. Principle of temperature compensation using CATs.

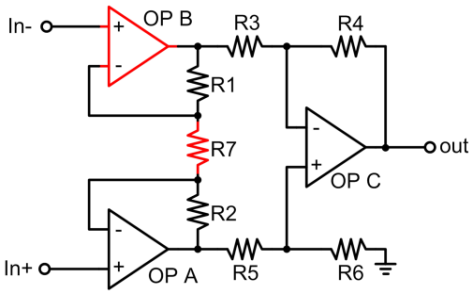


Figure 5. Circuit diagram of the instrumentation amplifier.

introduced in Section 1. However, instead of performing a Monte Carlo simulation across the process parameter space to gain stochastic information about the circuit's performance, a simple temperature sweep across the specified range is sufficient. Figure 4 shows a typical temperature dependence of a particular circuit performance, A, exhibiting a negative temperature coefficient. To find the optimal sizes of the CAT devices, the established optimal sizing algorithm can be used with the temperature dependence as an input distribution. The resulting CATs will be sized such that the mean deviation from the nominal value over the entire temperature range is minimized. In addition to optimal CAT sizing, the design process also outputs a configuration lookup table, mapping temperature to the CAT configuration.

For the purposes of illustration, a possible outcome of calibrating the example performance with a 2-bit CAT device is also shown in Figure 4. The CAT configuration that is active in a certain temperature range is indicated by numbers along the temperature axis. For very low temperatures, configuration 1 is chosen, which reduces the numerical value of the performance by ΔA_1. This reduction in value brings the mean of the performance between T_{low} and T_1 closer to the nominal performance, A_{nom}. If the temperature rises above T_1, configuration 2 is chosen. This reduces the performance only by ΔA_2, thereby bringing the performance closer to the nominal value, and so on. This example should also reinforce the point that neither the temperatures at which the configurations change nor the sizing of the CAT devices, corresponding to the change in performance, are arbitrary, but must be optimized during the design stage.

While this approach to temperature compensation is valid for a single chip at nominal device parameters, it does not consider the various parameter variation processes that occur in real circuits. A real circuit design, which includes optimally sized CAT devices, is replicated several times on a wafer to yield a large number of chips. While ideally all chips from a certain design have identical behavior, in reality the performance of any two chips and indeed identical devices on the same chip is not the same. These processes are referred to as process and mismatch variation, respectively and are modeled though stochastic processes in the fundamental device parameters.

The consequences of these variation mechanisms on the application of CATs to compensate temperature variation are two-fold. Firstly, because the designed CAT must provide good results on all produced chips of a given circuit, optimal sizing of the CAT must now consider both temperature and parameter variation. This brings the CAT from simple temperature sweeps back to its original stochastic domain, where the temperature can be considered as an additional random variable. Secondly, because the CAT must now compensate parameter and temperature variations, the achievable level of calibration will be lower than in the case where only temperature was considered. Nevertheless, the expected improvement in performance variation is still well defined through the stochastic processes.

A crucial difference between the temperature-only and variation-aware CAT application lies in the post-fabrication stage. In the case where only temperature is considered, it is sufficient to generate a single configuration lookup table from the simulations that is valid for all chips of a particular circuit. When considering additional parameter variations, not only must the initial CAT configuration be determined on a chip-by-chip basis, but also an individual lookup table generated for each chip. This is necessary because both the initial CAT configuration and the temperature behavior are likely different between chips. However, generation of the lookup table is computationally very inexpensive and follows directly from the initial CAT configuration. Therefore, this does not require any additional post-fabrication test equipment and does not significantly prolong post-fabrication calibration time.

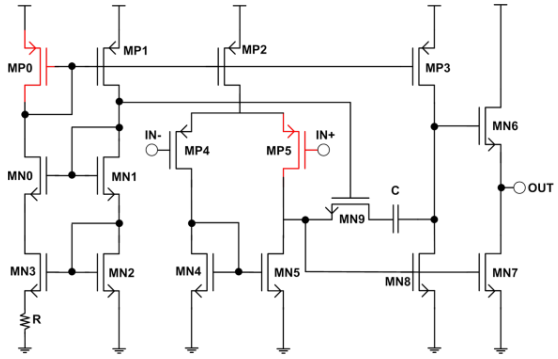


Figure 6. Circuit diagram of the operational amplifier

Demonstrator Circuit

In order to illustrate the merits of the CAT technique in the context of this paper, it is applied to a standard instrumentation amplifier comprised of three identical operational amplifiers. The amplifier block schematic is shown in Figure 5, while the operational amplifier used in the circuit is shown in Figure 6. The resistors in the instrumentation amplifier are modeled as diffusion resistors and n-well resistors rather than ideal resistors for two reasons. Firstly, the availability of accurate foundry models means that any nonidealities of on-chip passive components are modeled correctly, resulting in a realistic description of overall circuit performance after fabrication. Secondly, being structurally equivalent to a MOS device without a gate, diffusion resistors can be replaced by CATs and therefore tuned. As will be shown later, this is of great importance in this particular circuit.

TABLE I. OPAMP DEVICE SIZES AND COMPONENT VALUES

Device	Dimensions	Device	Dimensions
MP0	**8.75µm / 0.35µm**	MN2	35µm / 0.35µm
MP1	8.75µm / 0.35µm	MN3	35µm / 0.35µm
MP2	105µm / 0.35µm	MN4	52.5µm / 0.35µm
MP3	105µm / 0.35µm	MN5	52.5µm / 0.35µm
MP4	105µm / 0.35µm	MN6	175µm / 0.35µm
MP5	**105µm / 0.35µm**	MN7	175µm / 0.35µm
MN0	8.75µm / 0.35µm	MN8	98µm / 0.35µm
MN1	8.75µm / 0.35µm	MN9	50µm / 0.35µm
Component	**Value**	**Component**	**Value**
R	823kΩ	C	623fF

TABLE II. INSTRUMENTATON AMPLIFIER COMPONENT VALUES

Component	Value	Component	Value
R1	99k	R5	100k
R2	99k	R6	100k
R3	100k	**R7**	**2k**
R4	100k		

TABLE III. NOMINAL CIRCUIT PERFORMANCES

Performance	Symbol	Nominal value
Gain	G	94.35
Bandwidth	BW	1.99×10^{6}
Offset voltage	VOS	5.55×10^{-5}

Table I and Table II show the transistor sizes for the operational amplifier, and resistor values for the instrumentation amplifier respectively. The nominal circuit performances are listed in Table III. The circuit is implemented in a standard 0.35µm CMOS process, where the foundry device SPICE models are valid within the temperature range from -40°C to +125°C. The critical devices in this circuit, found using the algorithm in [13], are R7, MP0B and MP5B for gain, bandwidth and offset voltage, respectively. This diffusion resistor and the MOSFETs of operational amplifier B are set in bold in Tables I and II and colored red in Figures 5 and 6.

It is worth noting that one of the resistors has been identified as a critical device. Because the operational amplifiers operate under negative feedback, gain is primarily determined by the passive components with very little effect from active components. Therefore, the only viable way to adjust gain is to equip certain passive components with a CAT structure. In the case of resistor R7, this task is very straightforward because diffusion resistors lend themselves readily to the CAT structure. Bandwidth, on the other hand, is determined by the passive components and the gain-bandwidth product of the operational amplifier and can therefore be adjusted through transistors. Finally, offset voltage is not dependent on the passive components at all and can be adjusted by varying the transistors in the differential input stages.

3. RESULTS

Temperature and parameter variation

In this section, the results from applying online CAT calibration to a circuit that is subject to temperature and process parameter variations are presented. As described in Section 1, the CAT transistors are sized according to stochastic information about temperature and process parameter behavior. During operation, the best CAT configuration is chosen from a lookup-table, which is customized for each chip.

To illustrate this concept, a Monte Carlo simulation of the instrumentation amplifier was performed. In addition to process and mismatch variation, temperature was added as a random Monte Carlo variable. This means that each Monte Carlo iteration represents a particular circuit at a particular temperature. Since the CAT technique is based on stochastic information about circuit performance, it is possible to derive all necessary information for CAT sizing from this, without the need for a full temperature sweep for every set of Monte Carlo parameters. Indeed, such information would be meaningless for CAT sizing because the CAT is sized for a particular circuit, of which multiple copies are produced on a single wafer. CAT sizes can therefore not be optimized for a single chip, but for a particular circuit, for which only stochastic information is relevant. Figure 7 shows the results of the Monte Carlo simulation, with the circuit performances plotted against temperature and marked with the symbol ×. Gain and bandwidth clearly show a

dependency on temperature, with the performances of individual circuits scattered in a band around the ideal. This scattering represents the magnitude of process parameter variation. In the case of offset voltage, the influence of mismatch variation outweighs the temperature dependence by far. This means that for a particular chip, CAT will be able to reduce the inherent offset voltage well, but online temperature calibration will not be able to improve temperature drift greatly. This can be visualized by the fact that each CAT transistor only has a finite number of possible configurations. When the effects of process parameter or mismatch variation are small compared to the effects of temperature, only a small set of the possible configurations are required for the initial post-fabrication tuning, leaving ample free configurations for online calibration. However, if the effects of process parameter variation outweigh the effects of temperature, the majority of configurations are required for the initial calibration, leaving few or no free configuration states for online temperature calibration.

Without calibration, the standard deviation of gain is 6.85, the standard deviation of bandwidth is 222kHz and the standard deviation of the offset voltage is 3.69mV. These results are listed in the upper section of Table IV.

For each Monte Carlo parameter set, the optimal ideal adjustment in the critical devices is determined by applying the method described in [13] until all circuit performances are within 1% of their nominal value. However, it is clear that this adjustment is unrealistic because it would require transistor widths to be adjustable with infinite granularity. In practice, the number of calibration transistors in a CAT is limited. The optimal sizing algorithm for CAT is then applied to these ideal adjustments to give optimal sizes for CATs with three adjustment transistors, resulting in eight discrete adjustment steps for each critical device. The resulting sizes of the CAT devices are listed in Table V. For each Monte Carlo parameter set, the configuration of each CAT is chosen to be closest to the ideal adjustment.

The performances after CATs have been introduced in the circuit are marked with + in Figure 7. Over the entire temperature range, the standard deviations of gain, bandwidth and offset voltage are reduced to 2.81, 65.1kHz and 1.39mV, respectively. These calibrated performances are listed in the lower section of Table IV. As can be seen, introduction of CATs again significantly reduces variation in performances over temperature.

It is worth noting that a small number of performances after CAT calibration are significantly further from the nominal values than the majority. These points correspond to parameter sets for which no improvement in performance could be achieved within the adjustment constraints. Such instances will also occur in a real set of chips, where there will be a small number that cannot be calibrated at all. Because such circuits are already identified at the post-fabrication adjustment stage, they can be discarded as necessary.

Although the results obtained thus far show a significant reduction in performance standard deviation over the entire temperature range, no statement about parametric reliability has yet been made. For each performance, a pass band can be defined around the mean within which that performance is considered to operate to specification. For both Monte Carlo and CAT calibrated performances, parametric reliability can then be defined as the probability of a certain circuit at a certain temperature being within these bands. Although this definition is equivalent to the definition of yield, there is a practical difference introduced by the

TABLE IV CIRCUIT PERFORMANCE AFTER APPLYING CAT

Condition		**Performance**		
		gain	***BW***	***vos***
Nominal	Mean	94.35	1.99×10^{6}	5.55×10^{-5}
Monte Carlo	Mean	92.49	2.03×10^{6}	7.05×10^{-5}
	Standard deviation	6.85	2.22×10^{5}	3.69×10^{-3}
Adjustment with CAT	Mean	94.26	2.00×10^{6}	6.70×10^{-5}
	Standard deviation	2.81	6.51×10^{4}	1.39×10^{-3}
	Standard deviation improvement	59.0%	70.7%	62.3%

TABLE V. OPTIMAL SIZING OF THE CONFIGURABLE ANALOGUE TRANSISTORS

		R7	**MP0B**	**MP5B**
Nominal size		11.45µm	8.75µm	105µm
CAT size step		0.52µm	0.70µm	2.91µm
CAT device	***Main***	9.64µm	6.33µm	94.81µm
	1st	0.52µm	0.70µm	2.91µm
	2nd	1.04µm	1.39µm	5.82µm
	3rd	2.07µm	2.77µm	11.65µm

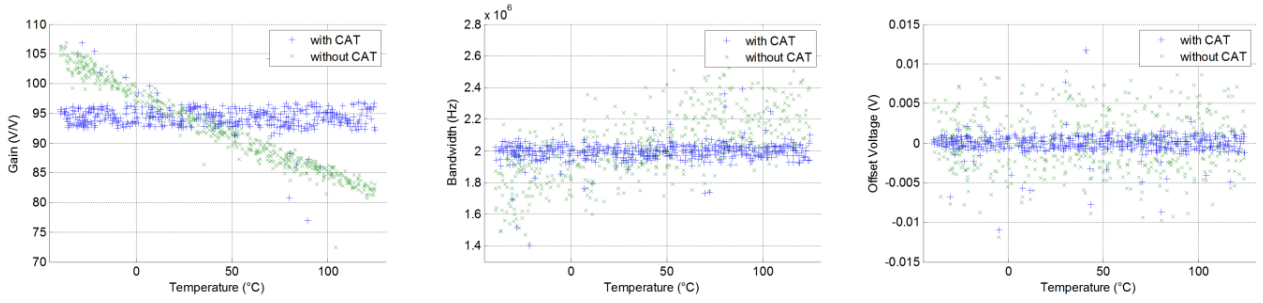


Figure 7. Temperature dependence of performances at with process and mismatch variation.

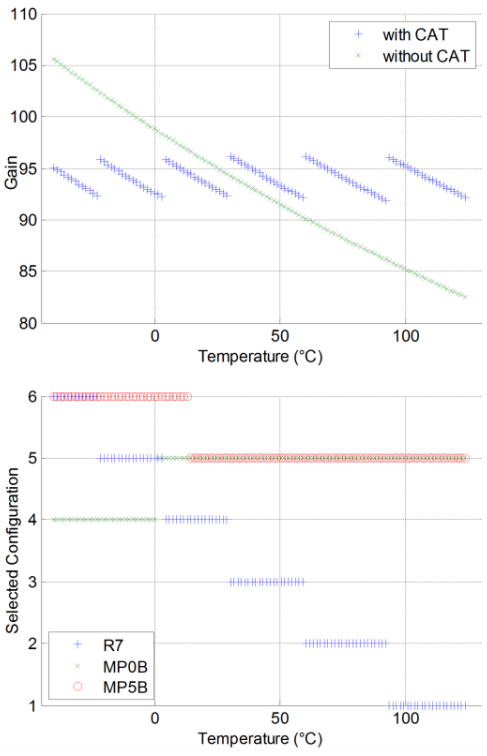


Figure 8. Temperature dependence of gain for one chip

inclusion of temperature.

Whilst yield is concerned with the probability of a circuit meeting specifications at static operating conditions, parametric reliability is concerned with the circuit meeting specifications over the entire range of operating conditions. For the purpose of illustration, the pass bands have been defined as ±5, ±100kHz and ±1mV, for gain, bandwidth and offset voltage, respectively. For the unadjusted Monte Carlo results, the system's parametric reliability is 4.0%. When CAT is applied to the system, parametric reliability increases to 80.8%. This is a significant increase in the parametric reliability of the system. Although this improvement in reliability seems exceptionally large when compared to the improvement in standard deviations, it is not surprising. Firstly, standard deviation is greatly affected by even a few outliers, while they do not contribute as greatly to a decrease in reliability. Secondly, reliability requires all three performances to be within the pass band, which is very improbable in the Monte Carlo case, leading to a low uncompensated reliability.

Application for temperature compensation

The results from the statistical analysis give an overview of the expected performance improvement as a statistical result over a large number of chips over the process parameter and temperature space. To illustrate the improvements achieved by CATs in single circuits, a single Monte Carlo parameter set is chosen and a full temperature sweep performed on it. This corresponds to the performance of an individual chip form a production run.

Similar to the previous section, the top section of Figure 8 shows the gain of the demonstrator circuit plotted against temperature. On the bottom of Figure 8, the currently chosen CAT configurations (0-7) for all three critical devices are plotted against temperature. This corresponds to the lookup table required for each individual chip and has been obtained by again finding the optimal ideal adjustment for each point and then selecting the closest configuration. Since there are only a finite number of CAT configurations, the temperature drift in gain can clearly be seen in Figure 8. However, as in the conceptual illustration of Figure 4, the absolute error in gain over temperature is improved.

4. CONCLUSIONS

In this paper, the application of the Configurable Analogue Transistor (CAT) was extended to online calibration of circuit performance for temperature variation. In an example circuit, the change in performance over the specified temperature range could be improved by between 59.0% and 70.7%, resulting in a significant improvement and potentially increasing the operational temperature range and performance of precision circuits. Furthermore, since the CAT is used to compensate for process parameter or mismatch variation at the same time, the inherent precision of circuit performance is also improved significantly. This mitigation of process variability effects enables high-temperature processes that suffer from great inherent device variability, such as SiC, to be used for precision analogue circuits. Enabling such technologies to be used in new applications will have a significant impact on the possible performance of systems in hostile environments, such as aerospace or defense.

REFERENCES

[1] Anderson, W.T., "Semiconductor device reliability in extreme high temperature space environments," IEEE Aerospace Conference, 2001

[2] Patterson, R.L.; Dickman, J.E.; Hammoud, A.; Gerber, S., "Electronic components and circuits for extreme temperature environments," IEEE Aerospace Conference, 2003.

[3] Cooper, J.A., Jr., "Silicon carbide electronic devices and integrated circuits for extreme environments," IEEE Aerospace Conference, 2004

[4] Sadwick, L.P.; Hwu, R.J.; Chern, J.H., "Extreme temperature electronics," IEEE Aerospace Conference, 2004.

[5] Grummel, B.; McClure, R.; Lei Zhou; Gordon, A.P.; Chow, L.; Shen, Z.J., "Design consideration of high temperature SiC power modules," 34th Annual Conference of IEEE Industrial Electronics, 2008. IECON 2008.

[6] Terry, S.C.; Blalock, B.J.; Jackson, J.R.; Suheng Chen; Durisety, C.S.A.; Mojarradi, M.M.; Kolawa, E.A., "Development of robust analog and mixed-signal electronics for extreme environment applications," IEEE Aerospace

Conference, 2004.
[7] Stoica, A.; Keymeulen, D.; Zebulum, R., "Evolvable hardware solutions for extreme temperature electronics," The Third NASA/DoD Workshop on Evolvable Hardware, 2001.
[8] Maunu, J.; Pankaala, M.; Marku, J.; Poikonen, J.; Laiho, M.; Paasio, A., "Current source calibration by combination selection of minimum sized devices," 2006 IEEE International Symposium on Circuits and Systems, 2006. ISCAS 2006
[9] Hyde, J.; Humes, T.; Diorio, C.; Thomas, M.; Figueroa, M.; , "A 300-MS/s 14-bit digital-to-analog converter in logic CMOS," IEEE Journal of Solid-State Circuits, vol.38, no.5, pp. 734- 740, May 2003
[10] P R Wilson and R Wilcock. "Yield improvement using configurable analogue transistors" Electronics Letters, 44(19):1132–1134, 2008
[11] Wilson, P.R.; Wilcock, R., "Operational calibration of mixed-signal integrated circuits in hostile environments," IEEE Aerospace conference, 2009.
[12] P Wilson and R Wilcock. "Optimal sizing of configurable devices to reduce variability in integrated circuits" In Design, Automation Test in Europe Conference Exhibition, 2009. DATE '09., pages 1385–1390
[13] R Rudolf, P Taatizadeh, R Wilcock and P Wilson, "Critical Device Identification for Configurable Analogue Transistors", to be published at DATE 2012

BIOGRAPHIES

Robert Rudolf (StM '07, GSM '10) graduated with a BEng in Electronic Engineering from the University of Southampton in 2010. He is currently pursuing a PhD in the Electronics and Electrical Engineering group in the School of Electronics and Computer Science at the University of Southampton, where he works on configurable analogue circuits. His research interests include analogue and mixed-signal circuit design.

Reuben Wilcock (M'02) is a Senior Research Fellow in the in the Electrical and Electronics Engineering Group at the University of Southampton. He graduated with a B.Eng. in Electronic Engineering at the University of Southampton in 2001, and in 2004 was awarded a PhD for the development of switched-current filters and phase locked loops. He spent a year working as a Senior Design Engineer for a high tech startup company, then rejoined the University of Southampton as a Research Fellow in 2006. Since working as a research fellow he has been a named researcher on a number of high profile grants from the EPSRC and ICUK to develop novel analogue ICs and more recently interface circuits and control systems for micro-machined sensors. He has over 25 peer reviewed journal and conference publications.

Peter Wilson (M'99, SM'05) is a Reader in Electronics in the Electrical and Electronics Engineering Group at the University of Southampton. He received the B.Eng. in Electrical and Electronic Engineering and Postgraduate Diploma in Digital Systems Engineering from Heriot-Watt University, Edinburgh, Scotland, in 1988 and 1992 respectively; an M.B.A from the Edinburgh Business School obtained in 1999 and obtained his PhD from the University of Southampton in 2002. He has published more than 80 papers in the areas of Electronic Design Automation, Power Electronics, Magnetics and Integrated circuit design. He has authored a number of books, and has managed numerous funded research projects in the areas of power electronics, integrated circuit design, modeling & simulation and renewable energy.

He worked in the Navigation Systems Division of Ferranti plc., Edinburgh, Scotland from 1988-1990 on Fire Control Computer systems, before moving in 1990 to the Radar Systems Division of GEC-Marconi Avionics, also in Edinburgh, Scotland. During the period 1990-1994 he worked on modeling and simulation of Power Supplies, Signal Processing Systems, Servo and Mixed technology systems. From 1994-1999 he worked as European Product Specialist with Analogy Inc. During this time he developed a number of models, libraries and modeling tools for the Saber simulator, especially in the areas of Power Systems, Magnetic Components and Telecommunications.

Since 1999 he has been working in the Electronic Systems and Devices group and more recently in the newly formed Electrical and Electronic Engineering research Group at the University of Southampton, UK. His current research interests include improving variability in integrated circuits, Analog and mixed Signal IC Design, modeling of magnetic components in electric circuits, power electronics, renewable energy systems, VHDL-AMS modeling and simulation, and the development of electronic design tools. He is a Member of the IEE, Senior member of the IEEE and a Chartered Engineer in the UK. He was Technical Program Chair for IEEE BMAS 2008, General Chair of IEEE BMAS for 2009, and finance chair for ETS 2006, EBTW 2006 and FDL 2010. He is the co-chair of IEEE Std 1076.1.1 and chair of IEEE Std 1573.

The following pages contain the paper "Fully differential electro-mechanical phase locked loop sensor circuit", which was published in Sensors and Actuators A: Physical, Volume 194. It is based on the work presented in Chapter 5.

Sensors and Actuators A 194 (2013) 240–246

Contents lists available at SciVerse ScienceDirect

Sensors and Actuators A: Physical

journal homepage: www.elsevier.com/locate/sna

Fully differential electro-mechanical phase locked loop sensor circuit

Robert Rudolf, Peter R. Wilson*, Li Ke, Reuben Wilcock, Andrew D. Brown, Nick R. Harris

School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK

ARTICLE INFO

Article history:
Received 5 November 2012
Received in revised form 30 January 2013
Accepted 30 January 2013
Available online 19 February 2013

Keywords:
Phase locked loop
Capacitive sensor circuit
MEMS sensor

ABSTRACT

Embedding a micro-machined sensing element in a closed loop, force feedback system is a technique commonly used to realize high performance MEMS (micro-electro-mechanical systems) sensors due to the advantages of better linearity, increased dynamic range and reduced parameter sensitivity. Electro-mechanical sigma delta modulators (EM$\Sigma\Delta$) have been proposed for this reason and high order loops have been shown to enjoy a good signal to noise ratio (SNR) of more than 100 dB. It is also well known that achieving stability in high order EM$\Sigma\Delta$s is a challenging task and in practice stability can be lost with large input signals or due to non-ideal effects in the circuits implemented. In this work we propose a novel differential frequency domain technique for closed loop control of micro-machined sensors. This method, called the electro-mechanical phase locked loop (EMPLL), uses a differential electro-mechanical phase locked loop to control and measure the deflection of micro-machined sensors. We believe that EMPLLs have the potential to have significant advantages over EM$\Sigma\Delta$s for high performance MEMS sensors. Preliminary research suggests that this novel approach will lead to significant benefits in signal to noise ratio, parameter sensitivity, and input signal range.

1. Introduction

The use of electro-mechanical sigma delta modulators (EM$\Sigma\Delta$) for high performance sensors in multiple applications has become pervasive over the last decade or so. One issue for designers of integrated electronic and MEMS circuits has been that while the EM$\Sigma\Delta$ approach offers generally good signal to noise ratio (SNR), of the order of 100 dB, this is often at the price of stability. It is also notoriously difficult to design stable and robust higher order sigma delta modulators. Another design constraint is that the input signal range tends to be limited due to the intrinsic tendency of the sigma delta modulator to become unstable for higher signal levels. In practical integrated systems, the overall design stability is very sensitive to parameter variation. Whether the variation is induced by environmental changes (such as temperature) or degradation over time, the circuits tend to be difficult to design with inherent robustness. This is a particular problem even with optimized parameter sets, as a nominal parameter set is often very sensitive to very small parameter changes.

Various options for improving these aspects of EM$\Sigma\Delta$s have led to a significant research effort in the area of parameter optimization and architecture design, however there is a tension ultimately between the high orders required to achieve good SNR and the resulting parameter sensitivity. The primary goal of this research has therefore been to identify a potential alternative electronic interface circuit to the conventional EM$\Sigma\Delta$ that could provide equivalent SNR performance, but would potentially be easier to design and offer better dynamic range and increased tolerance to parameter variation. This paper will use the standard principles of force feedback capacitive sensors, and offer an alternative to the standard EM$\Sigma\Delta$ modulator approach. A new fully differential technique for sensing changes in capacitive sensors will be introduced and the work will demonstrate how this can be used in a force-feedback control loop. The new approach is compared with a typical 5th order EM$\Sigma\Delta$ modulator to give an indication of the relative merits of the two methods.

2. Background

It has become common practice to include both a micro-electro-mechanical system (MEMS) sensor with an electronic sensor interface circuit, using a closed loop approach where the sensor itself is part of the control loop. MEMS inertial sensors are often based on a capacitive sensing element, and do have an advantage when linked to a $\Sigma\Delta$ modulator in that they provide a digital output that can be connected directly to a digital circuit for further processing. Using a MEMS sensor with a lower order $\Sigma\Delta$ modulator provides second order systems that are simple to design [1,2], relatively stable and have reasonable performance. Unfortunately, the inherent disadvantages of lower order $\Sigma\Delta$ modulators in the electronic domain are also well known and manifest themselves in the integrated MEMS sensor type, including quantization noise,

* Corresponding author. Tel.: +44 23 8059 4162.
E-mail address: prw@ecs.soton.ac.uk (P.R. Wilson).

0924-4247/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sna.2013.01.053

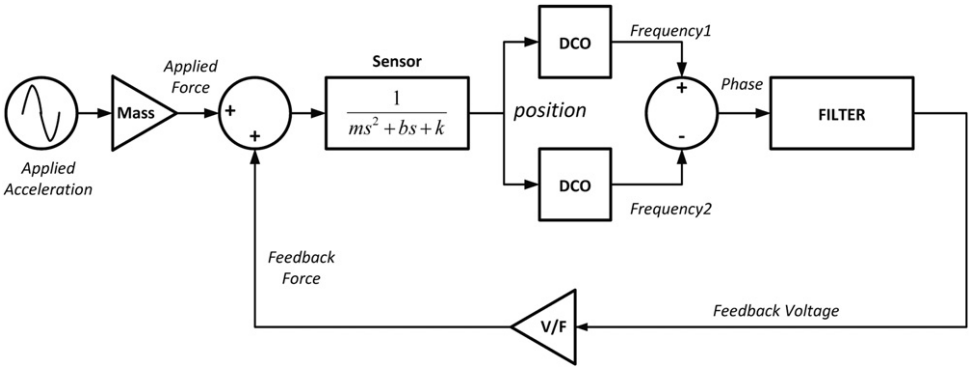


Fig. 1. Architecture of the EMPLL.

dead-zones and the issue of idle tones becoming apparent in the signal bandwidth.

With the increasing requirement for sensitivity and low noise, the basic approach has been to develop higher order $\Sigma\Delta$ modulators, with an integrated EM (electro mechanical) sensor in the loop, with 5th order systems providing excellent signal to noise ratio and overall performance [3,4]. In spite of these efforts, it has become necessary to provide advanced optimization tools [5] to establish the correct $\Sigma\Delta$ design parameters to ensure stability. While this type of approach can calculate the nominal optimum parameters for a circuit, in practice these can be extremely sensitive to component variation. Even if the parameters are designed to be more robust to variation, as suggested in [5], this involves a complex and time consuming optimization process. This research therefore provides insight into the possibility of using a phase locked loop (PLL) sensing circuit rather than $\Sigma\Delta$ based modulator, to establish whether there would be any potential advantages which would alleviate the sensitivity and difficulty in obtaining effective and robust design parameters. The method of applying the principles of a phase-locked loop to a differential force-balanced accelerometer interface circuit has not been previously described in literature. There are previous methods that use parts of this idea, but none show a combination of all three key elements (phase or frequency-based, differential and closed-loop) at the same time.

Matsumoto et al. [6] describes a single ended PLL system where the VCO contains the variable capacitor of the sensor. This means that the PLL will run normally and locked to a reference, but acceleration on the sensor causes an "error" in the VCO frequency which is compensated by changing the VCO input voltage. This system, however, works on a single-ended sensor capacitance and has no electromechanical feedback to the sensor. Matsumoto and Esashi [7] also presented a system which uses a PLL using the sensor as part of a variable-frequency oscillator with the feedback voltage as an input. The capacitor driving the oscillator is also single-ended, with a separate reference capacitor on the MEMS chip. Key differences to the proposed EMPLL are firstly that the capacitance is not differential, but instead use a fairly complex and precision-limited periodically updated counter to obtain the reference frequency for the PLL from a fixed reference on the sensor chip. Secondly the feedback system is also single-ended and therefore the feedback plates are not biased at a certain voltage and finally the loop topology in Matsumoto's system is quite different from the EMPLL, mainly due to the particular implementation of the oscillators.

Kitano et al. [8] use a differential sensor that drives two independent LC oscillators to account for errors from drift and parasitics. They suggest that the potential exists to tune the oscillators however there is no feedback to the mechanical system. In a system for resonant gyroscopes presented by Saukoski et al. [9], there are two control loops, one to drive the resonator and another one for compensation. However, although the resonator part looks like a PLL, this approach does not utilize a frequency that is generated from a variable oscillator. There are numerous other pickoff systems for gyroscopes that work in a very similar fashion such as [10]. Partridge [11] describe MEMS resonators of which they measure the resonance frequency, which changes under acceleration. They present a number of sensing circuits and also provide feedback to the mechanical system that is generated by an unspecified "Control Circuitry". Although they use PLLs to measure the frequency, the feedback is not directly within the PLL control loop – they merely use it to measure the frequency, among other approaches. Hati et al. [12] use stable PLLs and investigate the effects of vibration on frequency references. Although there is no electromechanical feedback as the topology is essentially the same as [6], it is still relevant because of the analysis of closed-loop system behaviour. Yoneoka et al. [13] look at random vibration noise in resonant systems and how it could be compensated, also by using a dedicated accelerometer. Again, there is no electromechanical feedback or indeed a full pickoff circuit, but this work is nevertheless relevant for noise considerations.

In summary, therefore, the basic idea of using a phase locked loop with the capacitive sensor providing some basic control of the oscillator frequency in the loop has some precedent, however the use of a fully differential system has not been described in the literature.

3. The electro-mechanical phase locked loop (EMPLL)

In a conventional electro-mechanical capacitive sensor circuit the variation in capacitance directly modulates the signal applied to the sensor (usually a high frequency "carrier" signal). This is demodulated by a $\Sigma\Delta$ circuit to generate a digital output and also provide a suitable force feedback signal to the sensor to keep it under control, and the system "in lock". In this section, we introduce the EMPLL concept, which is different in that unlike the electromechanical $\Sigma\Delta$ modulator, which measures sensor capacitance directly by means of a charge amplifier, the electromechanical PLL architecture shown in Fig. 1 uses a pair of oscillators with the sensor capacitance as the frequency determining element.

The two sensor capacitances determine the frequency of two oscillators, which are labelled DCO (displacement-controlled oscillator) in Fig. 1. The name displacement-controlled oscillator is in reference to a voltage-controlled oscillator (VCO), but indicates that the oscillation frequency is determined by the displacement of the accelerometer's proof mass rather than a control voltage. Since the capacitors change differentially when the sensor is subject to acceleration, the oscillator frequencies change likewise. The difference in frequency between the two oscillators is thus a measure for the

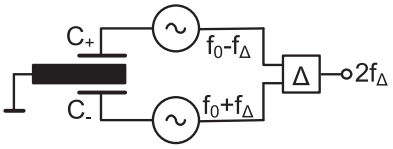


Fig. 2. Frequency difference as a measure for acceleration.

input acceleration. This principle is illustrated in Fig. 2. The outputs of the oscillators are input to a phase detector and a loop filter to suppress the "carrier" and provide a feedback control signal.

Effectively, this entire system behaves like a phase-locked loop, where one of the sensor oscillators represents the reference oscillator and the second oscillator is equal to the feedback controlled VCO. In this configuration, the EMPLL shows several advantages over ΣΔ modulators or conventional pick-off circuits. In the first place, no linear low-noise charge amplifier is required, which simplifies the circuit configuration. In contrast to the I:.6. System, the order does not need to be increased to gain SNR, and this has the secondary effect of making the system performance inherently more robust against parameter variation. Finally, there is no digital switching of feedback voltage, which has a positive impact on the noise in the system.The most important fact about the DCO to note is that ideally the difference in frequency is linearly dependent on acceleration, which is shown below. It should be noted that there is a well defined range of operation for any controlled oscillator and outside this range a practical system will limit the frequency of the oscillator – potentially introducing a non-linearity. In a practical system, therefore, the oscillators should be designed with the centre frequency and sensitivity to match the range of operation of the sensor. The force (F) experienced by the proof mass (m) in the accelerometer under acceleration (a) is:

$$F = ma \tag{1}$$

and the resulting displacement of the proof mass is that of a damped second-order mass-spring-damper system:

$$d = \frac{1}{ms^2 + bs + k} F = k_{\text{sens}}(s)a \tag{2}$$

with

$$k_{\text{sens}} = \frac{m}{ms^2 + bs + k} \tag{3}$$

where, m is the proof mass, b is the damping factor, k the spring constant of the mass-spring system that models the accelerometer and s is a complex number in the Laplace space. Recalling the generic capacitance of a parallel plate capacitor as an approximation for the sensing capacitance:

$$C = E_0 E_r \frac{A}{d} = k_{\text{cap}} \frac{1}{d} \tag{4}$$

with

$$k_{\text{cap}} = E_0 E_r A \tag{5}$$

where, A is the sensor capacitor plate area and d the spacing between the plates. Substituting the instantaneous displacement (d) from Eq. (2) around the nominal capacitor plate spacing (dnom), the sensor capacitance can be written as a function of acceleration (a).

$$C = E_0 E_r \frac{A}{d_{\text{nom}} \pm d} = k_{\text{cap}} \frac{1}{d_{\text{nom}} \pm k_{\text{sens}} a} \tag{6}$$

It is desirable to have a linear relationship between frequency and acceleration, which requires the oscillator frequency to be inversely proportional to capacitance making an RC oscillator the ideal choice. An LC oscillator would generally be the first choice from the point of view of low noise design, however the frequency of such an oscillator is inherently dependent on the inverse root of the capacitance.

$$f_{\text{osc}} = k_{\text{osc}} \frac{1}{C} \tag{7}$$

The parameter $k$osc relates the frequency to capacitance, depending on resistance values and switching thresholds. As can be see in Eq. (6), however, the capacitance is not exclusively dependent on the inverse of the acceleration, but also on the nominal capacitor plate spacing $d$nom. However, this constant term, which may well be larger in magnitude than the change in capacitance due to acceleration, cancels when two oscillators are driven from a pair of differential capacitors (in the symmetrical structure shown in (1) and are subject to the same acceleration. The frequency difference of two such oscillators is:

$$\Delta f = f_{+d} - f_{-d} = \frac{k_{\text{osc}}}{k_{\text{cap}}}(d_{\text{nom}} + d - d_{\text{nom}})$$
$$\Delta f = 2\frac{k_{\text{osc}}}{k_{\text{cap}}} d \tag{8}$$

As we have seen in Eq. (2), the displacement can be replaced with a function of acceleration and as a result the frequency difference can be seen in (9) to be linear with respect to acceleration.

$$\Delta f = 2\frac{k_{\text{osc}}}{k_{\text{cap}}} k_{\text{sens}}(s)a \tag{9}$$

The output voltage proportional to acceleration cannot be applied directly to the feedback plates. The reason for this is that the electrostatic force between the two feedback plates is only attractive, regardless of the polarity of the feedback voltage. It is therefore necessary to apply a certain bias voltage to both feedback plate pairs, causing oppositely directed forces on the proof mass resulting in no net force when under quiescent conditions. The feedback voltage is then superimposed on this bias voltage, resulting in an increase of one and decrease of the other force, resulting in a net feedback force on the proof mass, as depicted in Fig. 3.

Generally, the electrostatic force between two parallel plates is given by Eq. (10).

$$F = k_{\text{force}} \frac{V^2}{d^2} \tag{10}$$

where:

$$k_{\text{force}} = \tfrac{1}{2} E_0 E_r A \tag{11}$$

The total force on the proof mass is the sum of two oppositely directed electrostatic forces, each depending on the bias voltage and the differential feedback voltage as given in Eq. (12).

$$F = k_{\text{force}} \frac{(V_b + V_{fb})^2}{(d_{\text{nom}} - d_\Delta)^2} - \frac{(V_b - V_{fb})^2}{(d_{\text{nom}} + d_\Delta)^2} \tag{12}$$

$$F = k_{\text{force}} \frac{V_b^2 + 2V_bV_{fb} + V_{fb}^2}{(d_{\text{nom}} - d_\Delta)^2} - \frac{V_b^2 - 2V_bV_{fb} + V_{fb}^2}{(d_{\text{nom}} + d_\Delta)^2} \tag{13}$$

Due to the dependence of the force on the instantaneous spacing of the plates, the relationship between feedback force and feedback voltage is not linear. Assuming higher order effects can be ignored due to the small displacements involved, Eq. (14) shows that the feedback force is approximately linearly dependent on feedback voltage.

$$F \approx k_{\text{force}} \frac{4V_bV_{fb}}{d_{\text{nom}}^2} \tag{14}$$

This result means that the entire closed-loop system of the EMPLL is approximately linear, which is a very significant result. Unlike electromechanical ΣΔ modulators which are inherently non-linear and therefore require difficult analysis methods, initial

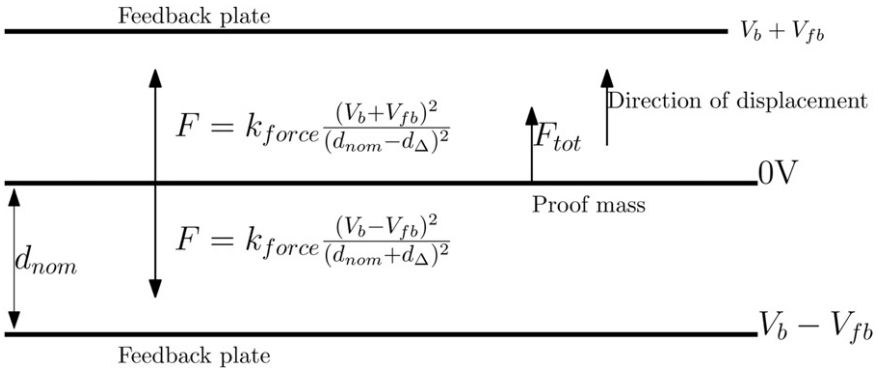


Fig. 3. Forces in the force-feedback system.

design calculations and simulations on the EMPLL can be achieved using a linear approximation. Once the system properties and parameters have been determined roughly using the linear approximation, final tuning can be conducted using the full non-linear models, which also give the most accurate performance metrics.

As an example of this it can be seen from Eq. (12) that in addition to non-linearity introduced by the instantaneous displacement, the feedback voltage also appears as a squared term in the numerator. This causes distortion of the signal, which can be seen in the simulations as a harmonic component. In the EMPLL, however, reduction of this component can be achieved more easily and without any additional circuitry. If the feedback plate bias voltage is increased to be relatively large with respect to the feed- back voltage, the squared term of the feedback voltage becomes insignificant relative to the constant bias voltage and the linear term.

4. Comparison of EMPLL and EMΣΔ systems

4.1. Introduction

In order to evaluate the performance of the proposed approach, three standard tests have been used to compare the new EMPLL circuit with a reference EMΣΔ system. The first test is to calculate power spectral density (PSD) of both the EMPLL and EMΣΔ circuits. The main criteria at this stage is to establish whether the basic performance is comparable between the two approaches. The second test was to compare the response of the two circuits to a wide range of accelerations, and therefore the sensor and circuit combinations were tested up to 15 g, which was the designed range of operation of the sensor. The final test was to evaluate the impact of parameter variations on the two circuit configurations, using equivalent basic parameter tolerances to see how well the two alternative approaches operated.

A selection of simulation results is presented which illustrate the performance and properties of the EMPLL. All simulations were conducted using models in Matlab and the system parameters were optimized using the Cheetah GA system [5]. In order to provide a comparison to existing methods, a 5th order ΣΔ based modulator system was analyzed as a representative reference circuit.

Table 1 lists the sensor parameter used for the simulation of both the EMΣΔ and the EMPLL, Table 2 lists the system values used for the simulation of the EMΣΔ and Table 3 lists the system values used for the simulation of the EMPLL. The variations indicated are used for the simulations in Section 4.4.

4.2. Output power spectrum

As in most sensor interface systems, the output voltage PSD is of most interest to compare system performance – especially when a ΣΔ approach is employed as one of the key advantages is the ability to shape the noise and achieve a significantly improved noise floor. The reference system was tested and the PSD of the reference EMΣΔ system from a time-domain simulation with an input amplitude of 2.5 g and an input frequency of 32 Hz is shown in Fig. 4a. In this reference system the simulated SNR was 108.5 dB for a signal bandwidth of 1024 Hz (it should be noted that the pick-off amplifier was modelled as noiseless, so the source of noise was discretization only).

The EMPLL system was tested in the same manner as the reference EMΣΔ system and the PSD obtained from a time-domain simulation with an input amplitude of 2.5 g and an input frequency of 32 Hz is shown in Fig. 4b. From this PSD plot, the simulated SNR of

Table 1
Sensor parameters.

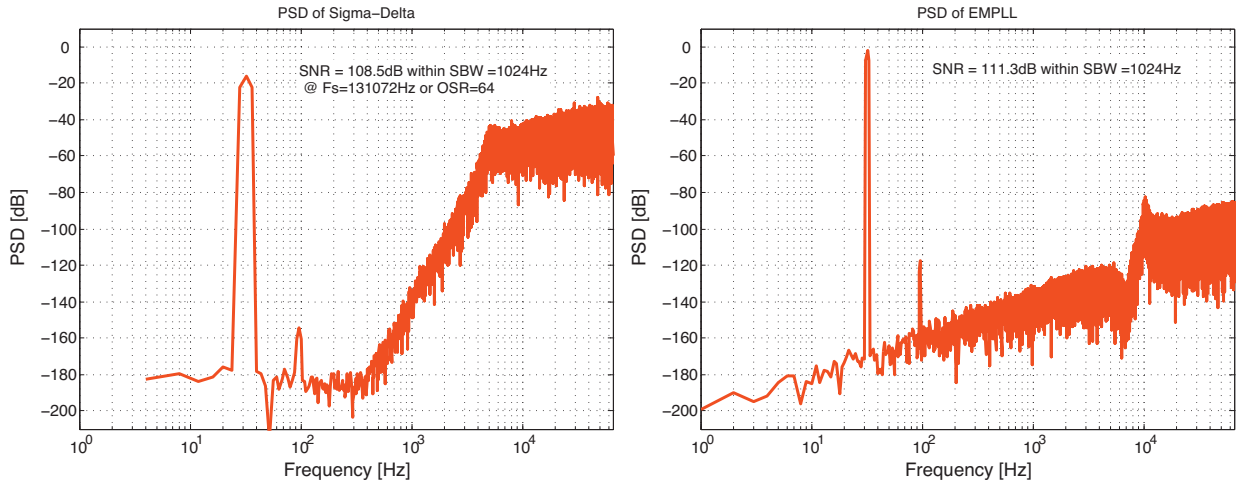
Parameter	Symbol	Value	Unit	Variation
Proof mass	m	1.23×10^{-6}	kg	2%
Damping coefficient	b	9.0×10^{-4}		25%
Spring constant	k	67	$N\,m^{-1}$	5%
Plate spacing	$d$nom	6.5×10^{-6}	m	
Sense plate area	area	2.85×10^{-6}	m2	
Feedback plate area	fb area	1.06×10^{-6}	m2	

Table 2
EMΣΔ system parameters.

Parameter	Symbol	Value	Unit	Variation
Compensator pole	pole	1.57×106	s^{-1}	5%
Compensator zero	zero	2.98×104	s^{-1}	5%
Pickoff amplifier gain	$k$po	4×105		5%
Boost amplifier gain	$k$bst	255.6		2%
Forward gain 1	$k1$	1.114		2%
Forward gain 2	$k2$	0.302		2%
Forward gain 3	$k3$	0.665		2%
Feedback gain 1	$kf1$	0.293		2%
Feedback gain 2	$kf2$	0.898		2%
Feedback gain 3	$kf3$	0.581		2%
Feedback voltage	vfb	18.20	V	2%
Force feedback linearization	$kffl$	7.55		2%

Table 3
EMPLL system parameters.

Parameter	Symbol	Value	Unit	Variation
Lead filter pole	$p_l$ead	6.43×105	s^{-1}	5%
Lead filter zero	$z_l$ead	1.47×104	s^{-1}	5%
Lag filter pole	$p_l$ag	232.84	s^{-1}	5%
Lag filter zero	$z_l$ag	7.54×103	s^{-1}	5%
Bias voltage	Vbias	18.5	V	5%
Compensator gain	$k$comp	0.95		5%
Oscillator gain	kDCO	4.25×1011	$H\,z\,m^{-1}$	5%



(a) PSD of EMΣΔ (b) PSD of EMPLL

Fig. 4. Comparison of PSD between EMΣΔ and EMPLL (a) PSD of EMΣΔ, (b) PSD of EMPLL

the output signal was 111.3 dB, which in this case is slightly better than the reference system.

It has to be noted that the presence of the harmonic component at 96 Hz is clearly visible and measurable. As discussed in Section 3, this harmonic results from the non-linearity in the feedback system and can be reduced by increasing the bias voltage. Furthermore, the noise floor rises with frequency and shows a marked increase above the sensor resonance. At first glance, this looks similar to the result of deliberate noise shaping that occurs in a ΣΔ system seen in Fig. 4a. In the EMPLL, however, this behaviour is not intentional and is merely an artefact from the combined sensor and loop filter transfer functions in the closed-loop system. Indeed, when a generic second-order system with appropriately placed complex poles is used in the loop filter, this increase of noise at high frequencies can be significantly reduced.

4.3. Input amplitude sweep

One significant problem with EMΣΔ systems are stability issues for large input amplitudes, as can be seen in Fig. 5. In this case, the 5th order system can typically cope with accelerations of up to 3 g and then the system will lose stability and become less able to detect the response of the sensor.

In contrast, the results of the EMPLL as shown in Fig. 5 demonstrate a much wider potential range of accelerations possible to be sensed. In a first approximation, the EMPLL is a linear system, which means there are no fundamental restrictions on amplitude performance. There is a certain point beyond which further increase of the input amplitude causes the quadratic term from Eq. (12) to rise above the noise floor which leads to a reduction of overall SNR. It also has to be noted that there is a maximum input amplitude for the EMPLL, although for different technical reasons. In closed-loop control, the accelerometer can be subject to much larger accelerations than without feedback because the feedback force on the proof mass counteracts the force to acceleration. The feedback force, however, is limited in amplitude by the bias voltage. If the acceleration results in a force that is larger than can be compensated by the feedback force, the proof mass can no longer be co ntrolled. Rearranging and evaluating Eqs. (10) and (1) at the bias voltage, while ignoring displacement results in the expression in (15) for the maximum acceleratio n.

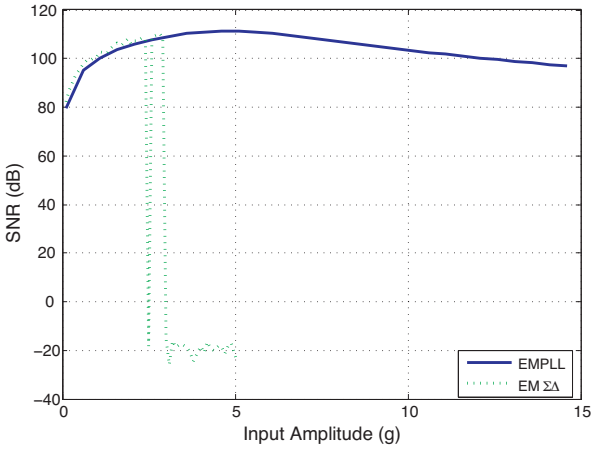


Fig. 5. Input amplitude sweep to 15 g for EMΣΔ and EMPLL.

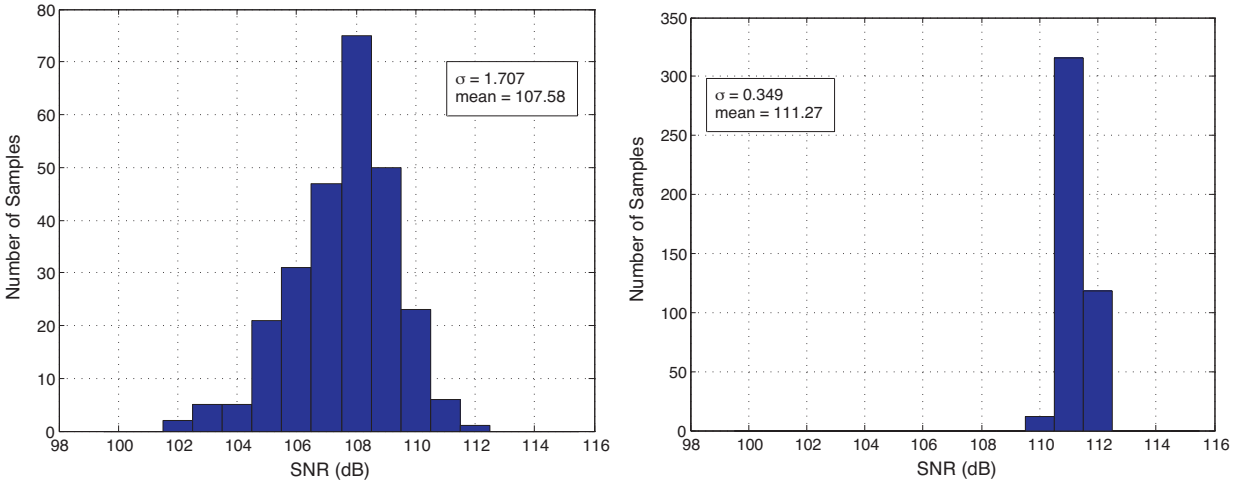
$$a_{\max} = 2\frac{k_{\text{force}}}{m}\frac{V_{\text{bias}}^2}{d_{\text{nom}}^2} \tag{15}$$

The factor 2 in the equation is a result of there being two feedback system driven by a differential feedback voltage, exerting twice the force of a single system. Evaluating Eq. (15) for nominal sensor parameters and a bias voltage of 20 V results in a maximum acceleration of about 70 g. Note that since the bias voltage influences this term quadratically, this figure drops significantly for lower bias voltages. At 12 V bias, for example, the maximum acceleration the system can handle is 26 g.

4.4. Parameter variation

In order to investigate the EMPLL and EMΣΔ circuit's susceptibility to device parameter variation, Monte Carlo simulations were performed based on the values given in Tables 2 and 3. The yield of EMΣΔ systems having a SNR of 90 dB or greater in 500 runs was 53.2%. The SNR distribution of the EMΣΔ systems is shown in Fig. 6a, with a mean SNR of 107.6 dB. The standard deviation in SNR is 1.707 dB, corresponding to 1.6% of the mean value. The yield of EMPLL systems having a SNR of 90 dB or greater in 500 runs was 89.2%. The SNR distribution of the systems is shown in Fig. 6b. It can be seen that the distribution is narrower than the one of the EMΣΔ, having a standard deviation of 0.349 dB, corresponding to 0.31% of the mean value.

This result is a significant improvement over ΣΔ systems, which are very susceptible to parameter variation. This can once again



(a) EMΣΔ (b) EMPLL

Fig. 6. SNR histograms of EMΣΔ and EMPLL under parameter variation (a) EMΣΔ, (b) EMPLL

be attributed to the approximately linear system structure of the EMPLL. The histograms in Fig. 6 show the effect of SNR degradation.

4.5. Summary of results

The comparison with a 5th order EMΣΔ System is useful, as this is a very typical system used in many applications. While it is true that each individual parameter could possibly be improved in the EMΣΔ system, looking at all three aspects (SNR, Amplitude Range and Variation Tolerance), the EMPLL simulation results indicate an exciting alternative approach which is currently being tested practically.

5. Conclusion

In this paper we have described a novel differential frequency domain technique for closed loop control of micro-machined sensors. This method, called the electro-mechanical phase locked loop (EMPLL), uses a differential electro-mechanical phase locked loop to control and measure the deflection of MEMS sensors. Preliminary results indicate that EMPLLs have the potential to have significant advantages over EMΣΔs for high performance MEMS sensors. In particular we have shown there are three areas where this novel approach will lead to significant benefits over previous approaches which are Signal to Noise Ratio, Parameter Sensitivity, and Input Signal Range.

Our tests have shown that for the same sensor, the EMPLL and a 5th order EMΣΔ circuit both provide an SNR performance of around 110 dB for the same signal bandwidth, which demonstrate the same fundamental noise performance. A striking difference, however, is the ability of the EMPLL circuit to tolerate much greater levels of acceleration, with nearly 100 dB of SNR achieved up to nearly 15 g, indicating a much higher tolerance than the equivalent EMΣΔ circuits. Finally, it is well known that the EMΣΔ circuits are extremely sensitive to parameter variations and the EMPLL circuits demonstrate an improved tolerance to those variations with an 80% reduction in variance of SNR of the EMPLL over the EMΣΔ circuit.

In summary, this paper has not only shown that the EMPLL approach can provide similar SNR performance to a conventional EMΣΔ, but that in addition it has the benefit of a much wider range of input acceleration for an identical sensor and also that the circuit offers a very robust system that is tolerant to variations in both the mechanical parts of the sensor, but also the electronic circuit.

References

[1] M.K.X. Jiang, J. Seeger, B. Boser, A monolithic surface micromachined *z*-axis gyroscope with digital output, in: Symposium on VLSI Circuits, Hawaii, USA, 2000, pp. 16–19.
[2] M. Kraft, C.P. Lewis, T.G. Hesketh, Closed Loop Silicon Accelerometers. IEE Proc.–Circuits, Devices and Systems 145 (5) (1998) 325–331.
[3] Y. Dong, M. Kraft, C.O. Gollasch, A high performance accelerometer with fifth order sigma delta modulator, Journal of Micromechanical and Microengineering 15 (2005) 1–8.
[4] J. Luo, H. Ding, M. Kraft, A new design methodology for electro-mechanical sigma-delta-modulators, in: 4th IEEE International Conference on Nano/Micro Engineered and Molecular Systems, NEMS 2009, 2009, pp. 881–884.
[5] R. Wilcock, M. Kraft, Genetic algorithm for the design of electro-mechanical sigma delta modulator MEMS sensors, Sensors 11 (2011) 9217–9232.
[6] Y. Matsumoto, M. Nishimura, M. Matsuura, M. Ishida, Three-axis soi capacitive accelerometer with pll c–v converter, Sensors and Actuators A: Physical 75 (1) (1999) 77–85.
[7] Y. Matsumoto, M. Esashi, Integrated silicon capacitive accelerometer with pll servo technique, Sensors and Actuators A: Physical 39 (3) (1993) 209–217.
[8] T.M. Mayo Kitano, H. Sumi, Capacitance difference detecting circuit and MEMS sensor, US Patent (2006) 7,119,550.
[9] M. Saukoski, L. Aaltonen, T. Salo, K.A. Halonen, Interface and control electronics for a bulk micromachined capacitive gyroscope, Sensors and Actuators A: Physical 147 (1) (2008) 183–193.
[10] R. Fang, W. Lu, C. Liu, Z. Chen, Y. Ju, G. Wang, L. Ji, D. Yu, A low- noise interface circuit for mems vibratory gyroscope, in: 10th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT), 2010, pp. 1456–1458.
[11] M.L. Aaron Partridge, Frequency and/or phase compensated microelectromechanical oscillator, US Patent (2007) 6,995,622.
[12] C.W.N.A. Hati, D.A. Howe, Phase-lock loops in vibration environments, in: 41st Annual Precise Time and Time Interval (PTTI) Meeting, 2009.
[13] S. Yoneoka, J. Salvia, G. Bahl, R. Melamud, S. Chandorkar, T. Kenny, Active electrostatic compensation of micromechanical resonators under random vibrations, Journal of Microelectromechanical Systems 19 (5) (2010) 1270–1272, http://dx.doi.org/10.1109/JMEMS.2010.2067444.

Biographies

Robert Rudolf graduated with a B.Eng. in Electronic Engineering from the University of Southampton in 2010. He is currently pursuing a Ph.D. in the Electronics and Electrical Engineering group in the School of Electronics and Computer Science at the University of Southampton, where he works on configurable analogue circuits. His research interests include analogue and mixed-signal circuit design.

Dr. Peter R. Wilson received the B.Eng. (Hons.) in Electrical and Electronic Engineering from Heriot-Watt University, Edinburgh, Scotland, in 1988; an M.B.A. from the Edinburgh Business School, Scotland in 1999, and Ph.D. from the University of Southampton, 2002. He worked for Ferranti in Edinburgh, Scotland from 1988 to 1994. From 1994 to 1999 he worked for Analogy Inc., based in Swindon, UK and Oregon, USA. Dr. Wilson is currently a Reader in Electronic and Electrical Engineering at the University of Southampton, UK. His current research interests include modelling of magnetic components in electric circuits, power electronics, renewable energy systems, integrated circuit design, modelling and simulation. He is a Senior Member of the IEEE, Member of the IET and a Chartered Engineer.

Dr. Li Ke obtained a B.Eng. in Electrical and Electronic Engineering (2004) from Beijing Union University, M.Sc. in Communications (2005) and Ph.D. (2010) from the University of Southampton. Since 2010 He has been working as a Research Fellow at the University of Southampton on Analogue IC design. His research interests include Phase Locked Loop design, PVT tolerant design and Silicon Photonics.

Dr. Reuben Wilcock received the B.Eng. degree in Electronic Engineering from the University of Southampton, UK, in 2001 and the Ph.D. degree in Electronic Engineering from the University of Southampton in 2004. He was appointed as a research fellow in the Electronic Systems and Devices research group in the University of Southampton in 2004. His current research interests include high frequency, low power, analogue circuit design in CMOS, and design for yield techniques for deep submicron process technologies. He is a member of the IET.

Andrew D. Brown obtained his first degree in Physical Electronics at Southampton University in 1976, and a PhD in Micro-electronics in 1981. He has been a member of academic staff in the Department since 1980 and became a Professor in 1999. He spent time at IBM Hursley Park UK as an IBM Visiting Scientist in 1983, at Siemens NeuPerlach (Munich) as a Visiting Professor in 1989, and at Multiple Access Communications Ltd. in 1994. He was appointed to the Established Chair of electronics at Southampton University in 1999. Dr. Brown is a Fellow of the IEE, a Chartered Engineer, a Senior Member of the IEEE and a registered European Engineer.

Dr. Nick Harris is a senior lecturer in the School of Electronics and Computer Science at the University of Southampton, Southampton, UK, and he has more than 125 publications in the fields of energy harvesting, wireless sensor networks, biosensors, microfluidic systems, and microsystem process development. He was awarded a Ph.D. degree in 1997 by the University of Southampton, Department of Mechanical Engineering. In 1998, Dr. Harris was appointed principal engineer at ERA Technology in the Microsystems and Materials Group before returning to the University of Southampton as a senior research fellow. Current research areas include self-powered health and usage monitors (HUMS), embedded condition-monitoring microsystems, electrochemical sensors and both electromagnetic and piezoelectric planar profile vibration energy harvesters. His other research interests include environmental wireless sensor networks, and analogue and digital electronic sensor circuit design. He is a member of the IET, a Chartered Engineer, director of ECS Partners, and a co-founder of Perpetuum Ltd., the world's leading vibration energy harvesting company.

Appendix B

Chip Schematics

This appendix contains select schematic diagrams of the DAC chip from Chapter 4.

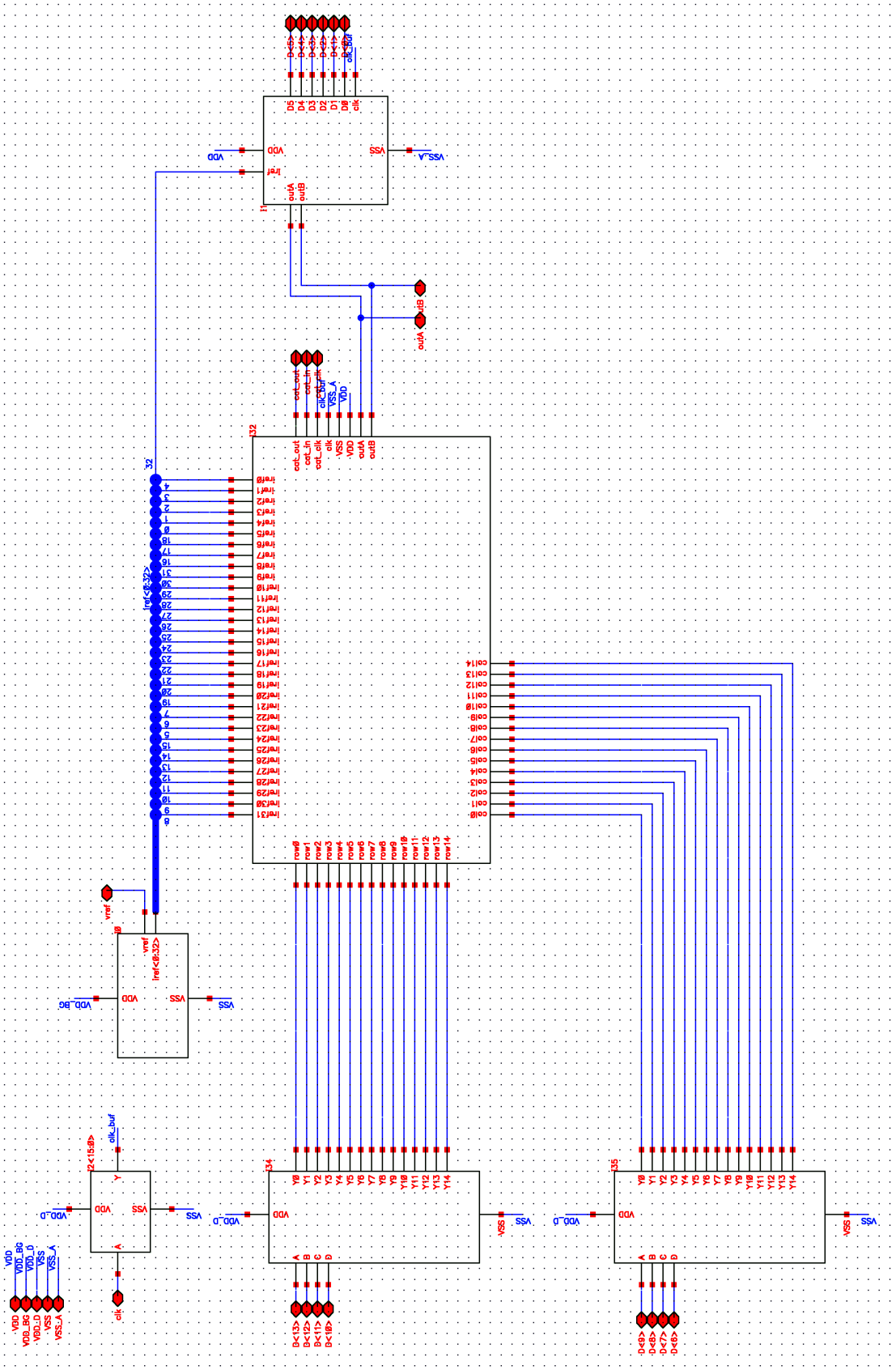


Figure B.1: Top-level schematic diagram of the DAC.

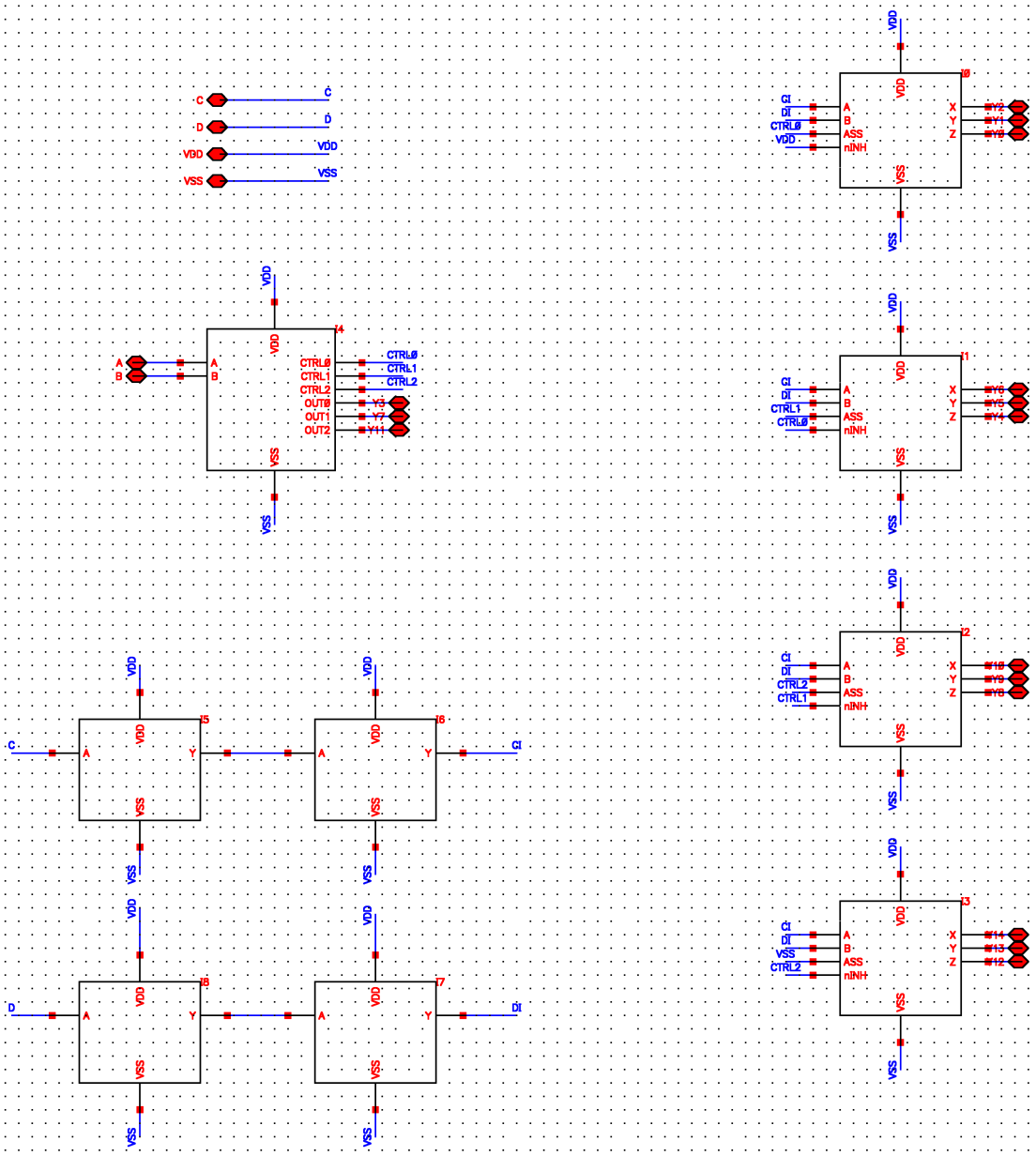


Figure B.2: High-level schematic diagram of the 4-to-15 thermometer decoder.

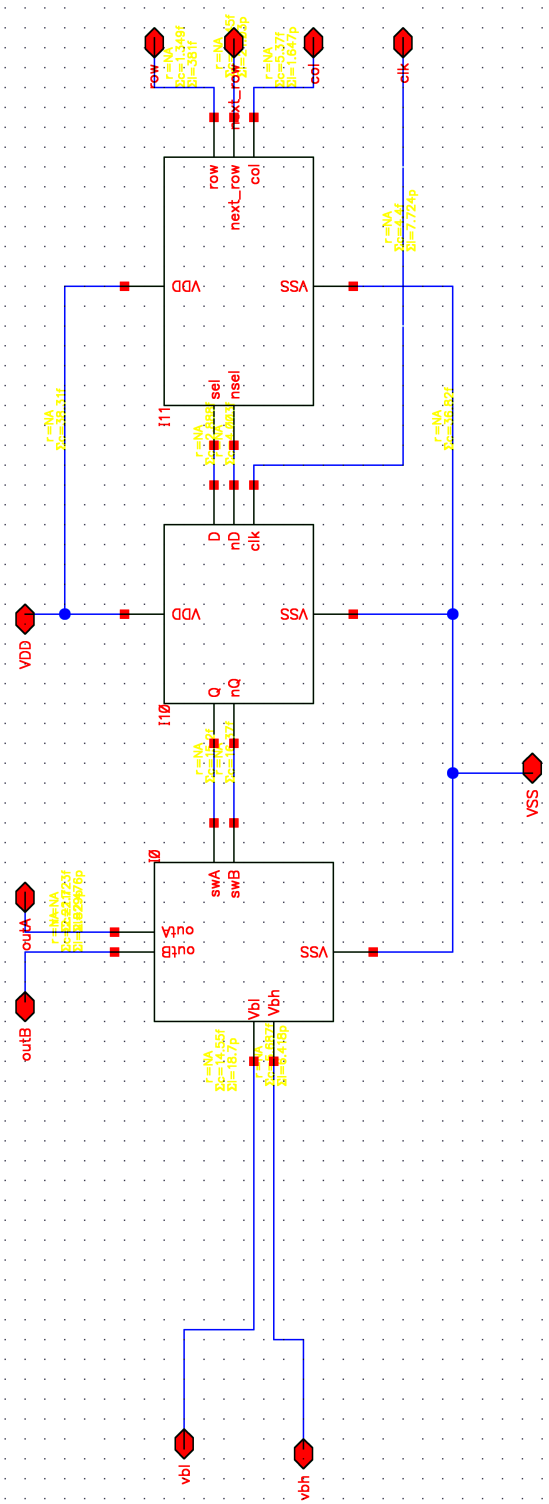


Figure B.3: High-level schematic diagram of a current cell in the unary array.

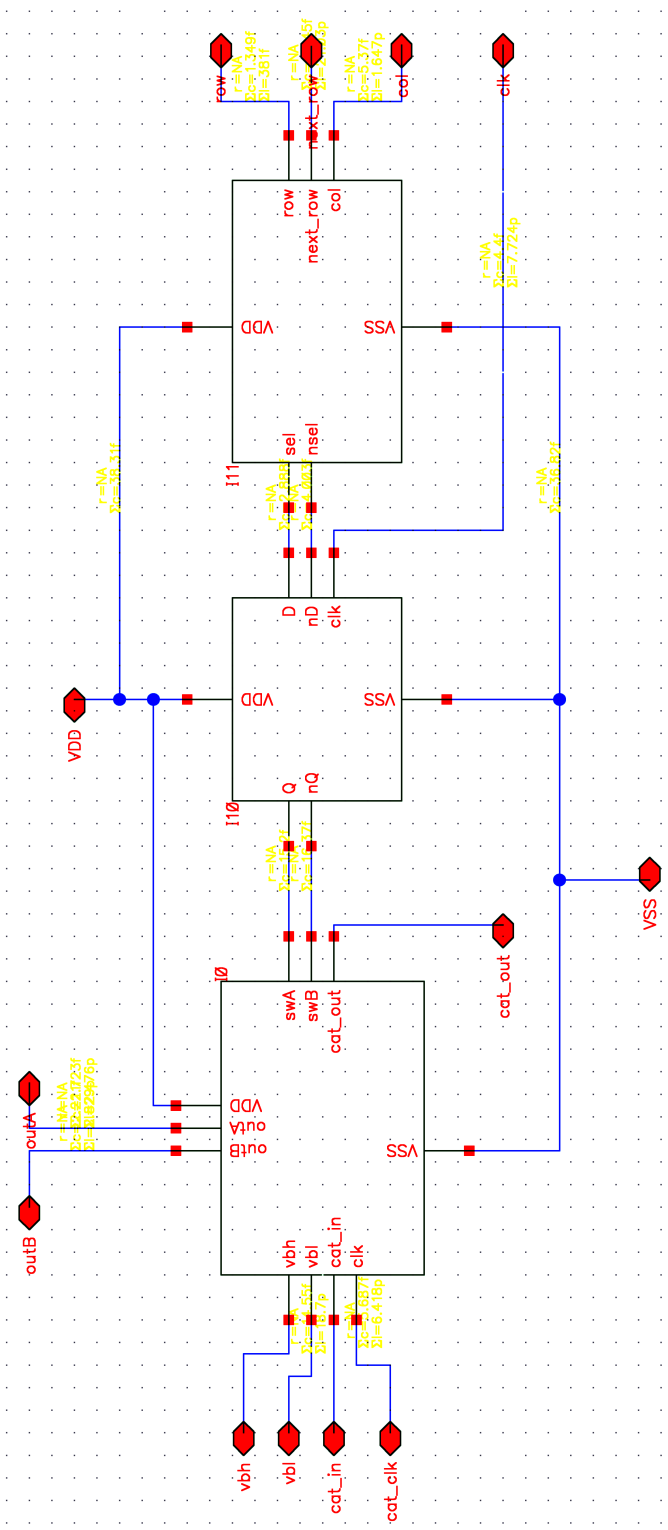


Figure B.4: High-level schematic diagram of a current cell in the unary array with CAT.

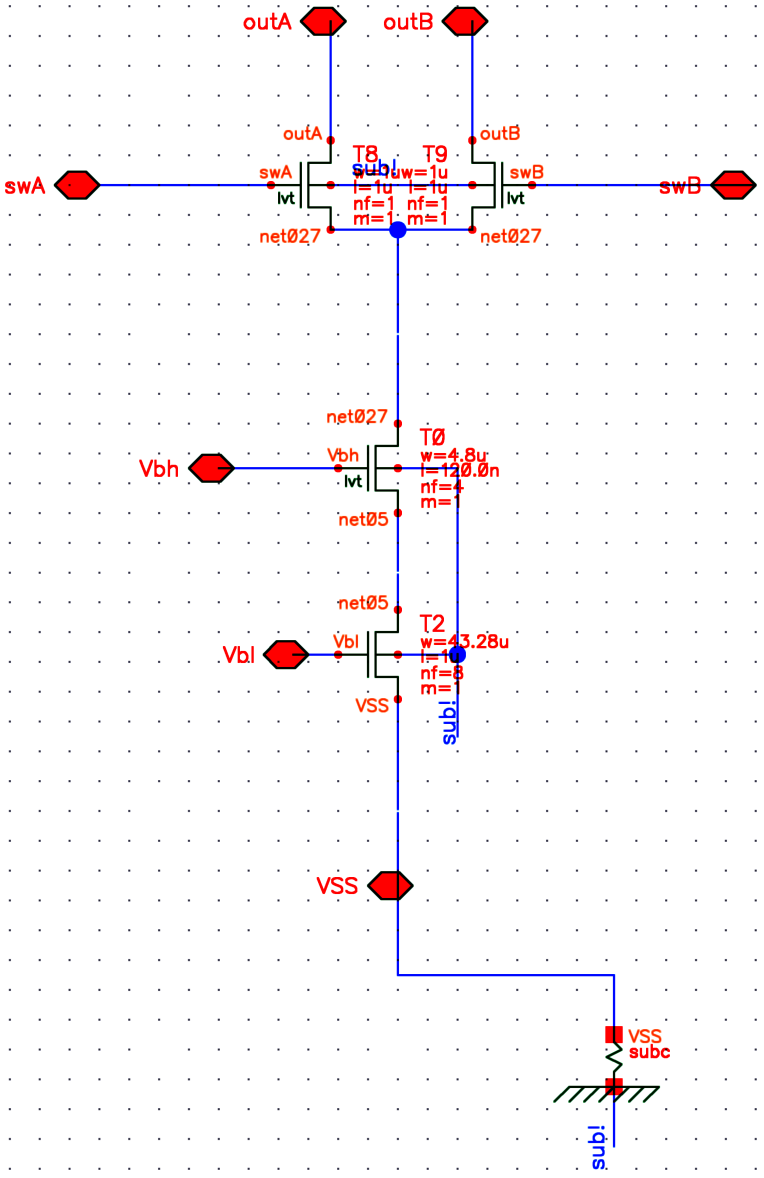


Figure B.5: Transistor-level schematic diagram of a $64\mu A$ current sink and switching transistors.

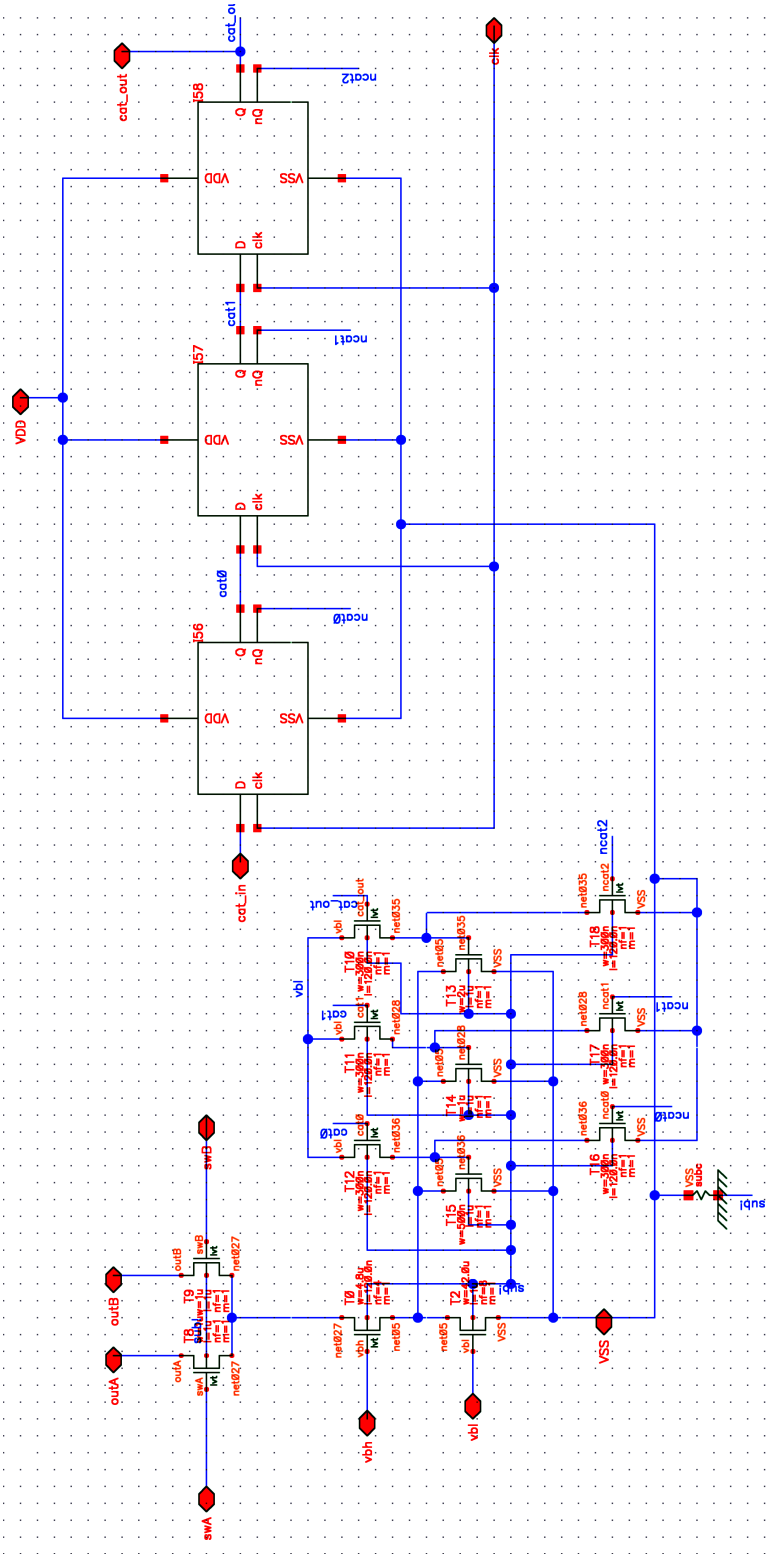


Figure B.6: Transistor-level schematic diagram of a $64\mu A$ current sink with CAT and switching transistors.

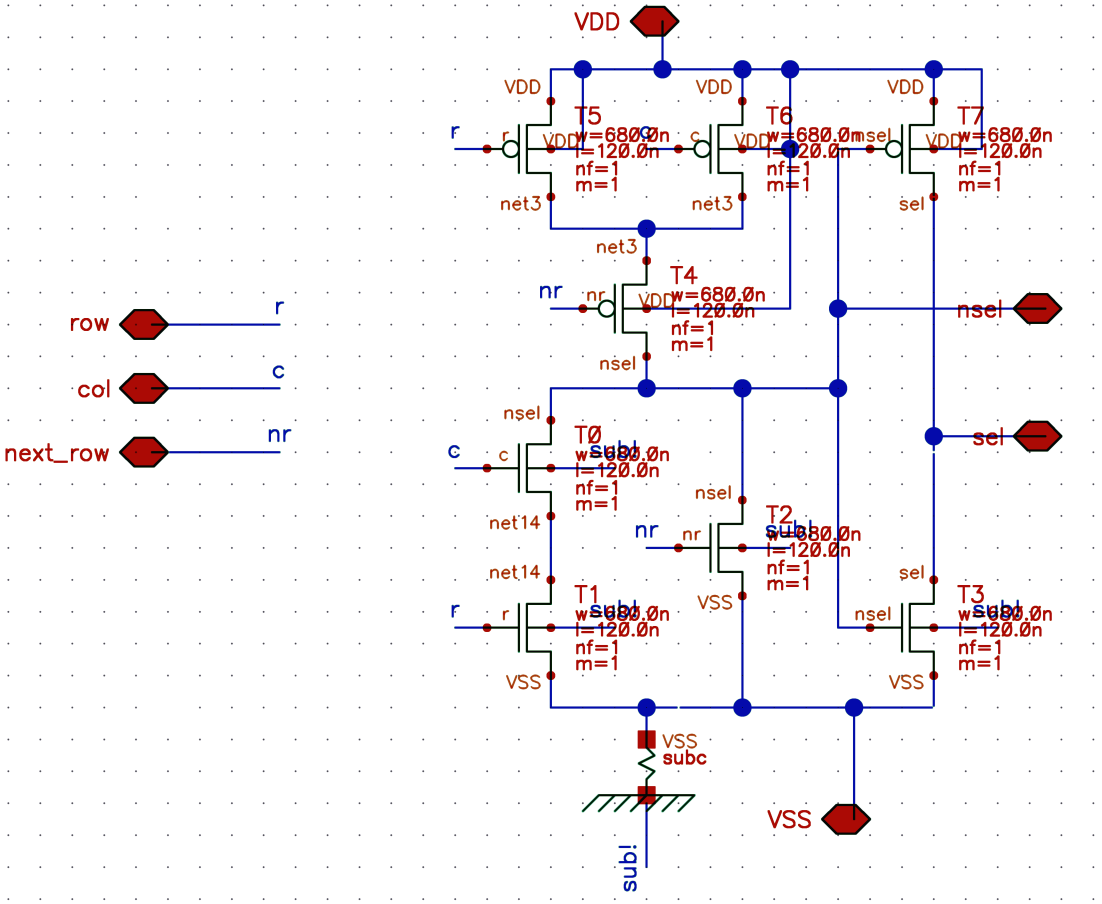


Figure B.7: Transistor-level schematic diagram of the cell select logic of a current cell.

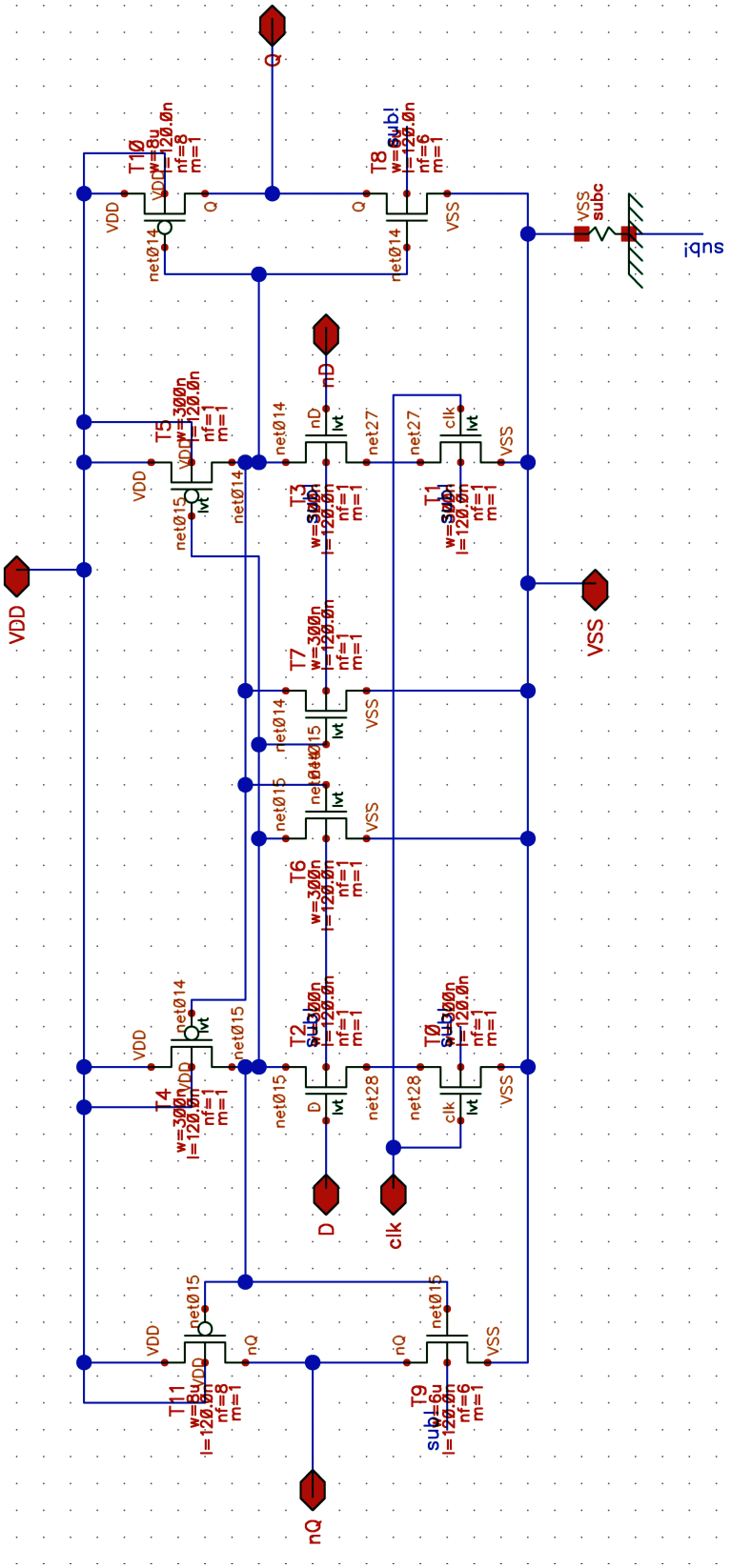


Figure B.8: Transistor-level schematic diagram of the data latch of a current cell.

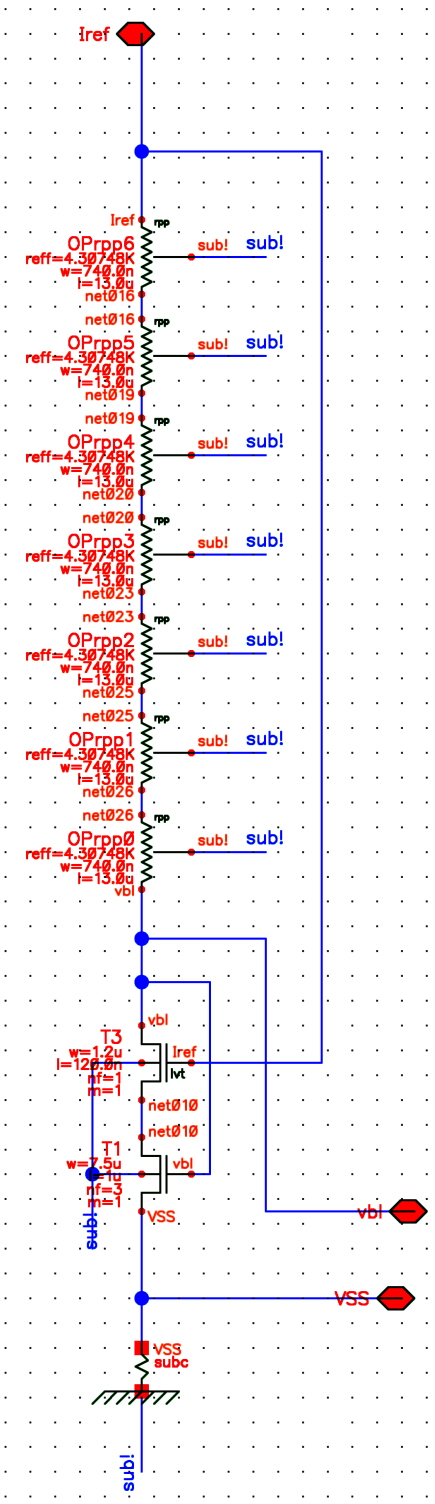


Figure B.9: Transistor-level schematic diagram of the bias generator for 8 current cells.

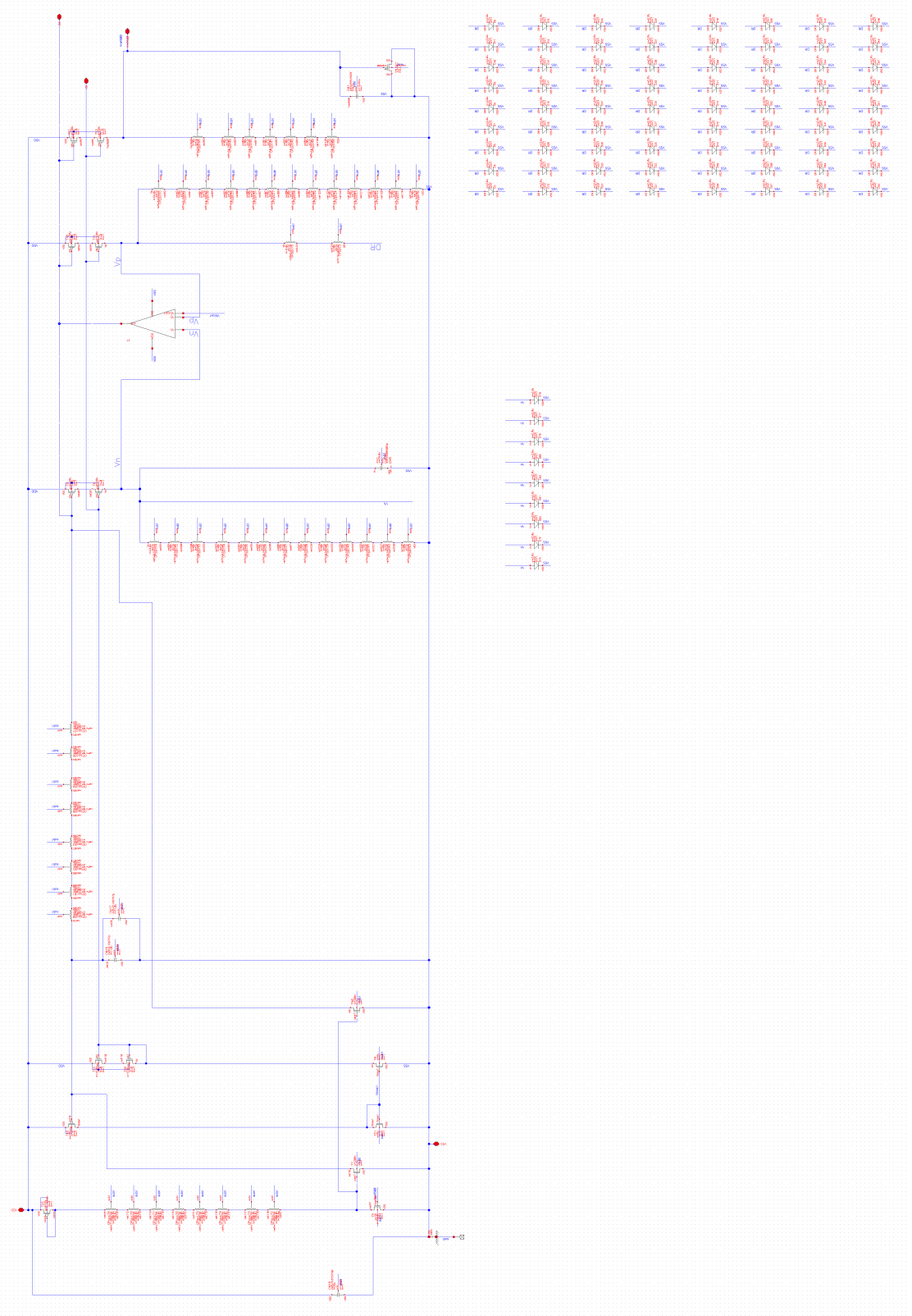


Figure B.10: Transistor-level schematic diagram of the bandgap core (designed by Dr Li Ke).

Appendix C

Chip Layout

This appendix contains layout details of the DAC chip from Chapter 4.

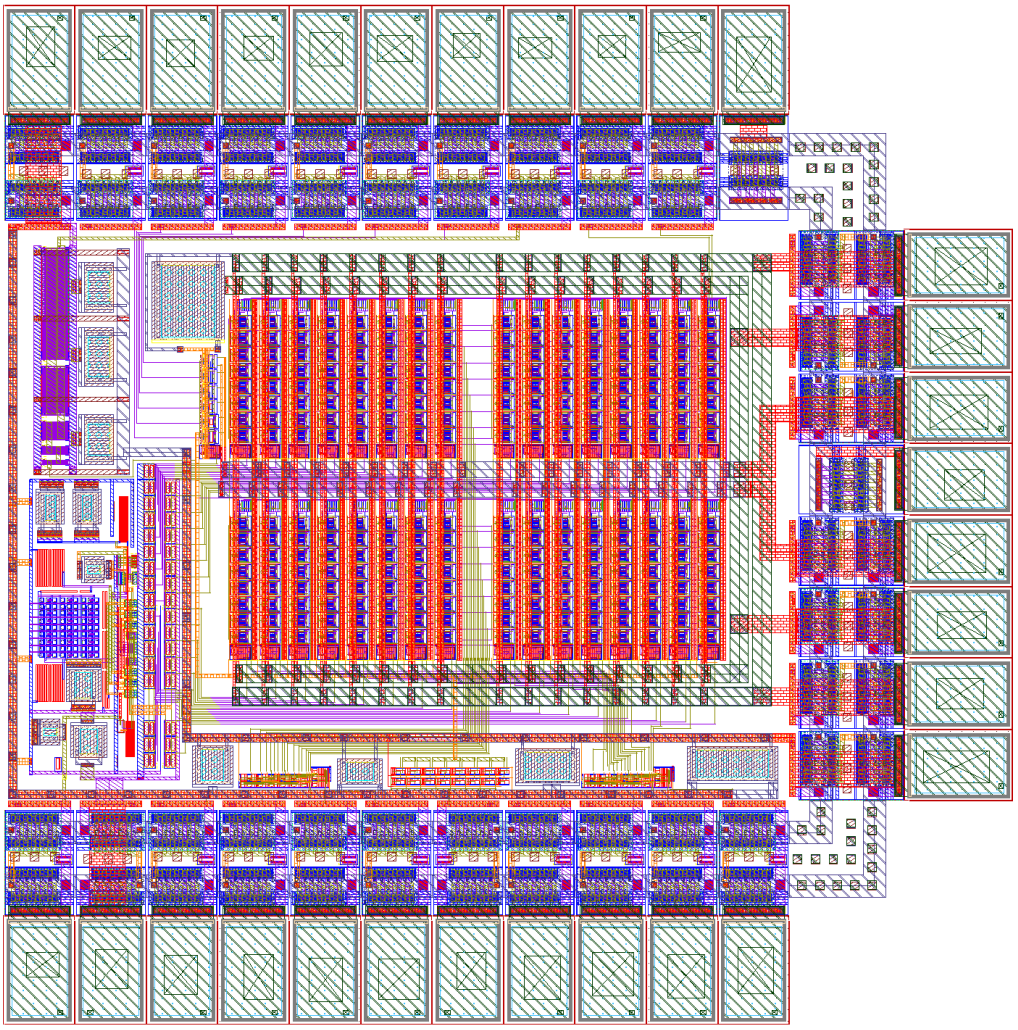


Figure C.1: Top-level view of the DAC with pads.

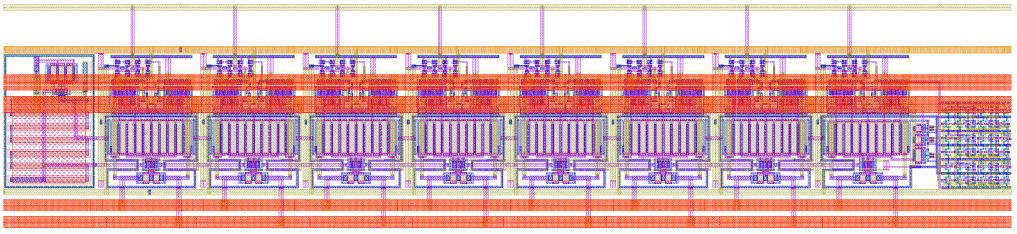


Figure C.2: Layout of a half-row block (8 current cells and bias generator).

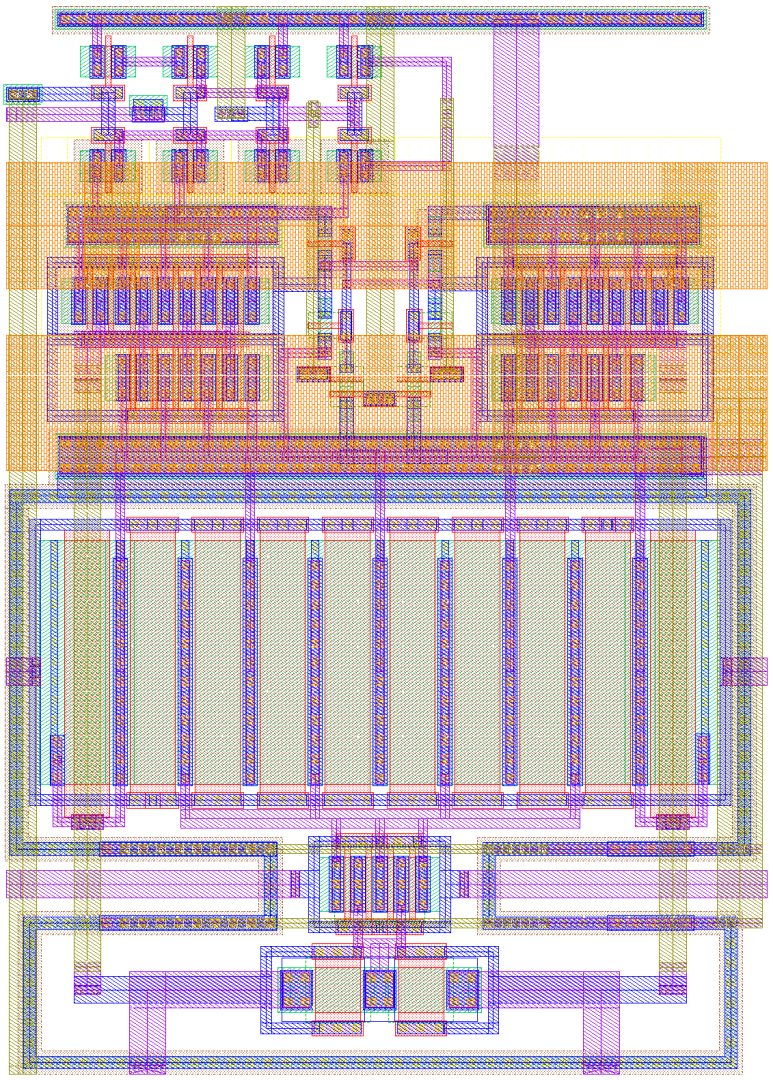


Figure C.3: Layout of a regular $64\mu A$ current cell.

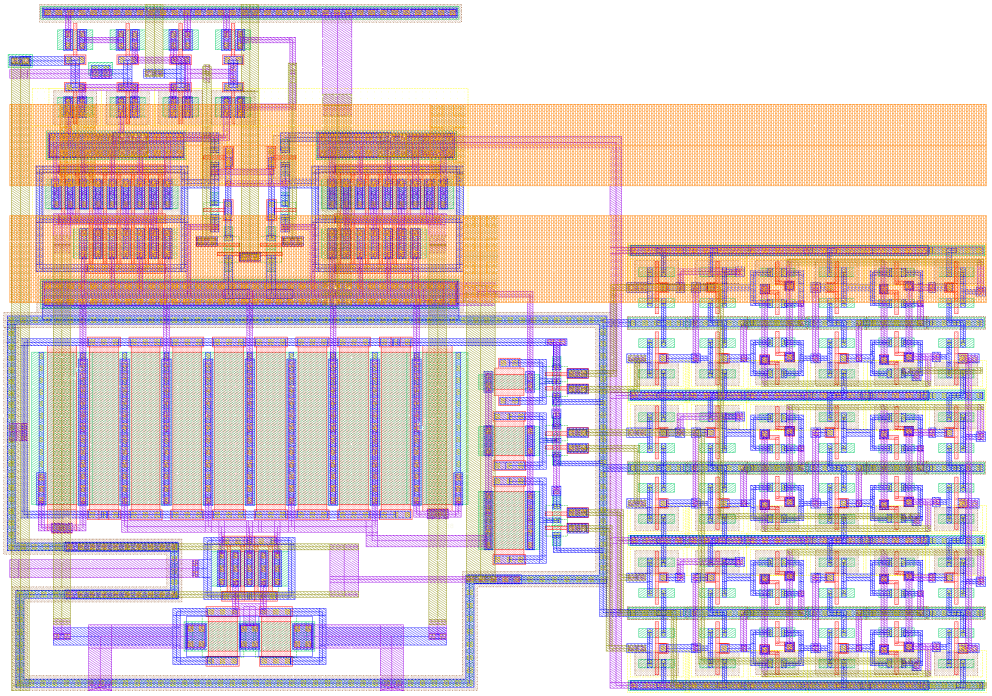


Figure C.4: Layout of a $64\mu A$ current cell with CAT (shift register on the right).

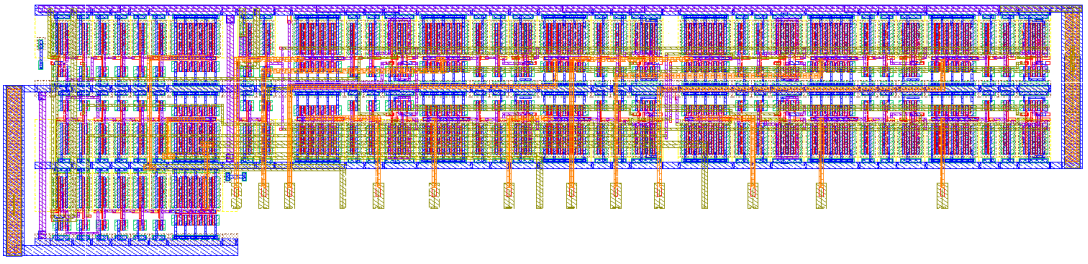


Figure C.5: Layout of the 4-to-15 thermometer decoder.

Appendix D

DAC Calibration Code

D.1 Calculation of INL

```
% Calculate INL of a transfer function at each point
%
% arguments:    meas        transfer function(s)

function inl = calc_inl2(meas)
    foo = meas - (diag(meas(1,:))*ones(size(meas')))';
    ideal = zeros(size(meas));
    for i = 1:size(meas, 2)
        ideal(:,i) = linspace(0,foo(size(foo,1),i),size(meas,1));
    end

    inl = foo-ideal;
end
```

D.2 Application of CAT to a Transfer Function

```
% Simulate INL of a differential transfer function with CAT applied.
%
% arguments:    cat         array of currents supplied by each CAT
%                           (e.g. [1 1 2 1].*1e-6 for four CATs)
%
% return value: inl_pp      peak-to-peak INL of the resultant transfer
% function
%
```

```
% base variabke: meas       uncompensated transfer function
%                ds_factor  ratio at which to insert CAT cells

function inl_pp = fitness_inl2_diff(cat)

    %read base variables
    meas = evalin('base', 'ref');
    ds_factor = evalin('base', 'ds_factor');

    %apply CAT
    for i = 2:length(cat)
       meas = apply_diffcat(meas, (i-1)*ds_factor, cat(i));
    end

    %calculate INL
    inl = calc_inl2(meas);
    inl_pp = max(inl) - min(inl);
end
```

D.3 Determine Optimal CAT Settings

```
%find CAT configuration that minimises INL for a given transfer function

%number of CATs
n_cats = 32;

%CAT current: 1/2 sum of all slices (e.g. 570nA LSB slices -> 4uA CAT ->
%cat_current = 2uA
cat_current = 2e-6;

%copy measurement to other variable
%m_meas(:,1) contains the TF with all CATs zero
%m_meas(:,2) contains the TF with CATs set to a pattern
ref = m_meas(:,1);

%downsample factor = ratio between CAT cells and all measured cells
ds_factor = size(m_meas, 1) / n_cats;

%bounds for solver
lb = -cat_current .* ones(1, length(ref)/ds_factor);
```

```
ub = cat_current .* ones(1, length(ref)/ds_factor);

%set up bins for discrete CAT configurations (3 slices -> 8 bins)
bins = linspace(-cat_current, cat_current, 8);

%set up solver
options = optimset('Display', 'iter', 'TolX', 1e-12);

%optimise peak-to-peak INL
solution = fmincon(@fitness_inl2_diff, initialpop, [], [],[],[],lb, ub,
    [], options);

%calcualte INL of uncompensated transfer function
%stored in inl2(:,1)
inl2 = zeros(length(ref), 2);
inl2(:,1) = calc_inl2(ref);

%apply ideal CAT to transfer function
meas = ref;
for i = 2:length(solution)
    meas = apply_diffcat(meas, (i-1)*ds_factor, solution(i)) ;
end

%calculate compensated INL with infinite granularity CAT
%stored in inl2(:,2)
inl2(:,2) = calc_inl2(meas);


%assign CAT currents to their respective closest bins
for i = 1:length(solution)
    best_bin = 0;
    best_val = 9e99;
    for j = 1:length(bins)
        diff = solution(i) - bins(j);
        if abs(diff) < abs(best_val)
            best_val = diff;
            best_bin = j;
        end
    end
    solution(i) = best_bin;
end
```

```
60
61 %apply CAT with binned currents to transfer function
62 meas = ref;
63 for i = 2:length(solution)
64     meas = apply_diffcat(meas, (i-1)*ds_factor, bins(solution(i))) ;
65 end
66
67 %calculate compensated INL with binned CAT
68 %stored in inl2(:,3)
69 inl2(:,3) = calc_inl2(meas);
70
71 %calculate compensated INL with real set CAT
72 %stored in inl2(:,4)
73 inl2(:,3) = calc_inl2(m_meas(:,2));
74
75 %draw INLs and calculate peak-to-peak INLs
76 xaxis = linspace(0,255*64,size(m_meas, 1));
77
78 figure;
79 plot(xaxis, inl2);
80
81 before_pp = max(inl2(:,1))- min(inl2(:,1));
82 ideal_pp = max(inl2(:,2))- min(inl2(:,2));
83 bin_pp = max(inl2(:,3))- min(inl2(:,3));
84 real_pp = max(inl2(:,4))- min(inl2(:,4));
85
86 legend(sprintf('before: %0.3e', before_pp), sprintf('ideal: %0.3e',
       ideal_pp), sprintf('bin: %0.3e', bin_pp), sprintf('real: %0.3e',
       real_pp));
```

Appendix E

DAC Interface Protocol

This appendix contains the specifications of the interface protocol used to control the DAC test board. It uses the USB-UART device of the mbed microcontroller board for communication.

Commands are sent from the PC, and end in either carriage return, line feed, or any combination. The controller does not respond to or acknowledge commands, expect when they involve reading back data.

The commands and their syntax are shown in Table E.1.

Command	Argument	Response	Description
`dac`	Integer	none	Sets the DAC input to the value given by the argument
`cat`	12 2-digit hex numbers	12 times `writing: 0xAA - read back: 0xBB`, where `0xAA` is the byte written into the CAT scan chain and `0xBB` is the byte received from the CAT scan chain	Sets the CAT bits to given values
`rel0`	none	none	Turns the external relay off
`rel1`	none	none	Turns the external relay on

Table E.1: DAC test board commands

Appendix F

EM-PLL Models and Parameters

F.1 Model Parameters

Parameter	Symbol	Value	Unit	Variation
Proof mass	m	$1.23 \cdot 10^{-6}$	kg	2%
Damping coefficient	b	$9.0 \cdot 10^{-4}$		25%
Spring constant	k	67	$N \cdot m^{-1}$	5 %
Plate spacing	d_{nom}	$6.5 \cdot 10^{-6}$	m	
Sense plate area	$area$	$2.85 \cdot 10^{-6}$	m^2	
Feedback plate area	fb_area	$1.06 \cdot 10^{-6}$	m^2	

Table F.1: Sensor parameters

Parameter	Symbol	Value	Unit	Variation
Compensator pole	$pole$	$2.43 \cdot 10^6$	Hz	5%
Compensator zero	$zero$	$34.19 \cdot 10^3$	Hz	5%
Pickoff amplifier gain	k_{po}	$400 \cdot 10^3$		5%
Boost amplifier gain	k_{bst}	205.7		2%
Forward gain 1	k_1	0.461		2%
Forward gain 2	k_2	0.811		2%
Forward gain 3	k_3	0.927		2%
Feedback gain 1	kf_1	0.419		2%
Feedback gain 2	kf_2	0.502		2%
Feedback gain 3	kf_3	1.132		2%
Feedback voltage	v_{fb}	23.11	V	2%
Force feedback linearisation	k_{ffl}	9.62		2%

Table F.2: EM-$\Sigma\Delta$ system parameters

Parameter	Symbol	Value	Unit	Variation
Lead filter pole	p_{lead}	$1.02 \cdot 10^6$	s^{-1}	5%
Lead filter zero	z_{lead}	$42.02 \cdot 10^3$	s^{-1}	5%
Lag filter pole	p_{lag}	$87.73 \cdot 10^3$	s^{-1}	5%
Lag filter zero	z_{lag}	$4.03 \cdot 10^3$	s^{-1}	5%
Bias voltage	V_{bias}	35.17	V	5%
Compensator gain	k_{comp}	$308.08 \cdot 10^{-3}$		5%
Oscillator gain	k_{DCO}	$425 \cdot 10^9$	$Hz \cdot m^{-1}$	5%

Table F.3: EMPLL system parameters

F.2 Simulation models

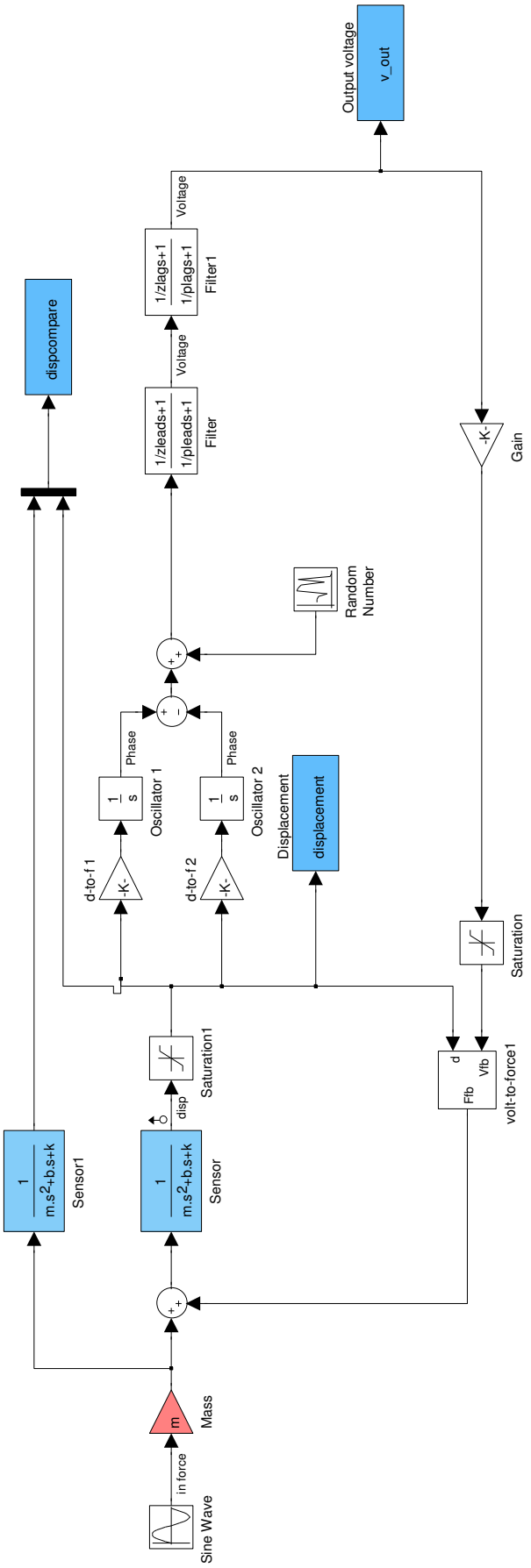


Figure F.1: Simulink block diagram of the EM-PLL

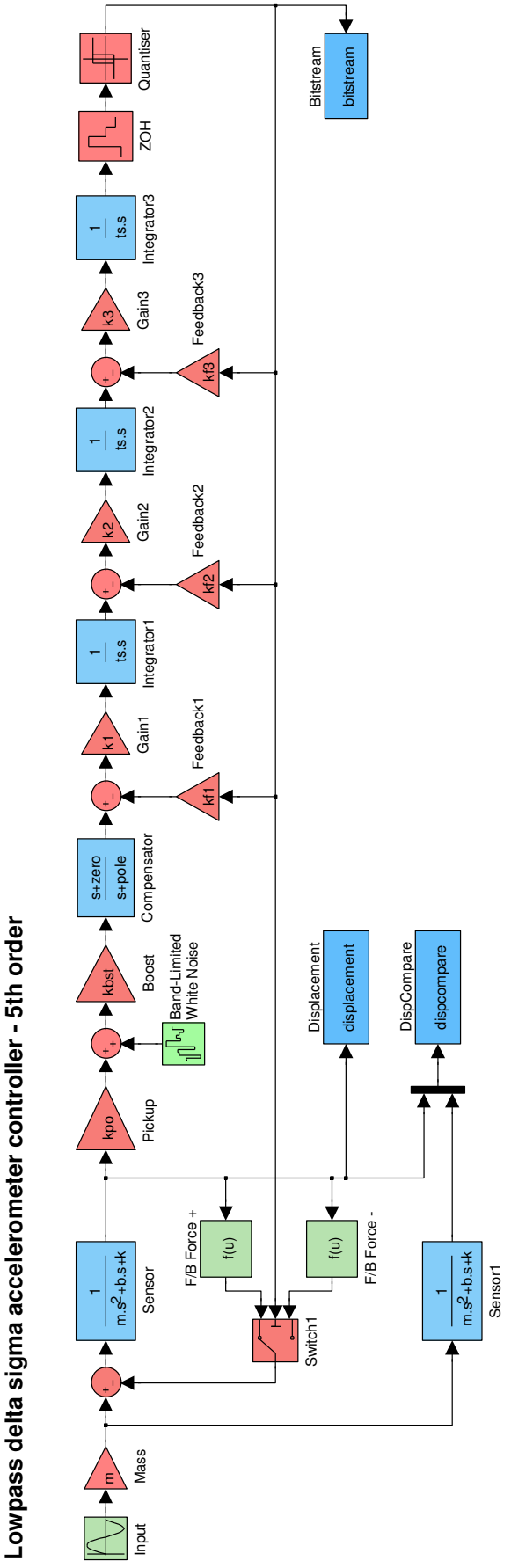


Figure F.2: Simulink block diagram of the EM-ΣΔ [95]

Bibliography

[1] R. Rudolf et al. "Automated critical device identification for configurable analogue transistors". In: *Design, Automation Test in Europe Conference Exhibition (DATE), 2012*. 2012, pp. 858 –861.

[2] R. Rudolf, R. Wilcock, and P.R. Wilson. "Reliability improvement and online calibration of ICs using configurable analogue transistors". In: *Aerospace Conference, 2012 IEEE*. 2012, pp. 1 –8. DOI: 10.1109/AERO.2012.6187250.

[3] Robert Rudolf et al. "Fully differential electro-mechanical phase locked loop sensor circuit". In: *Sensors and Actuators A: Physical* 194.0 (2013), pp. 240 –246. ISSN: 0924-4247. DOI: http://dx.doi.org/10.1016/j.sna.2013.01.053. URL: http://www.sciencedirect.com/science/article/pii/S0924424713000733.

[4] Robert Rudolf Peter Wilson and Reuben Wilcock. "Achieving Invariability in Analog Circuits Operating in Extreme Environments". In: *Extreme Environment Electronics*. Ed. by John D. Cressler and H. Alan Mantooth, pp. 509–517. ISBN: 978-1-4398-7430-1.

[5] Kwyro Lee et al. "The impact of semiconductor technology scaling on CMOS RF and digital circuits for wireless application". In: *Electron Devices, IEEE Transactions on* 52.7 (2005), pp. 1415–1422. ISSN: 0018-9383. DOI: 10.1109/TED.2005.850632.

[6] L.L. Lewyn et al. "Analog Circuit Design in Nanoscale CMOS Technologies". In: *Proceedings of the IEEE* 97.10 (2009), pp. 1687–1714. ISSN: 0018-9219. DOI: 10.1109/JPROC.2009.2024663.

[7] Colin. *Hewlett-Packard HP200A audio oscillator*. June 2010. URL: https://www.flickr.com/photos/woz2/4894370089/ (visited on 09/03/2014).

[8] William R. Hewlett. "Variable Frequency Oscillation Generator". Pat. US2268872. Jan. 1942.

[9] S. Saxena et al. "Variation in Transistor Performance and Leakage in Nanometer-Scale Technologies". In: *Electron Devices, IEEE Transactions on* 55.1 (2008), pp. 131–144. ISSN: 0018-9383. DOI: 10.1109/TED.2007.911351.

[10] K. Bernstein et al. "High-performance CMOS variability in the 65-nm regime and beyond". In: *IBM Journal of Research and Development* 50.4.5 (2006), pp. 433–449. ISSN: 0018-8646. DOI: 10.1147/rd.504.0433.

[11] A. Asenov et al. "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs". In: *Electron Devices, IEEE Transactions on* 50.9 (2003), pp. 1837–1852. ISSN: 0018-9383. DOI: 10.1109/TED.2003.815862.

[12] G. Roy et al. "Simulation Study of Individual and Combined Sources of Intrinsic Parameter Fluctuations in Conventional Nano-MOSFETs". In: *Electron Devices, IEEE Transactions on* 53.12 (2006), pp. 3063–3070. ISSN: 0018-9383. DOI: 10.1109/TED.2006.885683.

[13] Dieter K. Schroder. "Negative bias temperature instability: What do we understand?" In: *Microelectronics Reliability* 47.6 (2007). <ce:title>Modelling the Negative Bias Temperature Instability</ce:title>, pp. 841 –852. ISSN: 0026-2714. DOI: http://dx.doi.org/10.1016/j.microrel.2006.10.006. URL: http://www.sciencedirect.com/science/article/pii/S002627140600374X.

[14] K.A. Bowman et al. "Impact of extrinsic and intrinsic parameter fluctuations on CMOS circuit performance". In: *Solid-State Circuits, IEEE Journal of* 35.8 (2000), pp. 1186–1193. ISSN: 0018-9200. DOI: 10.1109/4.859508.

[15] P.R. Kinget. "Device mismatch and tradeoffs in the design of analog circuits". In: *Solid-State Circuits, IEEE Journal of* 40.6 (2005), pp. 1212–1224. ISSN: 0018-9200. DOI: 10.1109/JSSC.2005.848021.

[16] Jules Lehmann Edwin A. Goldberg. "Stabilized Direct Current Amplifier". Pat. US2684999. July 1954.

[17] G. Debyser and G. Gielen. "Efficient analog circuit synthesis with simultaneous yield and robustness optimization". In: *Computer-Aided Design, 1998. ICCAD 98. Digest of Technical Papers. 1998 IEEE/ACM International Conference on*. 1998, pp. 308–311. DOI: 10.1109/ICCAD.1998.144283.

[18] T. Tobita and H. Takasago. "New trimming technology of a thick film resistor by the pulse voltage method". In: *Components, Hybrids, and Manufacturing Technology, IEEE Transactions on* 14.3 (1991), pp. 613 –617. ISSN: 0148-6411. DOI: 10.1109/33.83952.

[19] D. Comer. "A monolithic 12-bit DAC". In: *Circuits and Systems, IEEE Transactions on* 25.7 (1978), pp. 504 –509. ISSN: 0098-4094. DOI: 10.1109/TCS.1978.1084505.

[20] J.J. Price. "A passive laser-trimming technique to improve the linearity of a 10-bit D/A converter". In: *Solid-State Circuits, IEEE Journal of* 11.6 (1976), pp. 789 –794. ISSN: 0018-9200. DOI: 10.1109/JSSC.1976.1050819.

[21] Precision Monolithics Incorporated. "AN-13 The OP-07 Ultra-Low Offset Voltage Op Amp". In: *Linear and Converters IC Products Databook*. 1976.

[22] S. Rioux et al. "Design methods for CMOS low-current finely tunable voltage references covering a wide output range". In: *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*. 2005, 4257 –4260 Vol. 5. DOI: 10.1109/ISCAS.2005.1465571.

[23] K. Kato, T. Ono, and Y. Amemiya. "A monolithic 14 bit D/A converter fabricated with a new trimming technique (DOT)". In: *Solid-State Circuits, IEEE Journal of* 19.5 (1984), pp. 802 –807. ISSN: 0018-9200. DOI: 10.1109/JSSC.1984.1052224.

[24] J.A. Babcock et al. "Precision electrical trimming of very low TCR poly-SiGe resistors". In: *Electron Device Letters, IEEE* 21.6 (2000), pp. 283 –285. ISSN: 0741-3106. DOI: 10.1109/55.843151.

[25] F. Assaderaghi et al. "A dynamic threshold voltage MOSFET (DTMOS) for very low voltage operation". In: *Electron Device Letters, IEEE* 15.12 (1994), pp. 510 –512. ISSN: 0741-3106. DOI: 10.1109/55.338420.

[26] M. Maymandi-Nejad and M. Sachdev. "DTMOS Technique for Low-Voltage Analog Circuits". In: *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on* 14.10 (2006), pp. 1151 –1156. ISSN: 1063-8210. DOI: 10.1109/TVLSI.2006.884174.

[27] Z Wang et al. "Coping with process variations in ultra-low power CMOS analog integrated circuits". In: *SoutheastCon, 2007. Proceedings. IEEE*. 2007, pp. 54–57. DOI: 10.1109/SECON.2007.342853.

[28] H Suzuki et al. "Post-silicon programmed body-biasing platform suppressing device variability in 45 nm CMOS technology". In: *Low Power Electronics and Design (ISLPED), 2008 ACM/IEEE International Symposium on*. 2008, pp. 15–20. DOI: 10.1145/1393921.1393931.

[29] A Sathanur et al. "Physically clustered forward body biasing for variability compensation in nanometer CMOS design". In: *Design, Automation Test in Europe Conference Exhibition, 2009. DATE '09.* 2009, pp. 154–159.

[30] J. Sweeney and R. Geiger. "Very high precision analog trimming using floating gate MOSFETS". In: *Circuit Theory and Design, 1989., European Conference on*. 1989, pp. 652 –655.

[31] E. Sackinger and W. Guggenbuhl. "An analog trimming circuit based on a floating-gate device". In: *Solid-State Circuits, IEEE Journal of* 23.6 (1988), pp. 1437 –1440. ISSN: 0018-9200. DOI: 10.1109/4.90044.

[32] D. W. Graham et al. "Indirect Programming of Floating-Gate Transistors". In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* 54.5 (2007), pp. 951 –963. ISSN: 1549-8328. DOI: 10.1109/TCSI.2007.895521.

[33] K. Ohsaki, N. Asamoto, and S. Takagaki. "A single poly EEPROM cell structure for use in standard CMOS processes". In: *Solid-State Circuits, IEEE Journal of* 29.3 (1994), pp. 311 –316. ISSN: 0018-9200. DOI: 10.1109/4.278354.

[34] S.A. Jackson, J.C. Killens, and B.J. Blalock. "A programmable current mirror for analog trimming using single poly floating-gate devices in standard CMOS technology". In: *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on* 48.1 (2001), pp. 100 –102. ISSN: 1057-7130. DOI: 10.1109/82.913192.

[35] Yanyi Liu Wong, M H Cohen, and P A Abshire. "A 1.2-GHz Comparator With Adaptable Offset in 0.35- m CMOS". In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* 55.9 (2008), pp. 2584–2594. ISSN: 1549-8328. DOI: 10.1109/TCSI.2008.918010.

[36] Yanyi Liu Wong, M H Cohen, and P A Abshire. "A floating-gate comparator with automatic offset adaptation for 10-bit data conversion". In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* 52.7 (2005), pp. 1316–1326. ISSN: 1549-8328. DOI: 10.1109/TCSI.2005.851389.

[37] Pui-In Mak, Seng-Pan U, and R.P. Martins. "On the Design of a Programmable-Gain Amplifier With Built-In Compact DC-Offset Cancellers for Very Low-Voltage WLAN Systems". In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* 55.2 (2008), pp. 496–509. ISSN: 1549-8328. DOI: 10.1109/TCSI.2007.910643.

[38] J A Lenero-Bardallo, T Serrano-Gotarredona, and B Linares-Barranco. "A Calibration Technique for Very Low Current and Compact Tunable Neuromorphic Cells: Application to 5-bit 20-nA DACs". In: *Circuits and Systems II: Express Briefs, IEEE Transactions on* 55.6 (2008), pp. 522–526. ISSN: 1549-7747. DOI: 10.1109/TCSII.2007.916864.

[39] J Maunu et al. "Current source calibration by combination selection of minimum sized devices". In: *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on.* 2006, 4 pp. –552. DOI: 10.1109/ISCAS.2006.1692644.

[40] G Gielen et al. "Emerging Yield and Reliability Challenges in Nanometer CMOS Technologies". In: *Design, Automation and Test in Europe, 2008. DATE '08.* 2008, pp. 1322–1327. DOI: 10.1109/DATE.2008.4484862.

[41] M Pastre and M Kayal. "Methodology for the digital calibration of analog circuits and systems using sub-binary radix DACs". In: *Mixed Design of Integrated Circuits Systems, 2009. MIXDES '09. MIXDES-16th International Conference.* 2009, pp. 456–461.

[42] Wei Jiang and V D Agrawal. "Designing variation-tolerance in mixed-signal components of a system-on-chip". In: *Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on.* 2009, pp. 2313–2316. DOI: 10.1109/ISCAS.2009.5118262.

[43] Hao-Chiao Hong, Fang-Yi Su, and Shao-Feng Hung. "A Fully Integrated Built-In Self-Test *Sigma−Delta* ADC Based on the Modified Controlled Sine-Wave Fitting Procedure". In: *Instrumentation and Measurement, IEEE Transactions on* 59.9 (2010), pp. 2334 –2344. ISSN: 0018-9456. DOI: 10.1109/TIM.2009.2034582.

[44] Jiun-Lang Huang and Kwang-Ting Cheng. "A sigma-delta modulation based BIST scheme for mixed-signal circuits". In: *Design Automation Conference, 2000. Proceedings of the ASP-DAC 2000. Asia and South Pacific*. 2000, pp. 605 –610. DOI: 10.1109/ASPDAC.2000.835172.

[45] Paul O'Leary. "Practical Aspects of Mixed Analogue and Digital Design". In: *Analogue-digital ASICs: Circuit Techniques, Design Tools and Applications*. Ed. by R.S. Soin, F. Maloberti, and J. Franca. IEE circuits and systems series. P. Peregrinus, 1991, pp. 213–238. ISBN: 9780863412592.

[46] Behzad Razavi. *Design of Analog Cmos Integrated Circuits*. Electrical Engineering Series. McGraw-Hill Higher Education, 2001. ISBN: 9780071188395.

[47] H. Camenzind. *Designing Analog Chips*. Virtualbookworm.com, 2005. ISBN: 9781589397187.

[48] P.E. Allen and D.R. Holberg. *CMOS Analog Circuit Design*. Oxford series in electrical and computer engineering. Oxford University Press, 2002. ISBN: 9780195116441.

[49] Ken Martin BDavid Johns. *Analog integrated circuit design*. McGraw-Hill Higher Education, 1997. ISBN: 9780471144489.

[50] S.K. Ghandhi. "Bias Considerations in Transistor Circuit Design". In: *Circuit Theory, IRE Transactions on* 4.3 (1957), pp. 194–202. ISSN: 0096-2007. DOI: 10.1109/TCT.1957.1086372.

[51] Barrie Gilbert. "A new wide-band amplifier technique". In: *Solid-State Circuits, IEEE Journal of* 3.4 (1968), pp. 353–365. ISSN: 0018-9200. DOI: 10.1109/JSSC.1968.1049924.

[52] C.S. Bill et al. "A temperature- and process-tolerant 64K EEPROM". In: *Solid-State Circuits, IEEE Journal of* 20.5 (1985), pp. 979–985. ISSN: 0018-9200. DOI: 10.1109/JSSC.1985.1052424.

[53] Texas Instruments Corporation. *LM199/LM299/LM399 Precision Reference Data Sheet*. April 2005. URL: http://www.ti.com/lit/gpn/lm199.

[54] Microsemi Corpoation. *1N4565 thru 1N4584A-1 6.4 Volt Temperature Compensated Zener Reference Diodes Data Sheet*. 2003. URL: http://www.microsemi.com/document-portal/doc_download/10936-sa6-21-pdf.

[55] R.J. Widlar. "New developments in IC voltage regulators". In: *Solid-State Circuits, IEEE Journal of* 6.1 (1971), pp. 2–7. ISSN: 0018-9200. DOI: 10.1109/JSSC.1971.1050151.

[56] Texas Instruments Corporation. *REF5010 Low-Noise, Very Low Drift, Precision Voltage Reference Data Sheet*. June 2010. URL: http://www.ti.com/lit/gpn/ref5010.

[57] L.T. Harrison. *Current Sources and Voltage References: A Design Reference for Electronics Engineers*. Elsevier Science, 2005. ISBN: 9780080455556.

[58] G. G E Gielen and R.A Rutenbar. "Computer-aided design of analog and mixed-signal integrated circuits". In: *Proceedings of the IEEE* 88.12 (2000), pp. 1825–1854. ISSN: 0018-9219. DOI: 10.1109/5.899053.

[59] R. Harjani. "OASYS: a framework for analog circuit synthesis". In: *ASIC Seminar and Exhibit, 1989. Proceedings., Second Annual IEEE*. 1989, P13–1/1–4. DOI: 10.1109/ASIC.1989.123242.

[60] F. El-Turky and E.E. Perry. "BLADES: an artificial intelligence approach to analog circuit design". In: *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on* 8.6 (1989), pp. 680–692. ISSN: 0278-0070. DOI: 10.1109/43.31523.

[61] G. James. *Modern Engineering Mathematics*. Pearson Prentice Hall, 2008. ISBN: 9780132391443.

[62] Sawal Hamid Md Ali. "System level performance and yield optimisation for analogue integrated circuits". 2009. URL: http://eprints.soton.ac.uk/69724/.

[63] K. Deb et al. "A fast and elitist multiobjective genetic algorithm: NSGA-II". In: *Evolutionary Computation, IEEE Transactions on* 6.2 (2002), pp. 182–197. ISSN: 1089-778X. DOI: 10.1109/4235.996017.

[64] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. "Optimization by simulated annealing". In: *SCIENCE* 220.4598 (1983), pp. 671–680.

[65] Colin Warwick. "In a Nutshell: How SPICE Works". In: *IEEE EMC Society Newsletter, Summer 2009 – Issue No. 222* 222 (2009), pp. 59–61.

[66] Michael Krasnicki et al. "MAELSTROM: Efficient Simulation-based Synthesis for Custom Analog Cells". In: *Proceedings of the 36th Annual ACM/IEEE Design Automation Conference*. DAC '99. New Orleans, Louisiana, USA: ACM, 1999, pp. 945–950. ISBN: 1-58113-109-7. DOI: 10.1145/309847.310104. URL: http://doi.acm.org/10.1145/309847.310104.

[67] G. Gielen, H. Walscharts, and Willy Sansen. "Analog Circuit Design Optimization based on Symbolic Simulation and Simulated Annealing". In: *Solid-State Circuits Conference, 1989. ESSCIRC '89. Proceedings of the 15th European*. 1989, pp. 252–255. DOI: 10.1109/ESSCIRC.1989.5468066.

[68] J. Bastos et al. "A 12-bit intrinsic accuracy high-speed CMOS DAC". In: *Solid-State Circuits, IEEE Journal of* 33.12 (1998), pp. 1959 –1969. ISSN: 0018-9200. DOI: 10.1109/4.735536.

[69] G.A.M. Van Der Plas et al. "A 14-bit intrinsic accuracy Q2 random walk CMOS DAC". In: *Solid-State Circuits, IEEE Journal of* 34.12 (1999), pp. 1708 –1718. ISSN: 0018-9200. DOI: 10.1109/4.808896.

[70] Yonghua Cong and R.L. Geiger. "A 1.5-V 14-bit 100-MS/s self-calibrated DAC". In: *Solid-State Circuits, IEEE Journal of* 38.12 (2003), pp. 2051 –2060. ISSN: 0018-9200. DOI: 10.1109/JSSC.2003.819163.

[71] Yuan-Lang Ma and Jiun-Lang Huang. "A self-testing and calibration technique for current-steering DACs". In: *VLSI Design, Automation and Test, 2008. VLSI-DAT 2008. IEEE International Symposium on*. 2008, pp. 295 –298. DOI: 10.1109/VDAT.2008.4542471.

[72] E. Bechthum, G. Radulov, and A. van Roermund. "A novel temperature and disturbance insensitive DAC calibration method". In: *Circuits and Systems (ISCAS), 2011 IEEE International Symposium on*. 2011, pp. 2003 –2006. DOI: 10.1109/ISCAS.2011.5937988.

[73] H. Hristov. "A low-complexity calibration algorithm for current steering DACs". In: *Electronics Technology: Meeting the Challenges of Electronics Technology Progress, 2004. 27th International Spring Seminar on*. Vol. 1. 2004, 162 –166 vol.1. DOI: 10.1109/ISSE.2004.1490398.

[74] K. Virtanen et al. "A 12-bit Current-Steering DAC with Calibration by Combination Selection". In: *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*. 2007, pp. 1469 –1472. DOI: 10.1109/ISCAS.2007.378580.

[75] Jen-Hung Chi, Shih-Hsuan Chu, and Tsung-Heng Tsai. "A 1.8-V 12-bit 250-MS/s 25-mW self-calibrated DAC". In: *ESSCIRC, 2010 Proceedings of the*. 2010, pp. 222 –225. DOI: 10.1109/ESSCIRC.2010.5619889.

[76] J. Hyde et al. "A floating-gate trimmed, 14-bit, 250 Ms/s digital-to-analog converter in standard 0.25 mu;m CMOS". In: *VLSI Circuits Digest of Technical Papers, 2002. Symposium on*. 2002, pp. 328 –331. DOI: 10.1109/VLSIC.2002.1015118.

[77] J. Hyde et al. "A 300-MS/s 14-bit digital-to-analog converter in logic CMOS". In: *Solid-State Circuits, IEEE Journal of* 38.5 (2003), pp. 734 –740. ISSN: 0018-9200. DOI: 10.1109/JSSC.2003.810049.

[78] D. Marche, Y. Savaria, and Y. Gagnon. "Laser Fine-Tuneable Deep-Submicrometer CMOS 14-bit DAC". In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* 55.8 (2008), pp. 2157 –2165. ISSN: 1549-8328. DOI: 10.1109/TCSI.2008.920152.

[79] Tao Chen and G G E Gielen. "A 14-bit 200-MHz Current-Steering DAC With Switching-Sequence Post-Adjustment Calibration". In: *Solid-State Circuits, IEEE Journal of* 42.11 (2007), pp. 2386–2394. ISSN: 0018-9200. DOI: 10.1109/JSSC.2007.906200.

[80] P R Wilson and R Wilcock. "Yield improvement using configurable analogue transistors". In: *Electronics Letters* 44.19 (2008), pp. 1132–1134. ISSN: 0013-5194. DOI: 10.1049/el:20081409.

[81] P Wilson and R Wilcock. "Optimal sizing of configurable devices to reduce variability in integrated circuits". In: *Design, Automation Test in Europe Conference Exhibition, 2009. DATE '09.* 2009, pp. 1385–1390.

[82] D. Marquardt. "An Algorithm for Least-Squares Estimation of Nonlinear Parameters". In: *Journal of the Society for Industrial and Applied Mathematics* 11.2 (1963), pp. 431–441. DOI: 10.1137/0111030. eprint: http://dx.doi.org/10.1137/0111030. URL: http://dx.doi.org/10.1137/0111030.

[83] Chi-Hung Lin and K. Bult. "A 10-b, 500-MSample/s CMOS DAC in 0.6 mm2". In: *Solid-State Circuits, IEEE Journal of* 33.12 (1998), pp. 1948–1958. ISSN: 0018-9200. DOI: 10.1109/4.735535.

[84] H. S. Black. "Stabilized Feed-Back Amplifiers". In: *American Institute of Electrical Engineers, Transactions of the* 53.1 (1934), pp. 114–120. ISSN: 0096-3860. DOI: 10.1109/T-AIEE.1934.5056493.

[85] M. Lemkin and B.E. Boser. "A three-axis micromachined accelerometer with a CMOS position-sense interface and digital offset-trim electronics". In: *Solid-State Circuits, IEEE Journal of* 34.4 (1999), pp. 456 –468. ISSN: 0018-9200. DOI: 10.1109/4.753678.

[86] W.R. Bennett. "Spectra of quantized signals". In: *Bell System Technical Journal, The* 27.3 (1948), pp. 446–472. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1948.tb01340.x.

[87] J.M. de la Rosa. "Sigma-Delta Modulators: Tutorial Overview, Design Guide, and State-of-the-Art Survey". In: *Circuits and Systems I: Regular Papers, IEEE Transactions on* 58.1 (2011), pp. 1–21. ISSN: 1549-8328. DOI: 10.1109/TCSI.2010.2097652.

[88] B.P. Brandt. *Oversampled Analog-to-digital Conversion*. Stanford University, 1991.

[89] J. Lota, M. Al-Janabi, and I. Kale. "Nonlinear-Stability Analysis of Higher Order Sigma Delta Modulators for DC and Sinusoidal Inputs". In: *Instrumentation and Measurement, IEEE Transactions on* 57.3 (2008), pp. 530–542. ISSN: 0018-9456. DOI: 10.1109/TIM.2007.911640.

[90] M. Kraft X. Jiang J. Seeger and B.E. Boser. "A monolithic surface micromachined z-axis gyroscope with digital output". In: *Symposium on VLSI Circuits*. Hawaii, USA, 2000, pp. 16–19.

[91] C.P. Lewis M. Kraft and T.G. Hesketh. "Closed loop silicon accelerometers". In: *IEE Proceedings Circuits, Devices and Systems* (1998).

[92] J.C. Candy and O. Benjamin. "The Structure of Quantization Noise from Sigma-Delta Modulation". In: *Communications, IEEE Transactions on* 29.9 (1981), pp. 1316–1323. ISSN: 0090-6778. DOI: 10.1109/TCOM.1981.1095151.

[93] Y Dong, M Kraft, and C O Gollasch. "A high performance accelerometer with fifth order sigma delta modulator". In: *J Micromech Microeng* 15 (2005), pp. 1–8.

[94] Jian Luo, Haitao Ding, and M. Kraft. "A new design methodology for electro-mechanical Sigma- Delta-Modulators". In: *Nano/Micro Engineered and Molecular Systems, 2009. NEMS 2009. 4th IEEE International Conference on*. 2009, pp. 881 –884.

[95] R. Wilcock and M. Kraft. "Genetic Algorithm for the Design of Electro-Mechanical Sigma Delta Modulator MEMS Sensors". In: *Sensors* (2011).

[96] Y. Matsumoto et al. "Three-axis SOI capacitive accelerometer with PLL C-V converter". In: *Sensors and Actuators A: Physical* 75.1 (1999), pp. 77 –85. ISSN: 0924-4247.

[97] Y. Matsumoto and M. Esashi. "Integrated silicon capacitive accelerometer with PLL servo technique". In: *Sensors and Actuators A: Physical* 39.3 (1993), pp. 209 –217. ISSN: 0924-4247.

[98] Mikko Saukoski et al. "Interface and control electronics for a bulk micromachined capacitive gyroscope". In: *Sensors and Actuators A: Physical* 147.1 (2008), pp. 183 –193. ISSN: 0924-4247.

[99] Ran Fang et al. "A low-noise interface circuit for MEMS vibratory gyroscope". In: *Solid-State and Integrated Circuit Technology (ICSICT), 2010 10th IEEE International Conference on*. 2010, pp. 1456 –1458.

[100] C. W. Nelson A. Hati and D. A. Howe. "Phase-Lock Loops in Vibration Environments". In: *41st Annual Precise Time and Time Interval (PTTI) Meeting*. 2009.

[101] S. Yoneoka et al. "Active Electrostatic Compensation of Micromechanical Resonators Under Random Vibrations". In: *Microelectromechanical Systems, Journal of* 19.5 (2010), pp. 1270 –1272. ISSN: 1057-7157. DOI: 10.1109/JMEMS.2010.2067444.

[102] Z. Ye et al. "High-Performance Closed-Loop Interface Circuit for High-Q Capacitive Microaccelerometers". In: *Sensors Journal, IEEE* 13.5 (2013), pp. 1425–1433. ISSN: 1530-437X. DOI: 10.1109/JSEN.2012.2236312.

[103] R.W. Rhea. *Discrete Oscillator Design: Linear, Nonlinear, Transient, and Noise Domains*. Artech House microwave library. Artech House, Incorporated, 2010. ISBN: 9781608070480. URL: http://books.google.co.uk/books?id=4Op56QdHFPUC.

[104] Texas Instruments Corporation. *LM231A/LM231/LM331A/LM331 Precision Voltage-to-Frequency Converters*. March 2013. URL: http://www.ti.com/lit/gpn/lm331.

[105] L. Ke, R. Wilcock, and P.R. Wilson. "120 nm 1.2 V Low-noise VCO with over seven octave tuning range". In: *Electronics Letters* 47.6 (2011), pp. 368–369. ISSN: 0013-5194. DOI: 10.1049/el.2010.3726.

[106] Yufeng Dong, M. Kraft, and W. Redman-White. "Higher Order Noise-Shaping Filters for High-Performance Micromachined Accelerometers". In: *Instrumentation and Measurement, IEEE Transactions on* 56.5 (2007), pp. 1666–1674. ISSN: 0018-9456. DOI: 10.1109/TIM.2007.904477.

[107] B.V. Amini, R. Abdolvand, and F. Ayazi. "A 4.5-mW Closed-Loop $\Delta\Sigma$ Micro-Gravity CMOS SOI Accelerometer". In: *Solid-State Circuits, IEEE Journal of* 41.12 (2006), pp. 2983–2991. ISSN: 0018-9200. DOI: 10.1109/JSSC.2006.884864.