UNIVERSITÀ DI PISA

Scuola di Dottorato in Ingegneria "Leonardo da Vinci"

Corso di Dottorato di Ricerca in
INGEGNERIA DELL'INFORMAZIONE
(SSD: Ing-Inf-01)

Tesi di Dottorato di Ricerca

VARIABILITY-AWARE DESIGN OF CMOS NANOPOWER REFERENCE CIRCUITS

Francesca Cucchi

Anno 2013

UNIVERSITÀ DI PISA

Scuola di Dottorato in Ingegneria "Leonardo da Vinci"

Corso di Dottorato di Ricerca in
INGEGNERIA DELL'INFORMAZIONE
(SSD:Ing-Inf-01)

Tesi di Dottorato di Ricerca

VARIABILITY-AWARE DESIGN OF CMOS NANOPOWER REFERENCE CIRCUITS

Autore:

Francesca Cucchi *Firma*_____

Relatori:

Prof. Giuseppe Iannaccone *Firma*_____

Prof. Paolo Bruschi *Firma*_____

Prof. Stefano Di Pascoli *Firma*_____

Anno 2013

INDEX

SOMMARIO

Il continuo scaling delle tecnologie CMOS pone numerose sfide ai progettisti, relativi alla crescente sensibilità delle grandezze elettriche alle variazioni di processo. Questa problematica riguarda sia la progettazione analogica che quella digitale, e si traduce nel fatto che le prestazioni tipiche dei circuiti non traggono pieno vantaggio dai miglioramenti nominali offerti dalle tecnologie molto scalate. Un'atra importante sfida per i progettisti circuitali è relativa alla progettazione di sistemi per le emergenti applicazioni portatili e impiantabili, come transponder RFID passivi o dispositivi medici impiantabili. Per queste applicazioni sono necessari un consumo di potenza estremamente basso e una ridotta sensibilità delle grandezze di uscita alle variazioni di processo.
Gli approcci proposti per affrontare il problema della variabilità spesso fanno uso sistemi in reazione di complessi e dispendiosi. Per quanto riguarda i circuiti analogici, in alcuni casi vengono utilizzate procedure di trimming ad hoc, che comunque possono essere costose per le applicazioni portatili e impiantabili sopra menzionate. Più recentemente, è stato presentato un metodo basato su una compensazione "interna", la cui efficacia è abbastanza limitata.
Per questi motivi proponiamo la progettazione di generatori di quantità di riferimento molto precise, basati sull'uso di dispositivi che sono disponibili anche in tecnologia CMOS standard e che sono "intrinsecamente" più robuste rispetto alle variazioni di processo. Seguiamo quindi un approccio "variability-aware", ottenendo una bassa sensibilità al processo insieme ad un consumo di potenza veramente basso, con il principale svantaggio di una elevata occupazione di area. Tutti i risultati sono stati ottenuti in una tecnologia UMC 0.18µm CMOS:
In particolare, abbiamo applicato questo approccio al progetto di un riferimento di tensione basato sulla classica architettura "bandgap" con l'uso di bipolari di substrato, ottenendo una deviazione standard relativa della tensione di riferimento dello 0.18% e un consumo di potenza inferiore a 70 nW. Questi risultati sono basati su misure su un set di 20 campioni di un singolo batch. Sono anche disponibili risultati relativi alla variabilità inter batch, che mostrano una deviazione standard relativa cumulativa della tensione di riferimento dello 0.35%.
Questo approccio è stato applicato anche alla progettazione di un riferimento di corrente, basato sull'uso della classica architettura bandgap a bipolari e di resistori di diffusione, ottenendo anche in questo caso una sensibilità al processo della corrente di riferimento dell'1.4% con un consumo di potenza inferiore a 300 nW. Questi sono risultati sperimentali ottenuti dalle misure su 20 campioni di un singolo batch.
I riferimenti di tensione e di corrente proposti sono stati quindi utilizzati per la progettazione di un oscillatore a rilassamento a bassa frequenza, che unisce una ridotta sensibilità al processo, inferiore al 2%, con un basso consumo di potenza, circa 300 nW, ottenuto sulla base di simulazioni.
Infine, nella progettazione dei blocchi sopra menzionati, abbiamo applicato un metodo per la determinazione e la stabilità dei punti di riposo, basato sull'uso dei CAD standard utilizzati per la progettazione microelettronica. Questo approccio ci ha permesso di determinare la stabilità dei punti di riposo desiderati, e ci ha anche permesso di stabilire che i circuiti di start up spesso non sono necessari.

ABSTRACT

The continuous scaling of CMOS technologies poses several challenges to circuit designers, related to the growing sensitivity of electrical quantities to process variations. This issue involves both digital and analog design, and translates in the fact that circuit performance is typically not able to take full advantage of the nominal improvements offered by aggressively scaled technologies. Another important challenge for circuit designers is related to the design of systems for the emerging portable or implantable applications, such as passive RFID transponders or implantable medical devices. For such applications, extremely low power consumption and low process sensitivity of output quantities are an essential requirement.

Proposed approaches to tackle variability often involve complex and power hungry feedback systems. For what concerns analog circuits, in some cases ad-hoc trimming procedures are envisaged, which however can be very expensive for the above mentioned portable and implantable applications. More recently, a method has been presented based on an "internal" compensation, whose effectiveness is quite limited.

This is why we propose the design of very precise reference quantities generators based on the use of devices that are available also in standard CMOS technology and are "intrinsically" more robust with respect to process variations. We therefore follow a "variability-aware" approach, obtaining a low process sensitivity together with a very low power consumption, with the main drawback of a sizeable increase in chip area. All results are obtained in a UMC 0.18 µm CMOS technology.

In particular, we apply this approach to the design of a reference voltage generator based on a "classical" bandgap architecture with the use of substrate bipolar transistors, obtaining a relative standard deviation of the reference voltage of 0.18% and a very low power consumption smaller than 70 nW. These results are based on measurements performed on a set of 20 samples from a single batch. Results related to inter-batch variations are also available, and show a cumulative relative standard deviation of the reference voltage of 0.35%.

This approach is also applied to the design of a current generator, based on the use of a "classical" bipolar-based bandgap architecture and of diffusion resistors, obtaining also in this case a very low process sensitivity of the reference current of 1.4% with a low power consumption smaller than 300 nW. These are experimental results obtained from measurements on 20 samples from a single batch.

The proposed voltage and current generators are then used for the design of a relaxation low-frequency oscillator, which couples a low process sensitivity smaller than 2% with a low power consumption of about 300 nW, as obtained by circuit simulations.

Finally, for the design of the above mentioned blocks, we apply a method for the determination and stability of circuit operating points, based on the use of the standard CAD tools used for circuit design. This approach allows us to determine the stability of the desired operating points, and it also allows us to establish the start up circuits are often not needed.

1. PROCESS VARIABILITY IN LOW POWER SCALED CMOS TECHNOLOGIES

In the last ten years, MOSFETs have reached deep decananometer dimensions: from 2010 Intel and AMD have started producing commercial chips using a 32 nm process (for example Intel processors Core i3, Core i5 and the dual core mobile Core i7). From 2012 a commercial 22 nm technology is used in the Intel Ivy Bridge family of processors. The use of 10-nm MOSFETs with conventional architecture has also been demonstrated in a research environment [1].
The continuous downscaling poses several problems to integrated circuit design, because the advantages of scaling, especially in terms of speed, area occupation (for digital circuits) and costs, come at the cost of a significant increase in power consumption, especially for complex digital systems like processors, and of very large sensitivity of transistor operation to process variations. Challenges are also posed by new emerging applications for integrated circuits in scaled technologies, like portable applications or medical implantable devices. For these kind of systems a low power consumption is necessary, but generally it is obtained with an increase of sensitivity to process variability, as we will see.
Therefore the challenge of the "portable devices era" is to obtain the desired performance, in terms of performance per watt, notwithstanding the increasing relevance of process variability
For digital systems like processors and memories, the main design trade-off is between speed and power consumption. While it can be difficult to obtain a good nominal trade off between these two requirements, process variability can change the operating point of the devices so that the final circuit does not comply with the specifications..
Process variability is usually taken into account in the design phase by performing corner analysis or Monte Carlo simulations, but the increased process variability can lead to an excessive reduction of the design space and to an unacceptable relaxation of other requirements. The problem stems from the fact that both corner and Monte Carlo analysis assume no correlation between different parameters, which is unrealistic and leads to overestimating the process variation and mismatch effects. Moreover, with technologies scaling there is a growth of the number of corners and a widening of their distribution.
These problems require new design techniques and approaches, beyond the traditional tools, in order to meet the nominal requirements in the presence of a large process variability. In digital systems, due to the high computational capability and resources, stability towards process is usually obtained with complex feedback systems which use run-time configurability and adaptive bias, as we will see. In analog circuits, the challenge posed by lowering of power consumption and increasing of process variability is very difficult to solve especially if we consider portable battery-operated devices with limited digital capabilities, for examples implantable medical devices or short-range low frequency communication systems like passive RFID transponders and wireless sensors networks. For these systems reducing power consumption implies an increased sensitivity to process variability, which, should be solved with simple and cheap solutions, so complex

compensation feedback or the classical post-processing trimming are not feasible. This is the challenging problem addressed in our thesis.
In what follows we will analyze the effect of scaling for digital and analog circuits, highlighting the growing process variability and the need and the effects of power consumption reduction.

For what concerns process variability, it is important to distinguish between inter-die and intra-die variability. The first one includes lot-to-lot, wafer-to-wafer and a portion of the within wafer variations: it equally affects every element on a chip. Intra-die variability, instead, is related to random and systematic components which produce a non-uniformity of electrical characteristics across the chip. Examples of the lot-to-lot and wafer-to-wafer variations include processing temperatures, equipment properties, wafer polishing and wafer placement [2]. The within-wafer variations are important both for inter-die and intra-die variability. For example, the resist thickness across the wafer is random from wafer to wafer, but deterministic within the wafer, while aberrations in the stepper lens are a cause of (systematic) intra-die variability. An example of random intra-die variability cause is the placement of dopant atoms in the device channel region, which varies randomly and independently from device to device. In the literature, a test chip has been proposed [3] capable of characterizing spatial variations in digital circuits: results from measurements data show that low-voltage domains lead to increased variability, that intra-die variations are uncorrelated and that die-to-die variations are strongly correlated but exhibit decreased correlation as the power supply voltage is lowered.

1.1 Variability-aware low power digital circuits in scaled technologies

1.1.1 Scaling of technologies and digital circuits

The continuous scaling of technologies poses several challenges to digital integrated circuit design. Indeed, down to the 90 nm node, scaling was accompanied by performance improvements and reduction of power consumption for a given functionality [1]. Scaling also implied a reduction of the intrinsic gate switching delay and an increased integrated circuit density, with consequent cost reduction for the same circuit complexity. The main drawback was the reduction of power supply voltage, which was usually accompanied by a reduction of the MOSFET threshold voltage that resulted in an overall decrease of switching times. If all voltages and all geometrical parameters were scaled by a factor 1/S, this implied an ideal S^2 density increase, an ideal 1/S decrease of the intrinsic delay and an ideal $1/S^2$ decrease in power consumption at a constant power per unit area [4].
However, from 65 nm and below, some physical and quantum mechanical effects, that were previously negligible, become very important and affect devices and circuits performances. The most important consequences of these effects are [4]:

- larger leakage currents;
- large increase of intra-die variability;

- increased influence of interconnect delay on circuit performance ;
- larger gate leakage currents

For what concerns subthreshold leakage current, the current I_{sub} of a MOSFET with V_{GS}=0 is given by:

$$I_{sub} = I_0 \exp\left(\frac{-V_{th}}{\eta V_T}\right), \tag{1.1}$$

where I_0 is inversely proportional to the MOSFET length L. This shows that a lower threshold voltage and lower sizes implies a larger subthreshold current, also because the threshold voltage reduces for very short transistors lengths.
For what concerns process variability, there are two main effects that increase intra-die variability:

- random dopant distribution [5], [6];
- line edge roughness [7] [8].

Random dopant distribution is related to the intrinsic spread of the number of dopants in the active area [1][9]: the decrease of the channel length and width implies a smaller number of dopants in the space charge regions, and therefore a larger impact of random variations on the total number of dopants. For example the number of dopant atoms in the channel inversion layer for a 0.1 µm feature size MOSFET is on the order of ten, hence, its standard deviation , in the case of a Poissonian process, reaches about 30% [6]. The variation in the number of dopants affects, heavily, MOSFETs threshold voltage [10]. In addition to random discrete dopants in the active region of transistors, it is important to consider also random dopants and grain boundaries in the polysilicon gate [11].

Line edge roughness is caused by tolerances inherent to materials and tools used in the lithography processes. Its importance is related to the fact that, as MOSFET length decreases, the same roughness at the edge of poly lines becomes more relevant, leading to a larger variation in the effective gate length.
In conclusion, digital design in scaled technologies must take into account the increase of power consumption due to the leakage currents and the increase of intra-die process variability.

1.1.2 Low power consumption in digital circuits

Reduction of power consumption is very important for digital systems like processors, because all improvements given by continuously scaled technologies, in terms of speed, devices density and cost, are limited by the continuous increase in power consumption per unit area, which poses problems related to heat removal and cooling and which can prevent circuits from reaching the high performance allowed by scaling.

Power consumption in digital circuits can be divided in three main parts: switching power, short circuit power (these two powers constitute the dynamic circuit power) and static leakage power. The total power consumption can be expresses as:

$$P = \alpha f C_L V_{DD}^{\ 2} + I_{SC} V_{DD} + I_{leak} V_{DD}, \tag{1.2}$$

where α is the switching activity (related to the number of commutations during which energy is absorbed from power supply), f is the switching frequency, C_L is the load capacitance, V_{DD} is the supply voltage, I_{SC} is the short circuit current and I_{leak} is the static leakage current. From this expression we can easily understand that an increase of the operating frequency f, allowed by scaling, also implies an increased switching power. Thus far, the most commonly used approach to reduce power consumption is related to the scaling of power supply voltage [13] because this has a quadratic effect on power consumption reduction. However, a reduction in power supply voltage also implies a reduction of speed, so we can use multiple power supply voltages (for example in memories) in order to obtain a good trade-off between these two requirements.
The MOSFETs threshold voltage reduction is useful in order to sustain performance with power supply voltage reduction because it allows to preserve overdrive voltages and MOSFET current drives; however, threshold voltage reduction implies a larger impact of leakage currents on the total power consumption.
In the literature, several techniques have been proposed in order to face the trade off between dynamic power (which can be reduced without performance reduction by lowering power supply voltage and MOSFETs threshold voltages) and leakage power consumption (which increases with the decrease of threshold voltage): we can use subthreshold devices [14], or multi-threshold devices [15]. The latter method can be applied in multiple-threshold technologies and it is based on the use of larger threshold-voltage MOSFETs in non-critical timing paths in order to reduce power consumption.
We can also use adaptive body bias [16] to control threshold voltages for the subthreshold leakage reduction: indeed reverse body bias (an increase of the pMOS n-well voltage with respect to the supply voltage or a substrate voltage reduction related to ground) is an effective technique for reducing the leakage power of a standby mode design [16]. We will better define the adaptive body bias in the next paragraph. It is important to note, however, that as technologies scale down, the bulk factor becomes smaller [4], and this implies that the adaptive bias is not effective in the reduction of power consumption.
In microprocessors we can control power supply voltage in order to reduce power consumption (switching power depends on the third power of V_{dd}) and we can control the substrate voltage in order to reach a faster speed, for a given power consumption requirement.

1.1.3 Variability-aware low-power digital circuits

Very common requirements for digital circuits such as memories and microprocessors are related to speed and power consumption. Due to process variability, the operating points of the implemented circuits form a "cloud" around

the nominal ones, so some dies cannot achieve the desired speed requirement, while others may fail the maximum leakage power specification [2]. Memories are among the most variability-sensitive components of a system, because most of the transistors in a memory are of minimum size and thus are more prone to variability. Additionally, memories are dominated by parallel paths (wordlines and bitlines) hence timing can be severely degraded by variability due to the dominance of a worst-case path over the rest. Finally, if we consider periodic multimedia applications, memories occupy the majority of the chip area and contribute to the majority of the digital chip energy consumption.

Classical worst case corner analysis is overly pessimistic and adds up to a high total safety margin, which may not ensure compliance with all requirements. For this reason, new techniques [17] have been developed to address the process variability problem with no degradation of circuit performance. They are based on a run-time configurability, because they monitor the configuration and/or system clock, detecting timing violations due to process variability, and they tune some circuit parts at run-time with the use of closed-loop control systems, in order to meet the system-level real-time requirements (in this sense they are “self-adaptive” systems). It is important to note that this approach is effective for the reduction of process variability not only related to technology (layers, atomic variations, lithography) or to electrical parameters (threshold voltage, mobility, leakage), but also to time (temperature drift, degradation, aging), because correction is at run-time.

In Fig. 1.1 an example of process variability effects on SoC is shown. Similar considerations are also true for memories [18], related to the requirements of energy and delay for read/write operations, and for microprocessors. Two operating points are considered: a nominal one, with the desired speed and power requirements, and a faster but more power-hungry one. As already said, due to process variability, the nominal operating point becomes a "cloud", so some dies cannot achieve the desired speed requirement. This cloud can also shift due to temperature or aging degradation. Moreover, if we consider temperature effects, some chips may fail the maximum leakage power specification [2]. Speed and power consumption of each circuit are measured and if the actual operating point does not satisfy speed requirements, a run-time knob is used to reconfigure it towards the faster operating point, by increasing power consumption, in such a way

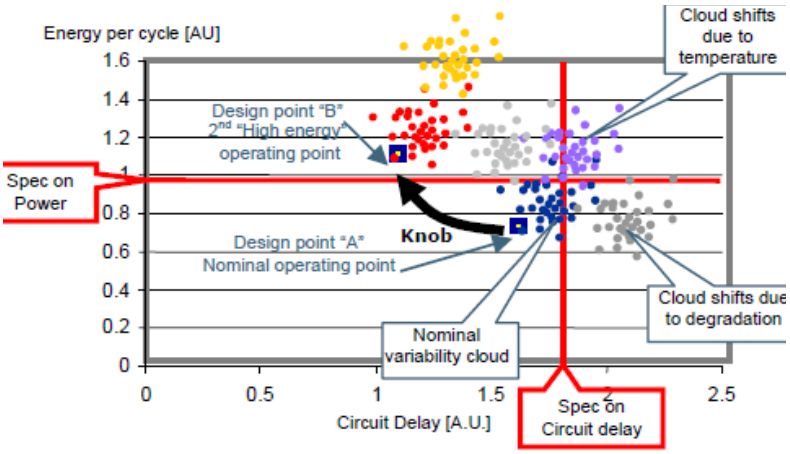

Fig. 1.1: SoC with two design points under variability, reliability and temperature effects, a knob is used to drift the clouds from one design point to another

that also considering the cloud due to variability, the desired speed requirement is met [17], [19]. In this process compensation approach circuits are designed for the nominal requirements of speed and power and they are optimized for the typical case, so no overly pessimistic design is considered: the reconfiguration is applied only for troublesome chips and only for the needed time, reaching gains in power consumption of up to 30% or gains in performance of up to 64% [18], according to targeted specification constraints. Alternatively, we can employ a clock frequency control, based on the use of PLLs, which can adjust the clock period to the delay of the slowest component in the system. This relaxing of the clock period requirement is a solution in order not to increase total power consumption, but has a negative impact on the application execution time, due to the slower clock. We can also combine the two techniques in order to obtain the desired functionality minimizing power consumption [19].

The self-adaptive systems approach is based on the use of process monitors in critical sections of the circuits, on the use of a control algorithm and of actuation knobs, which change circuit operating point obtaining a system reconfiguration so as to regain specifications. This implies the presence of tunable circuit parts.

In what follows we will discuss in more details the typical monitors, knobs and reconfiguration systems for digital circuits. Process monitors are used in order to detect quantities and parameters affected by process variations and critical from this point of view: since the important requirement for digital circuits, is speed, digital delay timing monitors are typical for these circuits are [17] [20], and are used to measure and predict setup time violations or to detect any significant shift in combinatorial logic delay.

Monitors are usually implemented as replicas of the monitored circuit critical paths [16], [21]. Delay chains, composed of multiple cascaded copies of the same cell, are used to amplify the propagation delay of interest. Alternatively, ring oscillators can be used [22][23]: they are closed-loop delay chains used to provide a periodic signal whose period is proportional to the propagation delay.

Delay chains and ring oscillators measure an average propagation delay, so they can be used in order to detect inter-die and inter-wafer variability, while they are

not useful for mismatch measurements. A mismatch detector for MOSFETs threshold voltage has been presented in [24], based on the comparison of the voltage drops on MOSFETs biased with a constant current source. These voltages are converted into a digital signal by means of a VCO followed by digital counters, in order to post process or scan out the threshold voltage measurement.
Some examples of leakage monitors have been proposed [25] [26], [27] [23], which measure on-chip leakage current. This is obtained by properly biasing the MOSFET under test. Indeed the basic principle of leakage sensing is based on the drain-induced barrier lowering (DIBL) [23]: when a high drain voltage is applied to a short-channel device, the barrier height between the drain and the source is lowered, resulting in a decreased threshold voltage and an increased leakage current. Inversely, when a constant bias current flows through a short-channel device, a smaller drain voltage is obtained on a device having a lower threshold voltage. On the basis of these considerations, a MOSFET biased in subthreshold region can be biased with a constant current source, and the leakage can be determined by measuring the drain voltage of the MOSFET itself. This can also be used to measure slew rate [28]. However, leakage-based sensing methods can suffer from on-chip temperature variations, especially in high-performance applications.
Some examples of **knobs** are:

- the circuit power supply voltage (this technique is called “dynamic voltage scaling” (DVS));
- the MOSFET transistor substrate voltage (this technique is called “adaptive body bias” (ABB));
- the power supply voltage together with the substrate bias (it has been shown that this control is more effective than the two previous individual controls for low-power high-performance microprocessors [29]);
- the control voltage of a PLL system which adaptively changes the system clock frequency [19].

The control of the power supply voltage V_{dd} can be efficiently used for the reduction of power consumption, as already said, but a low V_{dd} implies a significant performance degradation. In microprocessors, for example, frequency can be controlled through an adaptive change of the supply voltage, so can be also used to compensate process-induced frequency variations. However, since both switching and leakage components of power consumption have a super-linear dependence on power supply voltage, the change of V_{dd} can have a significant impact on the total power consumption [29].
For what concerns adaptive body bias, which has been presented as a technique for the reduction of power consumption, it is also an important way to control and to reduce process variability effects. The substrate voltage can be controlled in two ways:

- “Reverse body bias” (RBB): we use a pMOS well voltage larger than the power supply voltage or an nMOS substrate voltage smaller that the ground (or the negative power supply) voltage, with a consequent

increase of the MOSFET threshold voltage. It is used to reduce the circuit power consumption even at the cost of a lower speed.

- "Forward body bias" (FBB): we use a pMOS well voltage smaller than the power supply voltage or an nMOS substrate voltage larger that the ground (or the negative power supply) voltage. It is used to increase speed because it reduces the threshold voltage, at the cost of a larger power consumption.

We can also apply a bidirectional adaptive body bias (ABB), which means that we can lower the MOSFET threshold voltage for slow MOSFETS (FBB) and we can increase the threshold voltage for MOSFETs with a large leakage (RBB). In fact, if we consider digital circuits like microprocessors, while reverse body bias can be an effective way to reduce the maximum power consumption, however it does not affect the maximum circuit delay and so its minimum speed. On the other hand, forward body bias is effective in improving the minimum circuit speed [30]. The substrate voltage control poses some problems, for example the clusterization of MOSFETs with the same threshold voltage.
Monitors and knobs are used in closed-loops reconfigurable systems: for example in [30] a speed-adaptive threshold voltage CMOS scheme is presented, in which substrate bias is controlled so that critical paths delay in a circuit remains constant. The simplified view of this scheme, which can be considered an example of typical adaptive systems for digital delay control, is shown in Fig. 1.2.

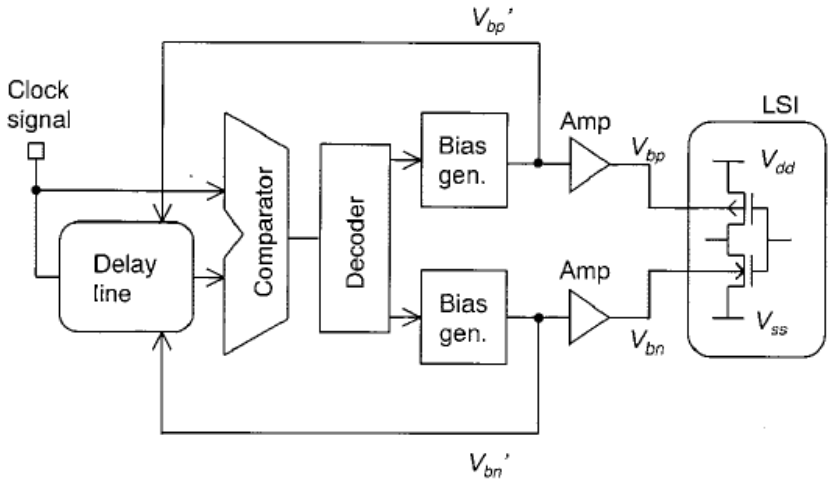

Fig.1.2: adaptive system for digital delay control

The feedback system is composed by a delay line, a comparator, a decoder and the body bias generators, which also control the propagation delay of the delay line. The comparator measures propagation delays between an external clock signal and the output signal from the delay line. The amount of delay is encoded and converted in an analog value which controls the bias generators, in such a way as to maintain a constant delay of the delay line, corresponding to the external

clock frequency at equilibrium. This delay line is usually a replica of the circuit with the most critical path.
The approach based on the use of replica circuits in a feedback loop is useful for the compensation of inter-die and inter batch variability, as already said [31].
However, scaling implies increased intra-die process variability, which limits the effectiveness of a global voltage control. For this reason several techniques [32], [33] have been proposed for an **in-situ characterization** of circuit blocks with small hardware overhead. For example [32] proposes local supply voltage adjustment for systems sub-blocks which are made over-critical: indeed a flip-flop with increased setup time is inserted in parallel to a regular flip flop (see Fig. 1.3), at the end of sub-block critical paths. Outputs of both flip flops are compared: when there is a timing failure for the increased setup time flip-flop (where it occurs first), an error prediction signal (ErrPre in Fig. 1.3) is triggered, which is used by a control logic in order to change the power supply voltage of the sub-block. An in-situ characterization is capable of taking all kinds of process variations, random, within die and die-to-die, as well as environmental and transient variations, into account. There is obviously a trade-off between the advantages in terms of process variability reduction, and the area and power consumption overhead due to the presence of over-critical flip flops.

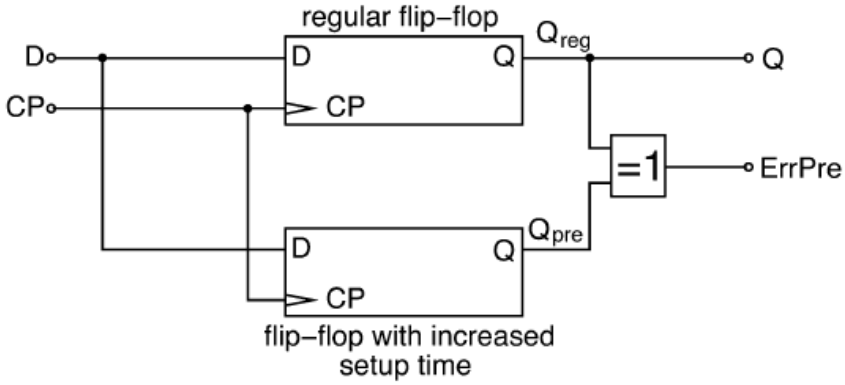

Fig. 1.3: basic structure of a flip flop for critical timing detection

1.2 Variability-aware low power analog circuits in scaled technologies

1.2.1 Scaling of technologies and analog circuits

We already said that scaling is advantageous for digital circuits, even if in sub nm technologies new problems arise which must be addressed by means of new techniques. For analog circuits, advantages and disadvantages of scaling are not so evident: scaled technologies are used in mixed-signal systems, where they offer advantages especially for the digital section. Also for analog circuits, however, an

advantage of scaling is the availability of transistors with higher speed, which can be used in RF circuits and high-speed analog blocks like data converters.
For analog circuits, scaling does not imply an area occupation reduction, because the area of transistors is chosen for noise, linearity or mismatch constraints. A decade-old work presents a detailed analysis of scaling effects for analog circuits [34]: it shows that MOSFET transistors improve with technologies scaling, and this implies also a low power consumption. On the other hand, the supply voltage decreases, and this implies a significant increase in power consumption in order to obtain the same signal to noise and distortion ratio over a certain signal bandwidth. The overall effect is that power consumption decreases with newer CMOS processes down to about 0.25 μm. In sub nm CMOS, either circuit performance decreases or power consumption increases.

1.2.2 Low power consumption in analog circuits

We already said that in digital circuits power consumption is related to the power supply voltage, so we can obtain a low power consumption by properly reducing line voltage. For analog circuits the reduction of power consumption is obtained with different techniques because it is not related to power supply voltage, but to SNR (Signal to Noise Ratio) requirements. In the following, we will derive an expression useful to underline the relation between power consumption and SNR of a circuit. We will consider linear circuits, so the main source of distortion is due to either voltage or current limitations, which produce slewing and/or clipping of the output signal. In a class A system, we start considering the expression for the minimum bias current I_{bias} in order to prevent slewing, which must be equal to the maximum signal current required to drive the output capacitance *C*. It can be expressed as [34]:

$$I_{bias} = 2\pi fCV\,, \tag{1.3}$$

where *f* is the signal frequency and *V* is its magnitude. The thermal noise related to the output node can be expressed as

$$v = \sqrt{\frac{KT}{C}}\,, \tag{1.4}$$

where *K* is the Boltzmann constant and *v* is the root mean square noise associated with the node voltage. So, if we consider relations (3) and (4), the required load capacitance for a specified SNR is:

$$C = SNR \cdot 2kT\,/V^2\,. \tag{1.5}$$

If the supply voltage is equal to the peak-to-peak value of the signal, the minimum power consumption *P* for class A systems is:

$$P = 2VI_{bias} = 8\pi kT \cdot SNR \cdot f\,. \tag{1.6}$$

For class B systems, the absolute minimum power consumption is a factor π lower than the limit given by (6) for class A systems. From expression (6) we can note that the minimum power consumption for an analog circuit is only related to the signal frequency and to the desired SNR, while it is not dependent on power supply voltage, if the input signal is rail-to rail. However, if the peak-to-peak value of the input signal is smaller than the line voltage, and if we call ΔV the part of the supply voltage not used for signal swing, we can express the minimum power consumption P as:

$$P = 2VI_{bias} = 8\pi kT \cdot SNR \cdot f \cdot \frac{V_{dd}}{V_{dd} - \Delta V} \ . \tag{1.7}$$

From (1.7) we can note that in this case the minimum power consumption is larger than in the case of a rail-to-rail input signal. In every case, the minimum power consumption is directly proportional to the desired SNR, so analog circuits become less power efficient when the desired SNR increases, independently of technology and of the power supply voltage.
The power consumption indicated by (1.6) and (1.7) is a theoretical value, which is increased considering:
- the power consumption of the bias circuit, which also adds to the output noise;
- the presence of additional noise sources, related to internal circuit components or to the power supply voltage, which imply an increase in power consumption in order to maintain the same SNR;
- the use of large devices (in order to reduce their mismatch), which adds to the total parasitic capacitance.

We already said that for systems that are not rail-to-rail there is a slight dependence of power consumption on the power supply voltage. Also the presence of a power consumption related to the bias circuit implies that a line voltage reduction can be helpful for in a power consumption reduction. These are the reasons why also in analog circuits the trend is toward a power supply voltage reduction. This reduction, together with the power consumption reduction, can be successfully obtained with the use of MOSFETs operated in subthreshold or weak inversion.
In subthreshold (or weak inversion) the relation between MOSFET drain current I_D and its gate-to-source voltage V_{GS} is exponential and it is expressed as:

$$I_D = \beta V_T^{\ 2} \exp\left(\frac{V_{GS} - V_{th}}{\eta V_T}\right)\left(1 - \exp\left(\frac{-V_{DS}}{V_T}\right)\right), \tag{1.8}$$

where V_T is the thermal voltage, V_{th} is the MOSFET threshold voltage, η is the subthreshold slope and β is the product between carrier mobility μ, gate oxide capacitance for unit area C_{OX} and MOSFET width W divided by MOSFET length L.
In weak inversion the minimum drain-to-source voltage in order to ensure saturation is lower than in strong inversion, because it is equal to the a small multiple of the thermal voltage. In addition, in weak inversion the

transconductance-to-current ratio of a transistor reaches its maximum value. This implies a maximum gain-bandwidth product for a given load capacitance when the current is limited or a minimum input-equivalent noise for a given output noise. In a differential pair, it implies the maximum gain per device and the minimum input-referred offset. However, the maximum value of the transconductance-to-current ratio of a MOSFET implies also a maximum mismatch of current mirrors, which is directly proportional to the MOSFET transconductance, for a given current, and a higher noise, because the drain current noise spectral density is proportional to the drain-source conductance, and proportional to the MOSFET transconductance. Finally, the transition frequency in weak inversion is lower than a few hundred MHz also for 0.18 μm CMOS, so it can not be used in high frequency analog applications.

1.2.3 Low power applications for analog circuits

Low power requirements are extremely demanding for emerging classes of applications, such as short-range low frequency communications systems, like RFID or wireless sensors networks, and biomedical implantable devices. These applications require small size and small weight, and require battery-operated systems with extremely low power consumption to minimize battery replacing or recharging, which can be very expensive or not practical, for example in the case of medical devices

1.2.3.1 Wireless sensor networks

A wireless sensor network consists [35] [36] of one or more base stations and a number of autonomous sensor nodes, from ten to thousands. These nodes are distributed in a physical space in order to monitor physical or environmental parameters. A wide variety of mechanical, thermal, biological, chemical, optical, acoustic and magnetic sensors may be attached to the sensor node. Each node is composed by one or more low-power sensing device, a limited memory and an embedded processor which is interfaced with the sensor, a power module and a radio transceiver, so it has capabilities of sensing, data processing and wireless communication. Wireless communication is used to minimize the infrastructure and to access otherwise inconvenient locations. The measured physical data are processed and transmitted through the network to the base station with the use of a wireless communication channel, so each node can share the collected and processed information through the series of wireless links between nodes, and the end-user can extract the collection of data gathered by several nodes. Depending on the application, actuators may be incorporated in the sensors [37].

One of the first applications for wireless sensor networks was in military environment, for example for the military target tracking and surveillance, because they can help in intrusion detection and identification [38].They can also monitor friendly forces, equipment and ammunition, they can be used for battlefield surveillance and nuclear, biological or chemical attack detection and reconnaissance [37]. Sensor nodes can also be used in environmental applications, for example for chemical or biological detection, for agriculture, or for the forest fire detection. Some applications in the context of seismic sensing have been proposed [39]. They are very useful in industrial application and especially in logistics, asset tracking and supply chain management [40][41].

With technology advances, sensor nodes can also be used for domotic applications, allowing the development of a smart environment self-adapting to the needs of the end user [42]. The concept of "ambient intelligence" [43], which is an environment sensitive and adaptive to people needs, can be applied in the contexts of homes, cars, and offices.
An important and emerging application of WSN is also biomedical and health monitoring [44]. In this field, they are also called body sensor networks (BSN) [45], and they are used for a continuous and non-invasive monitoring of health body parameters and physiological signals, such as temperature, heart rate, electrocardiogram (EKG/ECG), physical activity and respiration. Some examples of SoC platforms for the ubiquitous medical monitoring have been proposed, with very low power consumption [46] [47].
The data rate requirements of WSN are quite low because, especially if we consider environmental or health monitoring, changes are slow and so data rate is low.
A sensor node, as already said, is made up of four basic components [37]: a sensing unit, a processing unit, a transceiver unit and a power unit. The first unit consists of sensors and an analog front end interface with the internal blocks. Usually the first stage consists of an LNA with proper gain and high SNR, and analog-to-digital converters. The low output rates of sensors allow ultra low power amplifier design with devices operating in weak inversion. The processing unit, which is generally associated with a small storage unit, executes data processing and manages the communication protocol. The transceiver unit connects the node to the network. For what concerns communication, even the most energy efficient transceiver standards in commerce, for example those associated with the IEEE 802.15.4 and IEEE 802.11 physical layers, have power consumption of the order of tens of mW.
Wireless sensors networks nodes must have low sizes, in such a way to be embedded in the environment. They must also be low-cost to enable the realization of sensor networks with a large number of nodes, and this implies that the single node, the communication protocol and the network design must have low complexity to satisfy the low cost requirement. Finally they must have a very low power consumption, of the order of 100 μW, because nodes, which are typically battery operated, are in large number and they can be inserted in difficult to access location, so they must be energetically autonomous, because the battery replacing or recharging can be very difficult. Low power consumption must be obtained at each level of the system design, from the physical layer to the communication protocol (for example the operating range of each node is limited to a few meters and the data rate is limited to a few kbps).

1.2.3.2 Implantable medical devices

Implantable medical devices [48] monitor and treat physiological conditions within the body. Some example are hearing aids [49], pacemakers, implantable cardiac defibrillators, glucose meters, drug delivery systems and neurostimulators, which can help the treatment of hearing loss, cardiac arrhythmia, diabetes, Parkinson's disease.

The sensors of body area networks are also an example of implantable medical devices [50] for medical diagnostics. They are used in order to continuously measure internal health status and physiological signals.
These applications require wireless devices with small size and weight, and with low power consumption. Security and safety are also very important requirements. In fact, if we consider BSN, it is important for the communication to be reliable, secure and energy-efficient, and the network must be protected against injection or modification of measurements transmitted to the external devices.

1.2.3.3 RFID

Radio Frequency Identification (RFID) is a contactless technology for automatic identification and tracking which uses radio frequency communication. An RFID system is composed by a reader and a tag (or transponder) located on the object to be identified. RFIDs are relatively small and cheap and they are widely used [51] in asset tracking, real time supply chain management [52] and telemetry-based remote monitoring. They are also used for access control to buildings, parking [53], public transportation and open-air events, airport baggage, animal identification, express parcel logistic. The need for high volume, low cost, small size and large data rate is increasing, while stringent regulation of transmit power and bandwidth have to be met.
RFID systems are closely related to smart cards, because data (the identification information) is stored on an electronic device (the transponder), but their advantages with respect to smart cards are the fact that power supply to the transponder and data exchange with the reader is contactless: this circumvents all the disadvantages related to faulty contacting, so sabotage, dirt, unidirectional insertion, time consuming insertion, etc...
The RFID reader typically contains a radio frequency transceiver, a control unit and a coupling element to the transponder, while the latter is generally composed of a coupling element and an electronic microchip. An important classification of transponders is based on the type of power supply:
- active transponders have an on board battery which provides the power supply voltage;
- semi-passive transponders have an on board battery which however only supplies the logic and the memory management unit, while the reader radiated field is used for transmission, with the modulated backscattered radiation technique;
- passive transponders receive all power needed from field radiated from the reader. A fraction of such power is used by the transponder to communicate with the reader by modulating the backscattered radiation.

Both active and passive transponders benefit from a low power supply voltage: for the first ones, this implies longer battery lifetimes or smaller batteries sizes; for the second ones, this implies a longer communication distance (which can be up to a few meters).
Passive transponders with electromagnetic coupling operate in the UHF (868 MHz in Europe and 915 MHz in USA) or microwave range (2.45 GHz or 5.8 GHz). UHF transponders commercially available need to have about 150 μW input power, and reach a reading distance of 2-8 m, depending on the antenna and operating frequency. If we consider active transponders, the operating range can reach 15 m. The simplified view of a passive RFID transponder is shown in Fig. 1.4 [54].

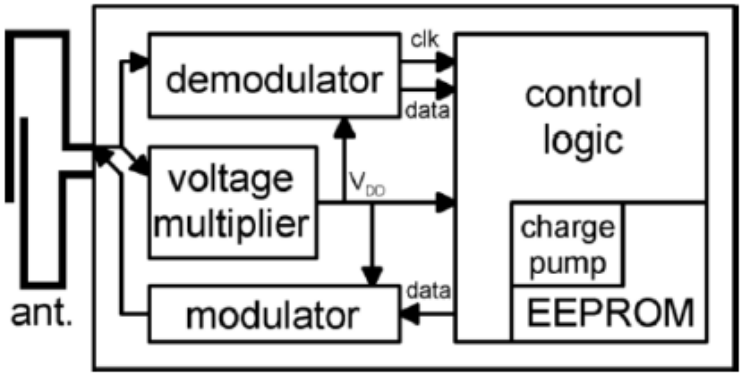

Fig. 1.4: simplified view of a passive RFID transponder

The system is composed by an external antenna (for example a printed loop antenna), which is power matched with the average input impedance of the voltage multiplier. The voltage multiplier is used in order to convert a part of the incoming RF signal into a dc power supply voltage for the internal transceiver blocks. The demodulator converts the pulse-width modulated input signal to digital data and generates a synchronous system clock. The transmission from transponder to reader (uplink) is based on modulating the backscatter of the continuous wave carrier transmitted by the base station. The changes in the IC input impedance are related to data from the control logic. The logic circuit handles the protocol, including anti-collision features, cycling redundancy checksums, error handling, enabling and disabling of analog circuits. A charge pump converts the dc supply voltage into the higher voltage needed for programming the EEPROM; it works at a frequency of approximately 300 kHz generated by an on-chip RC oscillator. The EEPROM contains the transceiver information, for example the identifier.
Further improvements in RFID systems in order to increase the operating range and/or the battery lifetime are related to a proper choice of the modulation technique or to the use of extremely low voltage and low power circuits.

1.2.4 Variability-aware low power analog circuits

The problem of process variability is relevant and challenging also for analog circuits. For example, [55] analyzes the distribution of the MOSFET drain current considering three different batches for the same 90 nm CMOS technology, showing a broad distribution of the current considering the same batch (mainly due to intra die variability), and a very different mean value from batch to batch.
This problem becomes more relevant if we consider subthreshold circuits, given the exponential relation between voltage and current.
Different techniques can be used in order to reduce process variability in analog circuits:

- trimming;
- digital calibration (we speak about digitally enhanced analog design);

- analog compensation circuits.

Trimming is a very common technique used in order to obtain high-precision output quantities, for example in reference voltage generators. It is a post-processing technique, which is based on the measurement of the error between the actual output quantity and the desired value, and on subsequent error correction through the tuning of a dedicated correction circuit. For example, if the output quantity is a current, the current mirror can be composed of additional branches which can be properly activated or deactivated in order to adjust the output quantity to minimize the error. The activation can be done with the help of a digital section including programmable fuses or memories. Another alternative is the use of laser [56] to adjust resistances. Another post-processing technique [57] in order to reduce the effects of process variations is based on the use of configurable transistors, which are composed by a certain number of parallel MOSFETs with switch-connected gates. The configuration of these switches, achieved through digital signals, provides a mechanism to adjust the overall device size. Configurability is used for transistors which are the most sensitive to process variations. In present day industrial design, however, the tendency is to avoid trimming, except for very precise blocks, such as reference generators, since post-processing is usually quite expensive.
Similar to digital trimming is the **digitally enhanced analog design** [58], whose aim is to leverage digital correction and calibration techniques to improve analog performance. The difference with respect to trimming is that in this case the calibration is at run time and is implemented with a digital section that controls the analog quantities and acts on the analog circuit in order to improve precision and stability towards process. It is a technique similar to the one applied in digital adaptive systems, previously described. Indeed digital calibration is obtained [59] with systems composed by an error-revealing node, an error configuration, and a compensation node. In the error-revealing node a signal is measured that is only function of the error we have to compensate. The error configuration is the condition with which we reveal the error and the compensation node is the node in which we insert a compensation signal, which reduced the error. For example, MOSFET bias currents can be digitally controlled and programmed [60] [61] using current mode ladder structures in order to compensate for process variations and components mismatch. The digital signal processing is used also to overcome shortcomings of analog design [62], for example in order to obtain digital calibration of A/D converters [63], [64], [65]. Digital enhancement techniques can be successfully applied also to power amplifiers [66], because they allow an improvement in linearity with a gain in power efficiency of the system. For example, in [67] it was shown that replacing opamps with open-loop gain stages and nonlinear digital correction can reduce amplifier power dissipation by 3-4 times.
Digitally enhanced analog design, however, is useful if a digital section is present, while it can be a very expensive solution for applications such as passive RFID transponders, sensors networks or biomedical devices, where the digital section is not present or it is very simple. For these applications, an analog compensation technique is more convenient.
In the recent literature [55], an alternative technique to trimming or to digitally enhanced circuits has been proposed. It only uses analog circuits with a proper compensation technique, based on the use of reverse correlated quantities. For

example in [55] the design has been proposed of current sources whose reference current is obtained as the sum of two reverse-correlated currents. The principle of operation is illustrated by the circuit in Fig. 1.5.

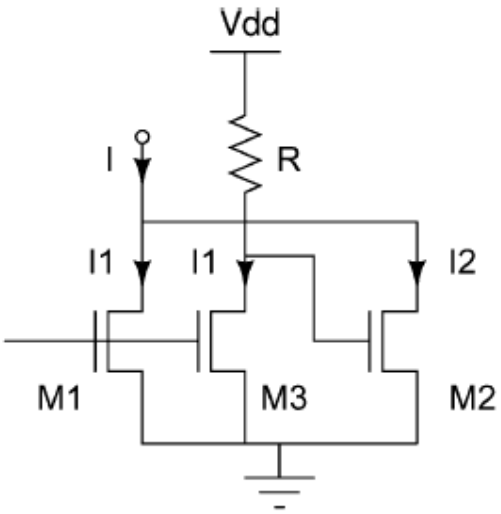

Fig. 1.5: addition-based current generator

The reference current is the sum of the saturation currents of MOSFETs M1 and M2 (I1 and I2, respectively). If current I1 increases, also the voltage drop across R increases, reducing the gate-to-source voltage of M2, and therefore reducing current I2. In this way the sum current I remains almost constant.
In order to obtain a process-invariant reference current, we must have:

$$\Delta V_{gs2} = -\frac{2\Delta I_1}{gm_2} \tag{9}$$

which requires that resistance *R* is equal to $2/gm_2$. With this approach, authors obtain with experiments a reduction of the relative standard deviation of the reference current from 11% to 6.5%, but still high in absolute value. Better results have been obtained with the same approach applied to the design of ring oscillators [68], where authors obtain a reduction of the oscillation frequency relative standard deviation of more than 65%, even if it is still of about 6%.
Another interesting example of process compensation analog loop is presented in [69], where a MOSFET threshold monitor is used in an adaptive biasing circuit in order to obtain a process-dependent control voltage for a process and temperature compensation of a clock oscillator. With this approach authors obtained an oscillation frequency very stable towards process, with a worst-case variation, considering process and temperature, of 2.64%. However the compensation mechanism is quite complex and it can add too much to total power consumption,

especially if we consider the previously mentioned RFID transponders and implantable devices.

1.3 Organization of the thesis

The problem of process variability in analog circuits becomes more and more important with the continuous scaling of technologies. Proposed solutions to mitigate it involve the use of digitally enhanced circuits or trimming, which however may be very expensive, considering systems with a limited or almost absent digital section, such as passive RFID transponders and implantable devices. For such kind of applications, a completely analog solution is preferable. The use of complex feedback systems is not effective, since they imply increased power consumption, while low power is an important requirement for these applications. The use of simple circuits which implement a compensation technique, such as the ones proposed by [55][68], may not be so effective, if very precise quantities are necessary.
This is why we propose an alternative approach, based on the use of devices which are intrinsically more stable towards process and which are available also in standard CMOS technologies, such as lateral bipolar pnp transistors and diffusion resistors. We apply this approach to the design of reference quantities (voltage, current and frequency) generators. With a proper design, we will show that we are able to obtain both extremely low sensitivities to process and a very low power consumption, with the main drawback of a large area occupation.
In Chapter 2 we propose the design of a reference voltage generator with extremely low process sensitivity and very low power consumption. The voltage generator uses substrate bipolar pnp transistors in a "classical" bandgap architecture, that as we will demonstrate is the best solution to obtain stability towards process. With a proper design, we meet this requirement together with the lowest power consumption when compared with bandgap generators proposed in literature. The main drawback is a large increase in area occupation. We made intensive statistical analysis on the reference voltage, considering two Silicon batches.
In Chapter 3 we propose the design of a reference current generator with very low process sensitivity. It is based on a bandgap architecture with the use of diffusion resistors, that represent one of the best options to minimize sensitivity to process variations. This requirement is coupled to very low power consumption, obtained with a proper design. Also in this case, the main drawback is a large increase in area occupation. We made statistical analysis on the reference current, considering a single silicon batch.
In Chapter 4 we propose the design of a relaxation oscillator, which uses the proposed voltage and current generators to obtain stability towards process. This is obtained together with low power consumption. The performance of the proposed oscillator is verified by simulations, because a silicon run is currently under fabrication.
In Chapter 5 we will propose a technique for the analysis of non linear loop circuits, which is useful also for the design of the previously mentioned blocks. This technique is very fast, simple and accurate, because it uses DC analysis of

commercial circuit simulators, and allows us to determine the number and the stability of circuit DC operating points. It also allows us to assess when it is necessary to include start up circuits. This technique has been successfully applied to the design of a current generator with the use of native transistors. Experimental results validate the usefulness of the method.
Finally in Chapter 6 we will present our conclusion.

2. DESIGN OF A LOW POWER, LOW PROCESS SENSITIVE REFERENCE VOLTAGE GENERATOR

2.1 Introduction

The reference voltage generator is an important building block for a wide range of analog and mixed signal circuits, such as A/D converters, DRAMs, flash memories, low dropout regulators and oscillators. It generates a reference voltage which has to be stable against process, temperature and line variations. The new kind of applications discussed in Chapter 1, such as implantable systems or passive and semi-passive transponders, also demand for a very low power consumption, because the reference voltage can be considered as a bias circuit, so its power consumption does not contribute to bandwidth or SNR of an analog system.

Until ten years ago, most reference voltage generators were based on a bipolar bandgap architecture proposed by Widlar in 1974 in [70], that achieves a very robust voltage against process, temperature and line variations, and is based on the principle illustrated in Fig. 2.1.

The reference voltage is obtained as the sum of two terms:

$$V_{ref} = V_{BE} + RI_{PTAT} \,, \tag{2.1}$$

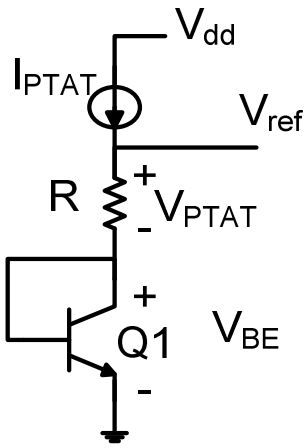

Fig. 2.1. Scheme of a bandgap voltage reference: a) Voltage mode. b) Current mode.

where V_{BE} is the base-emitter voltage of the bipolar transistor Q1. The temperature compensation is obtained by adding to V_{BE}, which has a negative temperature coefficient, a term RI_{PTAT} proportional to temperature: with the proper choice of

design parameters, V_{ref} can have a very low temperature sensitivity (as low as 11 ppm/°C [71]). In the standard bipolar bandgap voltage reference, the term RI_{PTAT} is obtained from the difference between the base-emitter voltages of two bipolar transistors operating with different current densities [70]. This architecture is also very useful in order to obtain a reference voltage with low process sensitivity (the typical relative standard deviation of the reference voltage is close to 1% [72][73]). The main drawback is represented by the difficulty to use a low power supply voltage. Indeed, to achieve temperature compensation, V_{ref} must be close to E_{gap}/q ~ 1.2 V, where E_{gap} is the silicon gap and q is the elementary charge [70]. This however implies that the reference voltage is "anchored" to the silicon energy gap, which is a physical property very stable towards process, and therefore enables to achieve a reference voltage with very low process sensitivity.
However, modern low-power low-voltage circuits need reference voltages well below 1 V. Several techniques have been proposed in order to design sub-1 V CMOS bandgap references [74]. Among them, a remarkable one is based on the bandgap principle but with a proper topology, in which compensation occurs by summing two currents, instead of two voltages, with opposite temperature coefficients. The drawback is a noisier reference voltage because of the contribution of noise due the current mirrors [71]. The principle is shown in Fig. 2.2, where I_{VBE} indicates a current proportional to the base-emitter voltage of a bipolar transistor, and was proposed by Banba et al. [75].

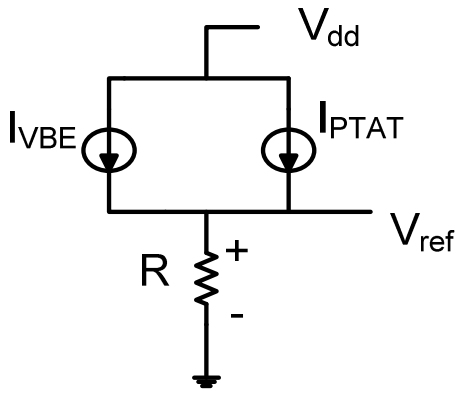

Fig. 2.2. Scheme of a bandgap voltage reference current mode.

It is based on the observation that even if the voltages in the bandgap generator are relatively large, currents can be made very small by properly choosing high resistance values, and can be effectively used for temperature compensation (the obtained temperature sensitivity is of about 120 ppm/°C). This however implies a trade off between power consumption and area occupation. In particular authors obtained a current consumption greater than 1 µA with a large area occupation of

0.1 mm^2. Another important technique is based on the reverse bandgap principle [71], which obtains the reference voltage as the sum of a PTAT voltage and a fraction of the bipolar transistor V_{BE} voltage.
However, in order to meet the requirements of low power consumption and reduced area occupation, in recent years voltage reference generators have also been proposed [76][77][78][79][80][81][82] based on the use of MOSFETs with standard CMOS process. In fact bipolar transistors can be replaced with MOSFETs biased in subthreshold region, and temperature compensation can be again obtained by properly summing two terms with opposite temperature coefficient.
A term proportional to temperature can be obtained as the difference between the gate-source voltages of two subthreshold MOSFETs. Indeed, in subthreshold, as already said, we can express the gate-source voltage V_{GS} of a MOSFET (biased with a current I_d) as:

$$V_{GS} = V_{th} + \eta V_T \ln\left(\frac{I_d L}{\mu C_{ox} W V_T^{\,2}}\right) \tag{2.2}$$

where V_{th} is the MOSFET threshold voltage, L and W are the channel length and width, η is the subthreshold slope factor, V_T is the thermal voltage, μ is the carrier mobility and C_{OX} is the gate oxide capacitance per unit area. By neglecting the mismatch between the threshold voltages of two MOSFETs M1 and M2 biased with the same current I_d, the difference between their gate-source voltages ΔV_{GS} is proportional to temperature:

$$\Delta V_{GS} = V_{GS1} - V_{GS2} = \eta V_T \ln\left(\frac{L_1 W_2}{W_1 L_2}\right). \tag{2.3}$$

The term with the negative temperature coefficient can obtained starting from the expression of the gate-source voltage of a subthreshold MOSFET biased with a constant current, since in (2.2) V_{th} is predominant and decreases with temperature.
Generators based on the use of only MOSFETs with standard CMOS process can have a very small power consumption (down to 2.2 pW in [82]), due to the MOSFETs bias in the subthreshold region, however they are intrinsically prone to large process variability, as in all cases the expression of the reference voltage contains as an addendum the MOSFET threshold voltage, which is subject to significant process variations. For example [77] estimates for V_{th} a tolerance of ±20%, which is confirmed by the fast and low corners of typical design kits.
If we accept as our priority a robust reference voltage with respect to process variability, the BJT-based bandgap topology is the most effective, because it is based on the use of quantities which are "intrinsically" more stable towards process. Actually, the BJT-based topology continues to be very popular, also in standard CMOS technologies, and it allows to obtain good results also in terms of area occupation [73][83], low supply voltage and low power consumption [83]. For example [83], which uses a reverse bandgap principle with a switch-capacitor

voltage sampling scheme, obtains a low power consumption of less than 200 nW with a process sensitivity of the reference voltage still of the order of 1%. A better result has been obtained by [73] (σ=0.8%), but with a current consumption of 1.4 μA at 1.1 V.
The power consumption of [73][83] can be however too large if we consider the low power applications above mentioned: a sub 100 nW total power consumption could be more useful. This is why we propose the design of a BJT-based topology which couples low power consumption (smaller than 100 nW) with a record-low dispersion of the reference voltage (0.18%). The main drawback is a sizeable increase in area occupation.

2.2 Description of the chosen topology

In order to operate with a low power consumption and a low power supply voltage, we consider the topology proposed by Banba [75], which was implemented in a standard CMOS process using substrate pnp transistors. This topology, (and in particular bipolar transistors), as already said, has also been chosen on the basis of our variability-aware approach, because the bipolar bandgap architecture is "intrinsically" more stable towards process.
In Fig. 2.3 the bandgap core is shown, which provides the current proportional to temperature. We call *m* the current mirror ratio of M1-M2 ($I_1 = mI_2$, where I_1 and I_2 are the emitter currents of Q1 and Q2, respectively), and *n* the ratio of the inverse saturation currents of Q2 and Q1, I_{s2} and I_{s1} ($I_{s2} = nI_{s1}$): for example, Q2 consists of n copies of transistor Q1 connected in parallel. Furthermore, we call α_1 the current mirror ratio of M2 and M3 ($I_{out1} = \alpha_1 I_2$) and $I = I_1 + I_2$. With these assumptions we can write (neglecting the base currents of Q1 and Q2 and the operational amplifier offset):

$$\Delta V = R_1 I_2 = V_T \ln\left(\frac{I_1}{I_{s1}}\frac{I_{s2}}{I_2}\right) = V_T \ln(nm) = \frac{IR_1}{(m+1)} \tag{2.4}$$

where ΔV is the voltage drop across R_1.
Furthermore:

$$I_{out1} = \alpha_1 \frac{V_T \ln(nm)}{R_1}. \tag{2.5}$$

This current is proportional to the absolute temperature, as required.

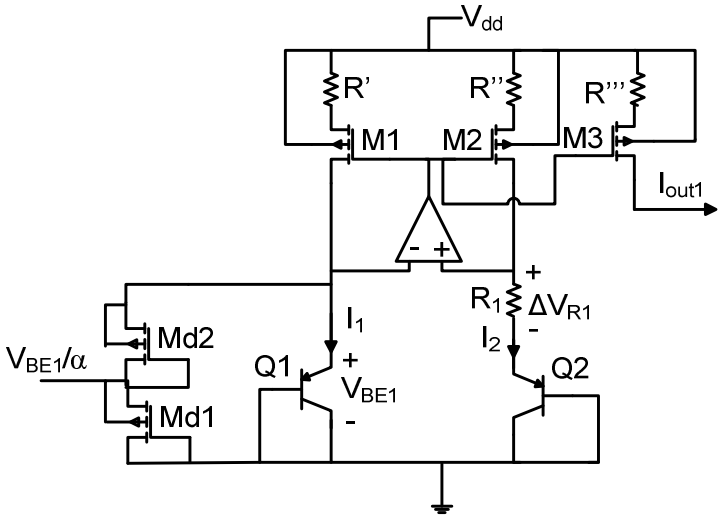

Fig. 2.3. Proposed bandgap core

Diode connected MOSFETs Md1 and Md2 constitute a voltage divider which will be explained later.
The complete bandgap voltage generator is shown in Fig. 2.4.

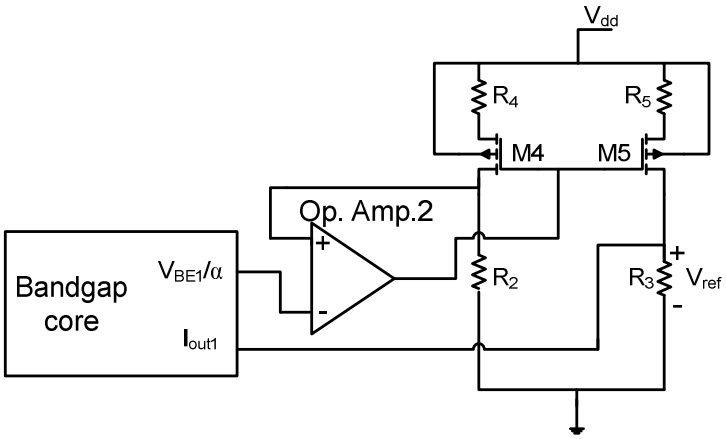

Fig. 2.4. Bandgap reference voltage generator

The second operational amplifier is used in order to impose on R_2 a fraction α of the base-emitter voltage of Q1, obtained from a voltage divider (Fig. 2.3), consisting of two diode-connected pMOSFETs in series, each realized in a different well in order to suppress the body effect. The divider itself has a negligible power consumption and it is used in order to reduce the voltage drop on - and the current through - R_2. The current in R_2 is mirrored to R_3 and added to I_{out1} in order to implement the current mode bandgap architecture. The reference voltage V_{ref} can be expressed the voltage drop across R_3, as:

$$V_{ref} = R_3\left(I_{out1} + \alpha_2 \frac{V_{be1}}{\alpha R_2} \right) = R_3\left(\alpha_1 \frac{V_T \ln(nm)}{R_1} + \alpha_2 \frac{V_{be1}}{\alpha R_2} \right) \quad (2.6)$$

where α_2 is the current mirror ratio of M5 and M4 ($I_5 = \alpha_2 I_4$). We note from (2.6) that we can minimize temperature sensitivity by properly choosing the coefficients of the two terms.
Expression (2.6) does not consider the effect of inter die and intra die variability (which affects operational amplifiers offset and current mirror ratios, for example). While our purpose is a very stable voltage towards process, in what follows we will present an accurate description of the various sources of variability in the reference voltage expression, in order to opportunely reduce their effect on the reference voltage itself.
For what concerns inter die and inter batch variability, from relation (2.6) we can note that it affects more heavily the reference voltage through the variation of the bipolar transistor Q1 base emitter voltage and through the resistances variations. Intra-die variability, instead, is the responsible of mismatch effects which can heavily affect the operational amplifier offsets and the current mirror ratios.

2.3 Analysis of variability sources

2.3.1 Bipolar transistors base-emitter voltage variation

The process sensitivity of the reference voltage (in terms of inter-die and inter-batch variability) mainly arises from the second term, in which the base-emitter voltage of Q1 appears. The base emitter voltage V_{be1} or the bipolar transistor Q1 can be expressed as:

$$V_{be1} = V_T \ln\left(\frac{I_1}{I_{s1}} \right) = V_T \ln\left(\frac{mV_T \ln(nm)}{R_1 I_{s1}} \right), \quad (2.7)$$

where

$$I_{s1} = \frac{qA}{Q_b}\left(N_C N_V e^{-\frac{E_g}{KT}} \right) V_T \mu_h , \quad (2.8)$$

and we have considered unity ideality factor of the pn junction. In this expression N_C (N_V) is the effective density of states of silicon conduction (valence) band, and is proportional to $T^{3/2}$ [84] through a constant K_1 (K_2). In addition, A is the junction area, Q_b is the base charge for area unit, μ_h is the carrier mobility [84].
Substituting (2.8) into (2.7) we have:

$$V_{be1} = V_T \ln\left(\frac{mV_T \ln(nm) Q_b}{R_1 A K_1 K_2 T^3 e^{-\frac{E_g}{KT}} KT\mu_h}\right) = \frac{E_g}{q} - V_T \ln\left(\frac{R_1 A K_1 K_2 T^3 KT\mu_h}{mV_T \ln(nm) Q_b}\right)$$

$$= \frac{E_g}{q} + V_T \ln\left(\frac{IQ_B \delta}{KT\mu_h A K_1 K_2 T^3}\right) \qquad (2.9)$$

where $\delta = m/(m+1)$. From this expression we can observe that the base-emitter voltage is the sum of E_g/q, which is a property of silicon (independent of process for medium-low doping), and of a second process-sensitive term, where the terms R_1, A, μ_h and Q_b can have a dispersion of up to 10-20%. In order to reduce the process sensitivity of V_{be1} it is important to reduce the weight of the second term. This requires large n, large m and small A and R_1. We can note, however, that the effect of m, n, A and R_1 in (9) is attenuated by their being arguments of a logarithm.
From this expression we can highlight the expected trade-off between the bandgap core current consumption I and the process sensitivity of V_{be1}.

2.3.2 Resistors variation

In (2.6) only resistance ratios appear explicitly, and can be made very precise with a proper layout. However, also V_{be1} depends on R_1, as can be seen in (2.9), and this introduce a source of variability in terms of inter-die and inter-batch variations (which produce correlated variations between resistor parameters). In order to mitigate the latter issue, we propose a method which is based on the fact that the relative process sensitivity of poly resistance increases in a predictable way with decreasing resistor width W. For example, Fig. 2.5 shows the relative process variation of the high-resistivity non-salicide poly resistors in the UMC 0.18 µm process as a function of their width W, considering the maximum resistors variation as predicted by corner analysis. We can note that this relative variation rapidly increases if W is reduced below 2 µm. This property can be successfully used in order to reduce the effect of resistor process sensitivity on the reference voltage. Indeed, we can write the second term of reference voltage expression in (6) as $V'=V(R_1)R_3/R_2$, where V is a function of R_1, resistance R_i $(i=1, 2, 3)$ is a function r of its width W_i, i.e. $R_i=r(W_i)$, and $\partial R_i/R_i=g(W_i)$.

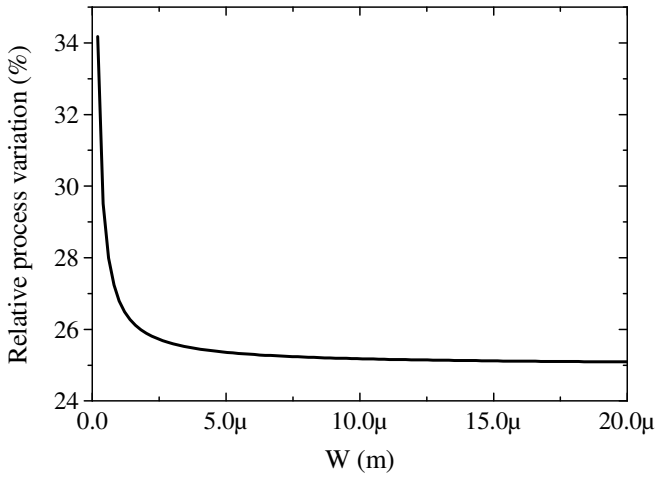

Fig. 2.5: Relative process variation of resistors as a function of their width W

We can express the relative variation of V' (∂V'/V') as:

$$\frac{\partial V'}{V'} = \frac{\partial V}{V} + \frac{\partial R_3}{R_3} - \frac{\partial R_2}{R_2} = \frac{\partial V(R_1)}{\partial R_1}\frac{1}{V(R_1)} g(W_1)R_1 + g(W_3) - g(W_2) \qquad (2.10)$$

From this expression we can note that by properly choosing resistor widths we can reduce the process sensitivity of V_{ref} in (2.6). The effectiveness of this method is based on the hypothesis that uncorrelated variations of resistor geometries are negligible with respect to the correlated ones.
In summary, since the two terms of (2.6) contain the resistance ratio as a multiplying factor, by properly selecting resistor widths we can adjust the process sensitivity of resistance ratios as to compensate the process sensitivity of other resistance-dependent terms (in this case V_{be1}).

2.3.3 Mismatch analysis

The main mismatch sources which affect the reference voltage stability towards process are:

- the input offset voltage of the two operational amplifiers;
- the mismatch in current mirrors;
- the mismatch in bipolar transistors
- the mismatch in the voltage divider.

2.3.3.1 Input offset voltage of the operational amplifier

The two operational amplifiers have a single stage, as shown in Fig. 6, and MOSFETs are biased in subthreshold in order to reduce the total power consumption.
We evaluate the effect of mismatch between Ma-Mb and Mc-Md on the input offset

voltage V_{os}.

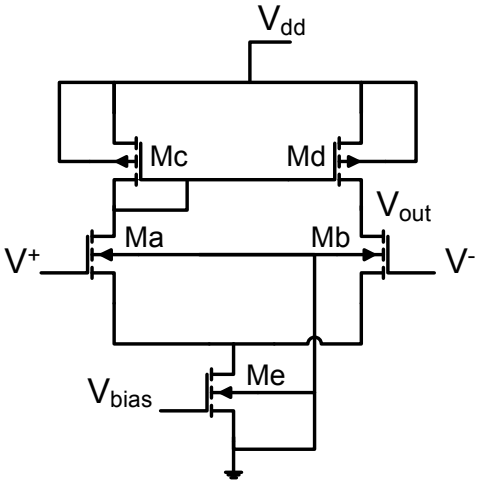

Fig. 2.6: single stage operational amplifier

By considering Ma-Mb as a subthreshold current mirror, we can express the ratio of the current in Mc, Ic, to the one in Md, Id, as:

$$\frac{I_d}{I_c} = \frac{\beta_d}{\beta_c} e^{\frac{\Delta V_{thc,d}}{\eta V_T}} \tag{2.11}$$

If we consider $\beta_d = \beta_c + \Delta\beta_{c,d}$ and $V_{thd} = V_{thc} - \Delta V_{thc,d}$ (due to mismatch), we obtain:

$$I_d = I_c \left(1 + \frac{\Delta\beta_{c,d}}{\beta_c}\right) e^{\frac{\Delta V_{thc,d}}{\eta V_T}} \tag{2.12}$$

We call $I_a=I_c=I$, $I_b=I_d=I+\Delta I$, and we consider the parameters of Ma as the nominal ones and those of Mb (nominally identical) as affected by a variation due to mismatch, so $\beta_a = \beta$, $\beta_b = \beta + \Delta\beta$, $\Delta V_{th} = V_{tha} - V_{thb}$.
We have:

$$V_{os} = V_{GSa} - V_{GSb} = V_{tha} + \eta V_T \ln\left(\frac{I_a}{\beta_a V_T{}^2}\right) - V_{thb} - \eta V_T \ln\left(\frac{I_b}{\beta_b V_T{}^2}\right) \tag{2.13}$$

and, as a first order approximation:

$$V_{os} = \Delta V_{tha,b} - \Delta V_{thc,d} - \eta V_T \frac{\Delta\beta_{c,d}}{\beta_c} + \eta V_T \frac{\Delta\beta_{a,b}}{\beta} \quad (2.14)$$

From this expression we can conclude that the input offset voltage mainly depends on the mismatch of MOSFET threshold voltages. Since the input offset voltage of the core operational amplifier directly adds to the voltage drop across R_1, as indicated by relation (2.15),

$$I_{out1} = \alpha_1 \frac{V_T \ln(nm) + V_{OScore}}{R_1}, \quad (2.15)$$

it is important to reduce the standard deviation of threshold voltage mismatch with the use of large size MOSFETs. Furthermore, considering (4), the voltage drop across R_1 can be increased (until reaching the desired error in the reference voltage due to the operational amplifier offset voltage, for example less than 0.5%) by increasing *n* and R_1 and by decreasing m. This allows us to obtain a large voltage drop across R_1 without increasing I too much, and so providing a good trade off between process sensitivity of the reference voltage and low power consumption of the core.
This choice of *n, m* and R_1 is in contrast with what needed in order to reduce the process variability of V_{be1}, which however was less sensitive to those parameters, since they were arguments of a logarithm.
The offset voltage of the second operational amplifier is directly added to V_{be1}/α, so - through coefficient $\alpha_2 R_3 / R_2$ - adds to the dispersion of the reference voltage.

2.3.3.2 Mismatch in current mirrors

We consider the current mirror M1-M2 (Fig. 3), with source resistors R' and R'', and assume a mismatch ΔI between the two currents I_1 and I_2 ($I_2 = \frac{I_1}{m} + \Delta I$), with $R'' = mR'$. We consider the M1 parameters as the nominal ones and the M2 parameters as affected by a variation due to mismatch: $\beta_2 = \frac{\beta_1}{m} + \Delta\beta$ and $V_{th2} = V_{th1} + \Delta V_{th}$.
Starting from:

$$R' I_1 + V_{GS1} = R'' I_2 + V_{GS2} \quad (2.16)$$

we obtain, at the first order:

$$\Delta I = \frac{-\Delta V_{th} + \eta V_T \ln\left(1 + \frac{m\Delta\beta}{\beta}\right)}{\left(mR' + \frac{m\eta V_T}{I_1}\right)}. \tag{2.17}$$

From this expression we can note the importance of the insertion of the source resistors in order to reduce ΔI. The ratio of the two currents can be expressed as:

$$\frac{I_2}{I_1} = \frac{\beta_2}{\beta_1}\exp\left(\frac{-mR'\Delta I - \Delta V_{th}}{\eta V_T}\right). \tag{2.18}$$

Similar considerations can be made for the other current mirror (M4-M5 of Fig. 2.5) and enable to dimension MOSFET areas and source resistors to achieve desired mismatch in current mirror.

2.3.3.3 Mismatch in bipolar transistors

The effect of mismatch between Q1 and Q2 can be important, especially by considering the large value of n and hence the very different current densities for transistor Q1 and each of the individual transistors in Q2. This could imply a difference in their inverse saturation current value and temperature and/or process dependence. The accurate model provided by the design kit allows us to evaluate these effects by means of corners analysis and Monte Carlo simulations.

2.3.3.4 Mismatch in the voltage divider

We consider the divider composed by two diode-connected pMOSFETs in series with short-circuited source and well (see Fig. 3). We consider $V_{thd2} = V_{thd1} + \Delta V_{thd}$ (so we neglect the difference in the beta value of the two nominally identical MOSFETs and we consider only a difference in the threshold voltage with respect to the one of Md1). We can express the V_{BE1}/α voltage as:

$$\frac{V_{BE1}}{\alpha} = \frac{1}{2}V_{BE1} - \frac{1}{2}\Delta V_{thd} \tag{2.19}$$

So, in order to reduce the effect of voltage divider mismatch, it is important to reduce the MOSFETs threshold voltage standard deviation.

All these relations and considerations support our choice of the design parameters and enable us to assess the effect of mismatch sources on the reference voltage.

2.4 Circuit design

The design of the proposed reference voltage generator [85] in a UMC 0.18μm CMOS process is based on the above discussion. In particular, we obtain a very good trade-off between power consumption and process sensitivity of the reference voltage. Obviously the choice of the design parameters not involved in this trade-off (as for example the current mirror ratios α_1 and α_2) has been done in order to obtain a low temperature sensitivity of the reference voltage.

We reduced the process sensitivity of V_{BE1} mainly with the use of pnp transistors with the minimum emitter area provided by the design kit. Their area (5 μm x 5 μm) is not the minimum area achievable on the basis of layout rules, but for transistors with a very small emitter area, due to three-dimensional edge effects, the current component proportional to the perimeter becomes predominant, limiting the advantages achievable with area reduction.

Our variability aware approach leads us to an accurate evaluation of all mismatch sources and to an accurate design. We estimate a standard deviation of the input offset voltage of the operational amplifier V_{OScore} of 0.72 mV using a single stage with M11-M22 width W=60 μm, length L=50 μm and M33-M44 W=10 μm, L=50 μm. On the basis of (1), in order to increase ΔV to make it insensitive to V_{OScore} without a large increase of the total current drawn from the power supply, it is important to have a small m (we choose m=3), large n (we choose n=50), and a large R_1 (R_1=20.94 MΩ). With this choice we obtain a ΔV=134.5 mV with a core current consumption of 26.16 nA. The main price to pay for this choice is the obvious large increase of the total area occupation due to large resistors and large n.

We also choose α_1=1, α=2, α_2=1/3, R_2=16.31 MΩ, R_3=19.5 MΩ, obtaining V_{BE1}/α =286.6 mV. The effect of the offset voltage on the reference voltage (in terms of relative standard deviation) has been assessed in 0.21% for the core operational amplifier and 0.12% for the second one.

The current mirrors have been realized with pMOS of width larger than 100 μm and length of 50 μm, and with source resistance R'=2 MΩ and R_4=4 MΩ. This gives I_1=19.7 nA and I_4=17.6 nA. With this choice we obtain a relative variation of the reference voltage (σ/μ) of 0.12% due to the core current mirror mismatch, and of 0.08% due to the other current mirror (composed by M4-M5). The main drawback of the use of large source resistors is the large area occupation, but they are very effective in the reduction of the mismatch impact on the reference voltage: without these resistors, the effect of core current mirror mismatch on the reference voltage is 0.33% (M1-M2) and 0.32% (M4-M5).

The mismatch in the voltage divider has been assessed in a negligible 0.09% relative variation of the reference voltage.

The effect of process variation of bipolar transistors and resistors on the reference voltage is respectively 0.25% and 0.19%, as obtained from Monte Carlo analysis.

The effects of the various sources of mismatch on the reference voltage are summarized in Table I.

Source of mismatch	*RSD (Relative Standard Deviation) on the reference voltage*
Core operational amplifier offset	0.21%
Second operational amplifier offset	0.12%
Core current mirror mismatch	0.12%
M4-M5 current mirror mismatch	0.08%
Voltage divider mismatch	0.09%
Bipolar transistors mismatch	0.25% (Monte Carlo)
Resistors mismatch	0.19% (Monte Carlo)

Table I: effect of the various sources of mismatch on the reference voltage RSD

Start up circuitry is not used because we verify with simulation that the circuit has no stability or start up problems. The problem of the number and the stability of the circuit operating points will be discussed in more details in Chapter 5. The stability and the absence of start up problems was also verified by experimental results. The start up time is of a few hundred ms, due to the presence of high impedance nodes in the circuit.

2.5 Experimental results

The chip layout is shown in Fig. 2.7: chip photo is not shown because dies are passivated with dummy layers which prevent us to see the circuit geometry. We can note the large area occupation (~0.28 mm^2) due to large resistors and the large number of BJTs in parallel. Measurements were performed on 20 packaged samples from a single batch, with the use of an Agilent E3631A DC Power Supply and an HP3478A digital multimeter. The input impedance of this multimeter, in the considered voltage range, is larger than 10 GΩ, and therefore much larger than the output resistance of the voltage generator (~20 MΩ). The average reference voltage is 240.9 mV (with respect to the 241.7 mV predicted by simulations), with a nominal supply voltage of 1 V at room temperature (27 °C). In these nominal conditions the current consumption is 68.34 nA.

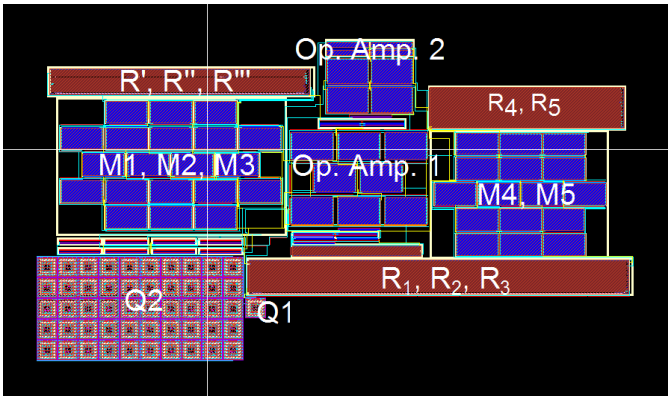

Fig. 2.7: chip layout (700 µm x 400 µm)

For what concerns line sensitivity, Figs. 2.8(a)-(b) show that the circuit properly operates for a supply voltage between 0.8 V and 1.4 V, with a mean line sensitivity of 0.12 %/V at room temperature. The PSRR of the proposed generator, which has been measured with the SR785 Dynamic Signal Analyzer, is of -68 dB at 100 Hz and it is lower than -50 dB at frequencies up to 100 kHz. At higher frequencies, also the pad and the input instrument capacitance contribute to maintain a low PSRR.

For a bandgap reference voltage it is also important the temperature stability, as already said. Fig. 2.8(c) shows the reference voltage as a function of supply voltage at different temperatures. The temperature sensitivity of the reference voltage is 97.7 ppm/°C from -25 °C to 80 °C, with the nominal supply voltage of 1 V. This temperature coefficient is not of the order of a few ppm/°C because we only apply a first-order temperature compensation, so second order effects, especially related to V_{BE} voltage, are not compensated.

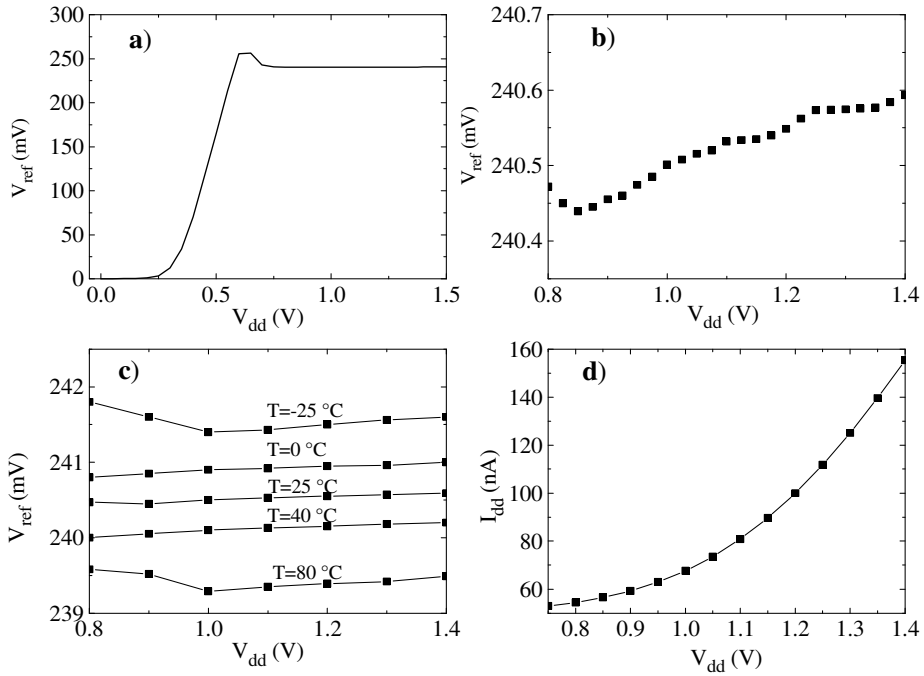

Fig. 2.8. a) and b) Reference voltage as a function of supply voltage. c) Reference voltage as a function of supply voltage for five different temperatures. d) Circuit current consumption as a function of supply voltage.

Current consumption at room temperature, which has a nominal value of 68.34 nA when the power supply voltage is 1 V, varies between 54.5 nA for a line voltage of 0.8 V, and 155.4 nA for a line voltage of 1.4 V, as shown in Fig. 2.6(d). From 1 Hz to 100 kHz, the noise power spectrum is flat and close to 2μV/√Hz.

Since our work is focused on the reduction of process variability, we show extensive statistical analysis of experiments on a single batch and compare them with results from Monte Carlo simulation to assess inter-batch variability. In particular we measure the statistical distribution of the reference voltage in nominal conditions (a power supply voltage of 1 V at 25 °C), which is shown in Fig. 2.9a, the statistical distribution of the current consumption (which is shown in Fig. 2.9b), and the statistical distribution of the line and temperature sensitivities (showed respectively in Fig. 9c and Fig. 9d). We can note the extremely low dispersion of the reference voltage value: in fact the maximum measured reference voltage variation is 0.67%, while the relative standard deviation is 0.18%. Also the total power consumption of the proposed bandgap has a low relative standard deviation of 1.6%, while line and temperature sensitivities have a greater dispersion between the 20 samples.

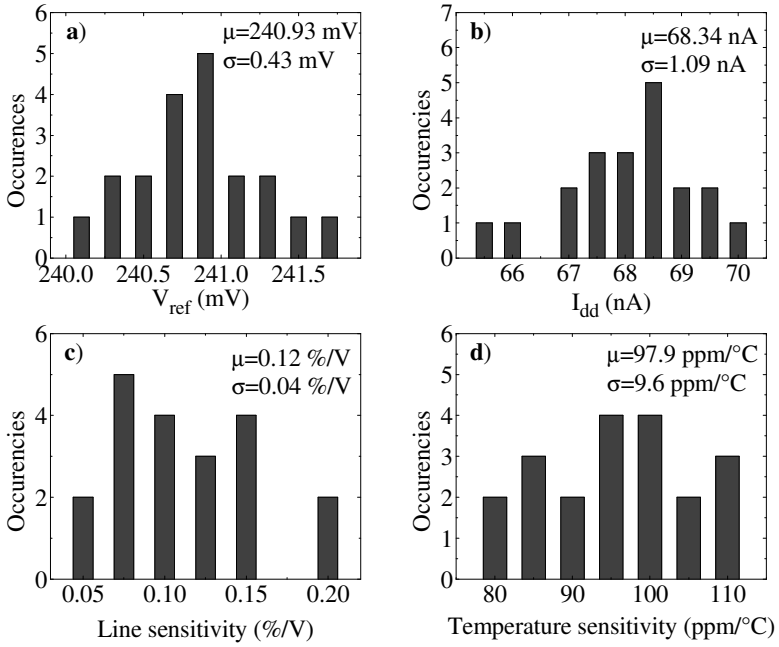

Fig. 2.9. Distribution over 20 samples at nominal conditions (V_{dd}=1 V, T=25 °C) of: a) Reference voltage. b) Supply current. c) Line sensitivity of the reference voltage. d) Temperature sensitivity of the reference voltage in the interval -25 °C÷ 80 °C.

These very good results in terms of process variability are also confirmed by Monte Carlo simulations considering MOSFETs, BJTs and resistors parameters, which show a reference voltage process sensitivity (in terms of σ/μ) of only 0.48%. As we can see from Fig. 2.10, the relative standard deviation of the reference voltage is very good also for power supply voltages and temperatures far from the nominal values. In particular, σ/μ= 0.17% when the line voltage is 0.8 V at room temperature and σ/μ = 0.19% when the line voltage is 1.4 V at room temperature. With the nominal line voltage (1 V), the σ/μ of the reference voltage value is 0.18% at -25 °C and 0.19% at 80 °C.

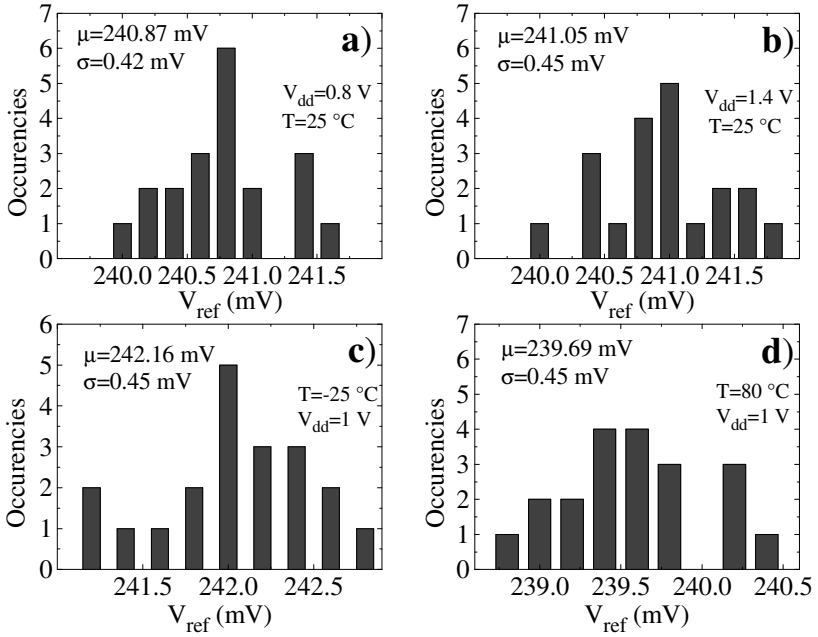

Fig. 2.10. Distribution of the reference voltage over 20 samples: a) At V_{dd}=0.8 V, T=25 °C. b) At V_{dd}=1.4 V, T=25 °C. c) At T=-25 °C, V_{dd}=1 V. d) at T=80 °C, V_{dd}=1 V.

Finally, we provide information on inter-batch variability by considering two batches of a slightly different voltage reference circuit. It is identical to the described circuit, except for the larger area occupation (due to the choice of n = 200), and lower-current voltage divider (leading to higher TC). The distribution of the reference voltage in the two batches is illustrated in Fig. 2.11: The relative standard deviation of the reference voltage is of 0.13% for the first batch, 0.19% for the second, and 0.35% in total.

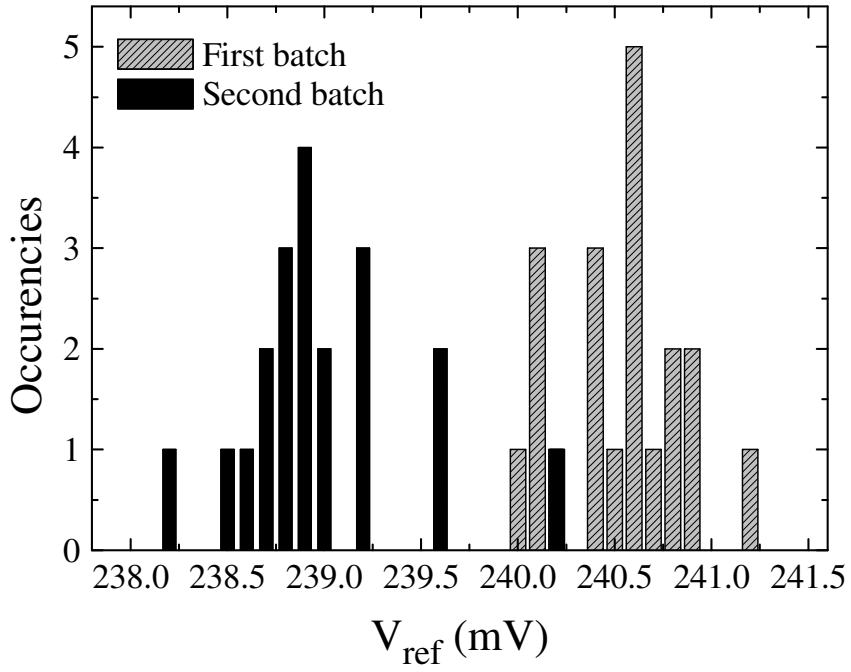

Fig. 2.11. Distribution over 40 samples from two batches of the reference voltage value in nominal conditions.

2.6 Comparison with literature

Table II compares performance figures of the proposed bandgap with those of voltage generators presented in literature with a power consumption smaller than 1 μW, which can be useful for ultra-low-power applications. Refs [83] and [86] are based on the bipolar bandgap architecture, while the other designs are based on subthreshold MOSFETs. Our solution exhibits - by far - the lowest relative standard deviation of the reference voltage, even if it was obtained considering only one batch.

	This work	[83] (2012) BJT	[86] (2010) BJT	[77] (2009)	[78] (2007)	[81] (2011)	[82] (2012)
Tecn. (µm CMOS)	0.18	0.13	0.35	0.35	0.35	0.18	0.13
V_{ref} (mV)	240.9	256	553	745	670	263.5	176
V_{dd} (V)	0.8÷1.4	0.75	1.1÷3.3	1.4÷3	0.9÷4	0.45÷ 1.8	0.5÷3
Power or I_{dd}	68.34nA	170 nW	110nW	300 nW	40÷55nA	7÷8nA	4.4÷81 pA
V_{dd} sens. (%/V)	0.12	0.005	0.11	0.002	0.27	0.44	0.033
Temp. Sens. (ppm/°C) *(°C)*	97.7 *(-25 ÷ 80)*	40 *(-20 ÷ 85)*	394 *(-20 ÷ 80)*	15 *(-20 ÷ 80)*	10 *(0 ÷ 80)*	142 *(0 ÷ 125)*	62 *(-20÷ 80)*
Process. sens. (σ/µ)	0.18% (meas.) 0.48% (MC)	1% (five batches)	1.63%	0.87%	3.1%	3.9% (three batches)	0.72% (cumulative on two batches)
Area (mm^2)	0.28	0.07	0.22	0.055	0.045	0.043	0.00135

TableII: comparison between the proposed voltage generator and results from literature

However, as already said, both Monte Carlo simulations and measurements on two batches on a previous version, based on the same principle, show very low dispersion even considering inter-batch variations. This is because we eliminated in the initial design phase any reliance on quantities - such as Vth - too sensitive to process variations.

BJT-based voltage references exhibit a power consumption comparable to our solution, but a higher process sensitivity, which for [86] is due to the use of MOSFET source-coupled pairs. Among the MOSFET-based generators, Ref [82] shows an extremely low power consumption with a very small area occupation. This reference voltage rely on the V_{th} difference of two different MOSFETs (thick oxide and native), and if we can assume some correlation between these threshold voltages, the impact of V_{th} process variability on the reference voltage is reduced with respect to a standard MOSFETs based solution. Results of [82] can be furtherly improved with digital trimming.

The choice of a circuit solution intrinsically less sensitive to process variability allowed us to obtain a manifold suppression of the relative standard deviation. It is more effective than both "internal" process compensation, which provides limited improvement since variability makes cancellation less effective, and "explicit" compensation due to feedback loops, which typically implies larger power consumption [87]. It is also an effective alternative to trimming if we consider applications, such as passive transponders and implantable applications, where trimming can be too costly.

In literature there is an example [71] of a BiCMOS reference voltage generator with a relative standard deviation of 0.19%, but with an extremely large current

consumption of 20 µA and with comparable area occupation with respect to our solution. Ref. [88] proposes a generator which, by using DTMOSTs, obtains a reference voltage relative variation of 0.3% with a power consumption of 2.5 µA. However this generator is based on a voltage-sum bandgap topology (the reference voltage value is 1.23 V), so it is not suitable for low power and low voltage applications.
It is also useful to compare in Table III our solution with voltage references using digitally controlled trimming.

	This work	[90]	[91]	[89]	[83]	[82]
Tecn.	0.18 um	0.16 um	0.35 um	0.6 um	130 nm	0.13 um
V_{ref} (V)	0.2409	1.0875	0.858	1.14205	0.256	0.1761
V_{dd} (V)	0.8÷1.4	1.8	1.4	2	>0.75	0.5÷3
Power/ Current	68.34 nW	55 µA	162 µW	23 µA	170 nW	59 pA
TC (ppm/°C) *(°C)*	97.7 *(-25÷80)*	5÷12 *(-40÷125)*	12.4 *(-20÷100)*	5.3 *(0÷100)*	40 *(-20÷85)*	5.3÷47.4 *(-20÷80)*
Process (σ/µ)	0.18%	0.05%	0.3%	0.08%	0.17%	0.16%
Area (mm^2)	0.28	0.12	1.2	0.057	0.07	0.0093

Table III: comparison between the proposed voltage generator and literature results using trimming

We can note that V_{ref} standard deviation is comparable at the price of larger power consumption of [90][91][89][83]]) and/or comparable (or larger) area occupation of [90][91]. Only [82] obtains a similar process sensitivity of the reference voltage with a much smaller power consumption and area occupation, but this solution is useful if a digital section is easy to implement.
We can note the proposed solution is in a middle position between the graph portions occupied by solutions respectively with and without trimming. It is important to underline that trimming can help us in achieving a stable reference voltage without excessive cost if a digital section is already available in the complete chip and if the trimming procedure does not increase cost too much. The proposed solution is a strong alternative, especially for systems with a very simple digital section (implantable systems, sensor interfaces), for which the implementation of a trimming procedure would be really expensive and not easy.

2.7 Conclusion

We apply a variability-aware approach on the design of a nanopower reference voltage generator, obtaining a record-low dispersion due to process variability and with a low power consumption. We have shown that a BJT-based bandgap topology is the most appropriate to the first aim, since the reference voltage is

anchored to a silicon physical property such as the energy bandgap, and does not rely on quantities sensitive to process variability as the threshold voltage.
We have derived design criteria that enabled us to obtain low power consumption of 68.3 nW and low relative standard deviation of the reference voltage of 0.18%, which is much smaller than all designs presented in the literature with sub-microwatt power consumption.
The main cost of our choices in the design space is a much higher area occupation, mainly due to the large resistances, needed to reduce the power consumption of the circuit, and the large n. Such cost is particularly acceptable in the case that most interests us, i.e. when one uses aggressively scaled CMOS technologies, which provide abundant margins in terms of die area.
A precise reference voltage generator such as the one proposed here, can be effectively used as a basic building block to provide robustness with respect to process variability to more complex circuits and systems, where one prefers not to use alternative reference circuits (such as quartz oscillators, for example) or expensive trimming procedures.

3. DESIGN OF A LOW POWER, LOW PROCESS SENSITIVE REFERENCE CURRENT GENERATOR

3.1 Introduction

The reference current generator is an important block for a broad variety of analog and digital systems. For example, it is used as a bias source for oscillators, amplifiers and it is very important for A/D and D/A converters. Its requirements are a very low reference-current sensitivity to supply voltage, temperature and especially to process. Also low power consumption is needed, if we consider portable and implantable systems.
Stability towards process can be effectively obtained by “anchoring” the reference quantity to an intrinsically stable physical quantity: for example, a reference voltage can be anchored to the silicon bandgap if we choose a bipolar bandgap topology, that was proposed by Widlar in [70] and is one of most popular architectures to realize a voltage generator. The typical obtained RSD of the reference voltage is of the order of 1% [93]. Better results can be obtained with the use of trimming.
While this approach is useful for voltage references, it is not as effective for the design of reference current generators. A reference current is usually obtained starting from a reference voltage [94], so the expression of the current consists of a voltage multiplied by a transconductive factor. This factor is usually much more sensitive to the process than the voltage term, and is therefore responsible for the large process variability of the reference current. Our approach is based on the consideration that even if we cannot "anchor" the transconductive factor to a reference quantity, we can reduce its process sensitivity by using devices that are intrinsically less sensitive to process and available also in standard CMOS technologies.
In the following subsections we will evaluate the available options, considering that the transconductive factor can be based on a resistance or on a MOSFET beta.

3.1.1 Transconductive factor based on a resistance

The transconductive factor can be obtained using a resistor, if we choose a bandgap architecture based on the sum of two currents with opposite temperature coefficients [95][96][97]. It can also more generally be obtained by imposing a reference voltage on a resistor [98]. In this case, if we want to obtain a reference current with low power consumption, we must use large resistances, which require a large area. For example [96] has an area occupation of 0.3 mm^2, while [99], which uses only MOSFETs, has the much smaller area of 0.035 mm^2 and also much smaller power consumption. Resistors-based current references generally have lower process sensitivity with respect to MOSFET-based ones: for example, the RSD of [98] is only 1.5%. This is because resistance variability depends on dopant activation and on geometry, and does not depend on the properties at the silicon-dielectric interface. Furthermore, we can use diffusion resistors instead of poly resistors. Indeed, polycrystalline silicon is more challenging from the point of

view of controlling dopant activation and mobility. Diffusion resistors usually have lower doping, larger volume and they are constituted by monocrystalline material. Lower doping implies that impurity activation is more complete and reproducible. Larger volume and monocrystalline material imply a reduced impact of localized defects and grain borders [100].
We use diffusion resistors in the implementation of a bipolar bandgap current generator. We use this very "classic" architecture because it is the most promising in order to obtain a reference voltage with low process sensitivity. In recent years several architectures for voltage references have been proposed based on the use of only MOSFETs biased in subthreshold region, in order to reduce power consumption [101]. Indeed, MOSFET-based generators are very effective for this purpose, but they are not optimized in order to obtain a low process sensitivity of the reference voltage. Indeed, the reference voltage expression includes as a dominant term the MOSFET threshold voltage, that exhibits a large spread of roughly 20%. The base emitter voltage of a bipolar transistor, instead, is “anchored” to the silicon energy gap, which is an intrinsic property of the material. This implies a lower process sensitivity of the reference voltage. Our objective is to obtain a variability-aware reference voltage with very low power consumption. We will see that we can satisfy both requirements with the drawback of a large increase in area occupation.

3.1.2 Transconductive factor based on the MOSFET beta

The transconductive factor can also be related to the MOSFET beta $\beta = \mu C_{ox} W/L$, where μ is the carrier mobility, C_{ox} is the gate oxide capacitance for unit area, W and L are the MOSFET width and length, respectively. Different architectures can be used for this purpose: for example [102] imposes a bandgap reference voltage as the gate-to-source voltage of a load MOSFET biased in saturation region. Variations of the threshold voltage of the load MOSFET can be partially compensated by using a reference voltage also depending on the threshold voltage. Compensation is not perfect, so the reference current can still have high process sensitivity (for example an RSD of 2.3% in [102]).
A reference current depending on the MOSFET beta can also be obtained by imposing a reference voltage as the drain-to-source voltage of a MOSFET biased in the triode region [103][104]. This solution is very effective if we want to reduce power consumption (the power consumption is 2 nW in [103]), but the obtained process sensitivity of the reference current is quite high (RSD of 3.3% [103]).
The self-biased architecture [105] is the typical choice in order to obtain a reference current depending on the MOSFET beta. Its simplified schematic is shown in Fig. 3.1.

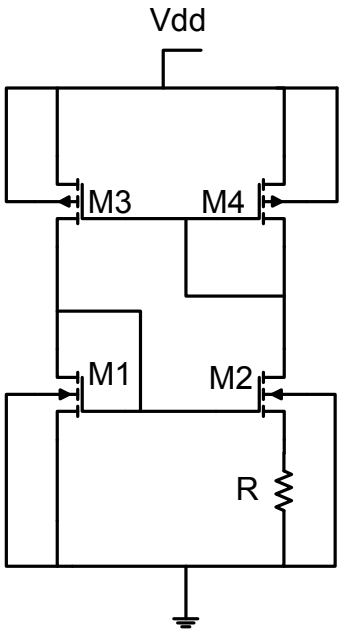

Fig. 3.1 self biased current generator

Current of M4 is mirrored in M3. Transistors M1, M2 and resistor R form a non-linear mirror that is connected back to back to the upper mirror in a positive feedback loop. Therefore the current of M1 is bootstrapped to the current of M4. This architecture can be implemented with MOSFETs and resistors, by using a circuit schematic similar to the one of Fig. 3.1 [106][107]: in this case the transconductive factor is also related to the resistance R. It can also be implemented by using only MOSFETs, replacing resistor R with a MOSFET biased in the triode region [108], or with MOSFETs having their gate and drain voltages controlled by diode-connected MOSFETs [99]. This architecture is useful in order to obtain a reference current with low power consumption (for example 54.84 nW in [99]) and low supply voltage.

The transconductive factor contains mobility, whose process variability depends on partial dopant activation and on the properties at the interface between Si and Si oxide. We generally obtain a process sensitivity of the reference current larger than 2% (see [99]), by considering only a single batch.

Following a variability-aware approach for the design of a MOSFET-based current generator, we can consider a transconductive factor based on native MOSFETs instead of the standard ones, which have been successfully used for the design of an integrated time reference [109]. Native transistors have a lower doping in their active area that should lead to reduced mobility spread. In addition, transport is less concentrated at the Si-SiO_2 interface, because of their lower threshold voltage, therefore mobility should be less sensitive to interface roughness than in the case of standard transistors. The self-biased architecture can be used, with required architectural changes in order to ensure that native transistors are biased in saturation. In fact, since native MOSFETs have negative threshold voltage, a

constant positive drain-to-gate voltage must be applied.
However, in our design kit native transistors were poorly characterized, therefore this approach was not successful. We still consider the general concept to be applicable, but for the scope of the present paper we focus on the solution based on diffusion resistors.

3.2 Design of the reference current generator

The proposed bandgap core, which generates a current proportional to temperature, is shown in Fig. 3.2, whereas the complete current generator is shown in Fig. 3.3.

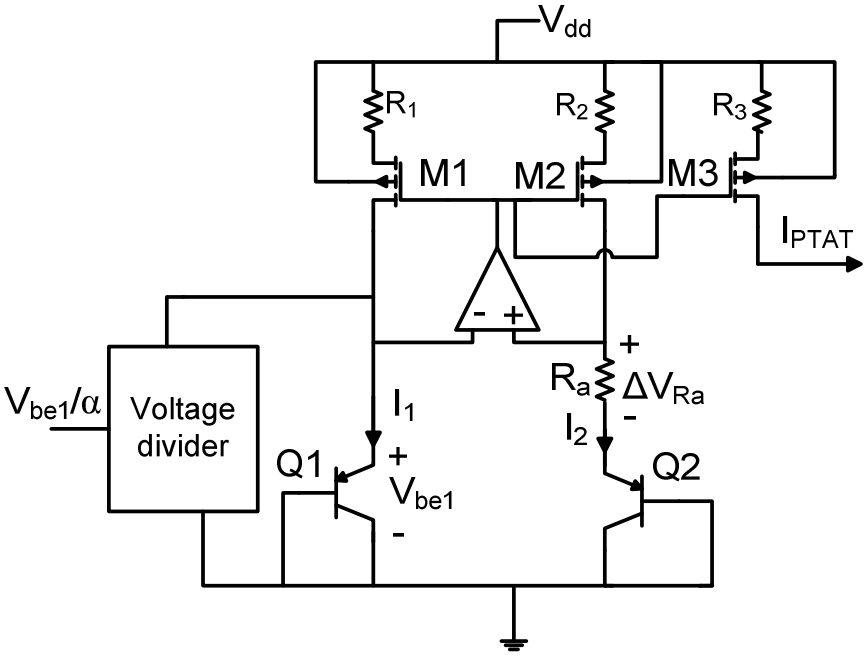

Fig. 3.2 bandgap core

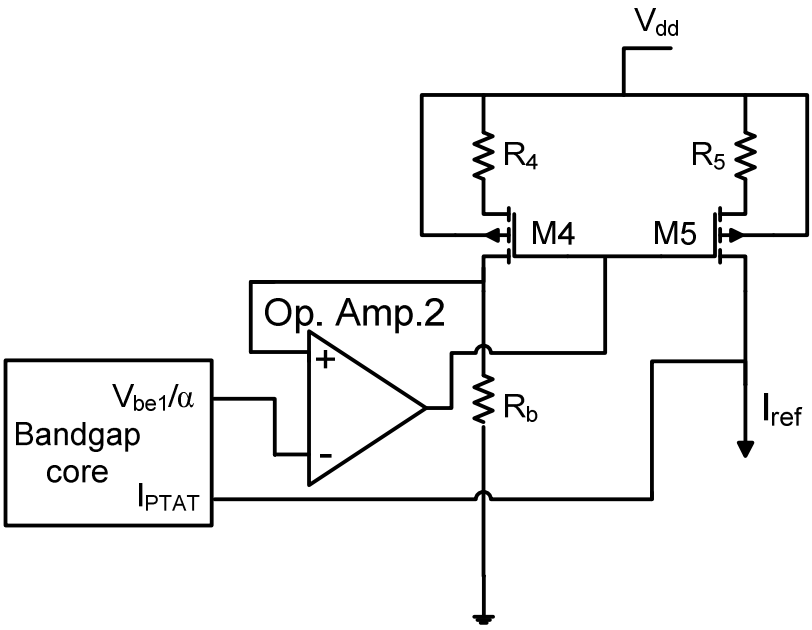

Fig. 3.3 bandgap reference current generator

The temperature stability of the reference current is obtained by properly summing two currents with opposite temperature coefficients. Neglecting the offset voltages of operational amplifiers and the other sources of mismatch, the reference current can be expressed as:

$$I_{ref} = \alpha_1 \frac{V_T \ln(nm)}{R_a} + \alpha_2 \frac{V_{be1}}{\alpha R_b} \tag{3.1}$$

where α_1 is the current mirror ratio of M2 and M3, m is the current mirror ratio of M1-M2 ($I_1=mI_2$, where I_1 and I_2 are the emitter currents of Q1 and Q2, respectively), n is the ratio of the inverse saturation currents of Q2 and Q1, I_{s2} and I_{s1} ($I_{s2}=nI_{s1}$), α_2 is the current mirror ratio of M4 and M5 and α is a resistive partition. The first term of (1) is the PTAT current I_{PTAT}, whose expression is obtained considering a nearly zero input differential voltage for the core operational amplifier of Fig. 2. The negative TC current is obtained by imposing a fraction $1/\alpha$ of the V_{be1} voltage on resistor R_b, as can be seen in Fig. 3.3.

From Eq. (3.1) we can see that R_a and R_b represent the main sources of variability for the reference current (in terms of inter-die and inter-batch variations). This is the reason we implement them with components that are intrinsically more robust towards process. Poly resistors, available in standard CMOS technologies, are heavily used for their low sensitivity to substrate noise, however their resistance exhibits a large process spread. For example, we have performed a corner analysis on the poly resistance in a standard UMC 0.18 µm technology by varying

resistor width *W*, obtaining both a high resistance sensitivity to *W*, especially for very low *W*, as shown in Fig. 3.4, and a high process sensitivity of the resistance itself, whose minimum relative process spread is about 25%.

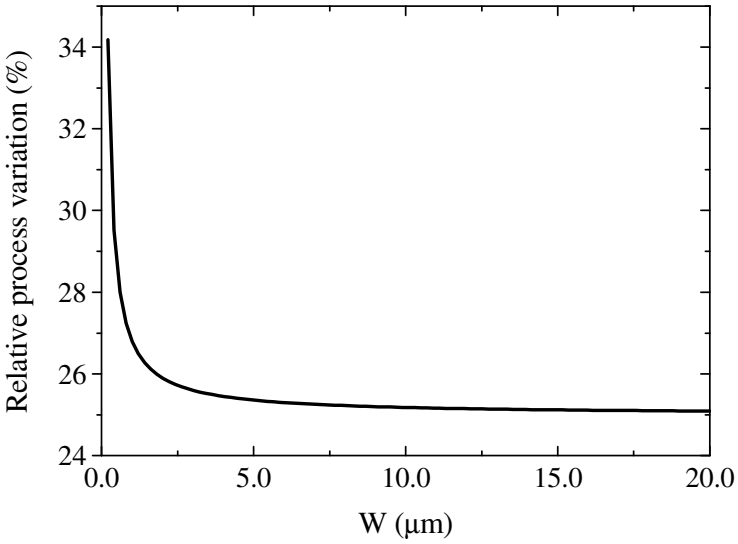

Fig. 3.4 relative process spread of resistors as a function of their width W

We also made a Monte Carlo analysis on the poly resistance, considering a nominal value of 1.42 MΩ. We obtained a relative standard deviation for the resistance of 8.76%. The design kit does not provide information on process variability of diffusion resistors, therefore we performed statistical analysis on a nominal resistance of 0.39 MΩ by measuring it over 23 samples from a single batch, obtaining a 2% RSD.

The base emitter voltage of Q1 is another source of variability for the reference current, but it can be anchored to the silicon energy gap, so it can be very stable versus process. Indeed we can express V_{be1} as:

$$V_{be1} = V_T \ln\left(\frac{I_1}{I_{s1}}\right) = \frac{E_g}{q} - V_T \ln\left(\frac{R_a A K_1 K_2 T^3 K T \mu_h}{m V_T \ln(nm) Q_b}\right) \tag{3.2}$$

where K_1 (K_2) are the proportionality constants of the effective states density of silicon conduction (valence) band N_C (N_V), and $T^{3/2}$ [23]. In addition A is the junction area, Q_b is the base charge for area unit, μ_h is the carrier mobility [23]. From this expression we can observe that the base-emitter voltage is the sum of E_g/q, which is a property of silicon (independent of process for medium-low doping), and of a second process-sensitive term, where the terms R_1, A, μ_h and Q_b have a large dispersion of up to 10-20%, but they are argument of a logarithmic function.

In the variability-aware design of the current generator it is important also to reduce

the effect of the offset and of mismatch sources. Referring to the core circuit of Fig. 2, we can note that the voltage drop ΔV_{Ra} across R_a can be heavily affected by the input offset voltage V_{io} of the core operational amplifier, as shown below:

$$\Delta V_{Ra} = V_T \ln(nm) + V_{io} = \frac{(I_1 + I_2)R_a}{(m+1)} + V_{io} \tag{3.3}$$

We reduced the standard deviation of V_{io}, obtaining $\sigma_{V_{io}} = 0.59$ mV, with the use of large input nMOSFETs of width W=120 μm and length L=50 μm for the operational amplifier. Furthermore, on the basis of (3), in order to increase ΔV_{Ra} to make it insensitive to V_{io} without a large increase of the total current drawn from the power supply, it is important to consider a small m (we choose m=2), large n (we choose n=20), and large R_a (R_a=2.9 MΩ). The use of a large R_a and large n (transistor Q2 is constituted by n transistors like Q1 connected in parallel) implies an obvious increase in area occupation.
We also choose α_1=1, α=4, α_2=1/6, R_b= 1.45 MΩ, obtaining ΔV_{Ra}=99.31 mV and V_{be1}/α=151.6 mV. With theses large voltage values the effect of the core and the second operational amplifiers offsets have been estimated in a 0.39% and 0.12% relative standard deviation of the reference current, respectively.
For the reduction of power consumption, apart from the trade off with process sensitivity indicated by relation (3), we use large resistances, even if this implies a large area occupation, we use single-stage operational amplifiers biased with a current of few nA, and we use a pMOS voltage divider to impose on R_b only a fraction $1/\alpha$ of the V_{be1} voltage. The voltage divider, shown in Fig. 3.5, consists of diode-connected pMOSFETs in series, with each well at source potential, in order not to have body effect. Its current consumption is of only about a hundred pA, negligible with respect to the bandgap core current.

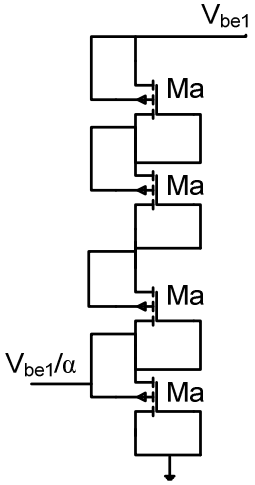

Fig. 3.5 voltage divider

It is important to consider and suppress also the effect of mismatch in current mirrors, which adds to the intra-die process variability. This is obtained with the use of large MOSFETs and with the use of source degeneration resistors. The last ones are really effective for this purpose (even if they increase total layout area), because the current mismatch, in absolute value, is inversely proportional to source resistance. We used mirror pMOSFETs with minimum W=100 µm (the effective width of each mirror transistor also depends on current mirror ratio) and L= 50 µm: this implies σ_{Vth}=0.25 mV. We also use source resistors of order 100 kΩ. We estimate a reference current RSD of 0.31%. Also the mismatch in the voltage divider MOSFETs can be neglected: simulations predict a RSD of V_{be1}/α of 0.066%, by considering pMOSFETs with W=60 µm and L=20 µm, which implies σ_{Vth}=0.32 mV. We evaluate the effect of pnp transistors mismatch with Monte Carlo analysis, obtaining a 0.2% RSD of the reference current. The results obtained from the mismatch analysis are summarized in Table I.

Variability source	Corresponding RSD (Relative Standard Deviation) on the reference current
Core operational amplifier offset	0.39%
Second operational amplifier offset	0.12%
Current mirrors mismatch	0.31%
Voltage divider mismatch	0.066%
pnp transistors mismatch	0.2% (Monte Carlo)

Table I: analysis of the effect of the various sources of mismatch on the reference current RSD

Start up circuitry is not used because we verify with simulation that the circuit has no stability or start up problems. The problem of the number and the stability of the circuit operating points will be discussed in more details in Chapter 5. The stability and the absence of start up problems was also verified by experimental results. The start up time is of a few hundred ms, due to the presence of high impedance nodes in the circuit.

3.3 Experimental results

We realized the current generator in the UMC 0.18 µm CMOS technology with the use of substrate pnp transistors and performed measurements on 23 packaged samples from a single batch. Chip layout is shown in Fig. 3.6. The chip photo is not shown because dies are passivated with dummy layers that prevent us to see circuit geometries. Measurements have been performed with the use of an Agilent E3631A DC Power Supply and an HP3478A digital multimeter: the output node of the current generator has been closed on a known external output resistor and the

output voltage has been measured instead of the output current. Different values of output resistances have also been used, in order to verify the impact of the load on the reference current. The average reference current value is 54.08 nA at a nominal supply voltage of 1 V, with a standard deviation of 0.76 nA, so the RSD is 1.4%. The mean line sensitivity of the reference current is 0.21%/V by varying the power supply voltage from 0.8 V to 1.4 V, while the temperature sensitivity is 66 ppm/°C from -25 to 80 °C. The current consumption is 288.84 nA when the power supply is 1 V.

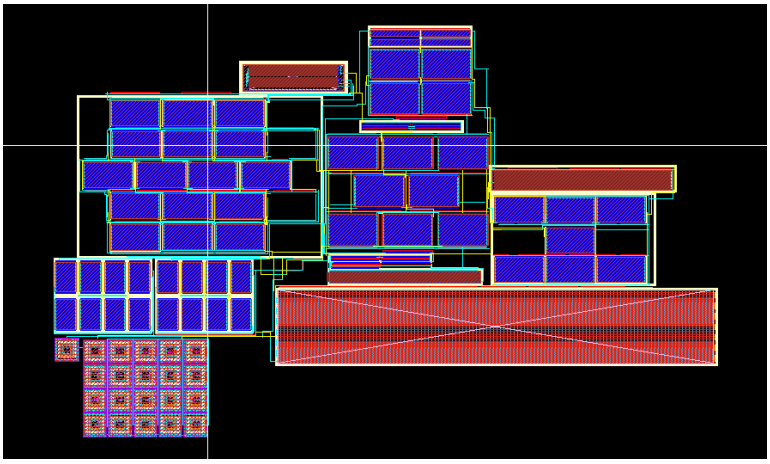

Fig. 3.6 chip layout (approx. 700 µm x 350 µm)

The performance of the proposed current reference by varying power supply voltage and temperature are summarized in Fig. 3.7, where also the power consumption as a function of power supply voltage is shown. The load sensitivity is 0.5%/V by varying the output voltage between 0 V and 0.8 V. The PSRR of the proposed generator is measured by using a SR785 Dynamic Signal Analyzer (DSA): we obtain a PSRR of -50 dB at low frequencies, -65 dB at 100 Hz and lower than -40 dB at higher frequencies. At these frequencies also pad and input instrument cross-capacitances contribute to reduce the PSRR. Flat band noise has been evaluated with the use of the DSA: it results in 0.3 pA/√Hz.

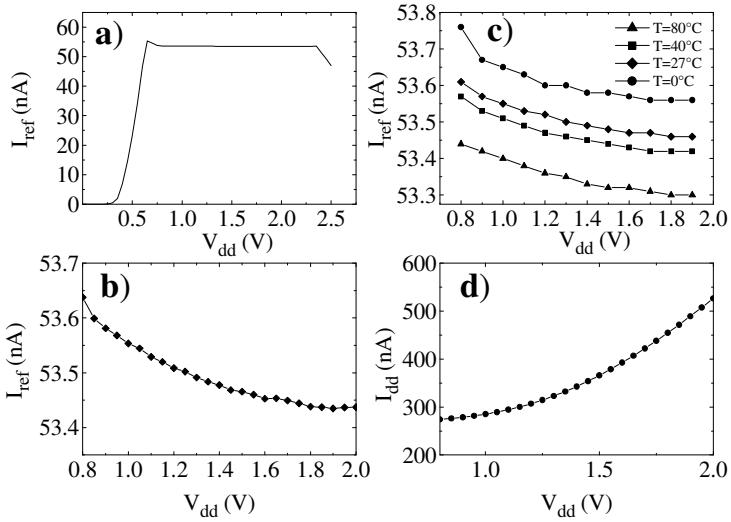

Fig. 3.7 a) and b) reference current I_{ref} as a function of supply voltage V_{dd}. c) I_{ref} as a function of V_{dd} for four different temperatures. d) circuit current consumption as a function of supply voltage.

Table II compares our results with the most relevant ones available in the literature.

	This work	[111][90]	[99][91]	[96]	[98]
Tecn.	0.18 µm	3 µm	0.35 µm	2 µm SIMOX	0.18 µm
I_{ref}	54.08 nA	774 nA	9.14 nA	19.5 µA	7.81 µA
V_{dd} (V)	0.8÷2	3.5	1.5	5	1.2
I_{dd}	288.8 nA @1 V	2 µA@5 V	36.6 nA @1.5 V	300 µA @5 V	27.2 µA @1.2 V
TC (ppm/°C) *(°C)*	63 (0÷80)	375 (0÷80)	44 (0÷80)	12 (-15÷90)	24.9 (0÷100)
Line Sens. (%/V)	0.21	0.013-0.015	0.0569	-	0.13
Process (σ/µ)	1.4%	2.58%	2.17%	1.67%	15%
Area (mm^2)	0.245	0.2	0.035	0.3%	0.123
Variability factor	P+ resistors	β std MOSFET	β std MOSFET	P+ resistors	Poly resistors

Table II: current generator performance figures and comparison with literature

References [99] and [111] present MOSFET-based architectures, while [96] and [98] use resistors. The process sensitivity of the proposed generator is comparable with the one of other current generators using resistors, but with a much lower power consumption and reference current. Generators using MOSFETs beta in their transconductive factor allow to obtain a low power consumption with a very low area occupation (see [99]), because MOSFETs can be biased in subthreshold, but with a larger process sensitivity of the reference current, as we can see from Table II. Better results in terms of process sensitivity can be obtained introducing programmability [111], or with digital trimming [97], but this implies an additional phase of calibration after chip fabrication, which can be expensive.
We made statistical analysis on a small but significant set of 23 samples from the same batch (the relative root mean square error on the standard deviation is therefore $1/\sqrt{(2 \cdot 23)} = 14.74\%$). The obtained results are reported in Fig. 3.8 and Fig. 3.9: we can note that the relative standard deviation of the reference current is very good also for line voltages and temperatures far from the nominal ones.

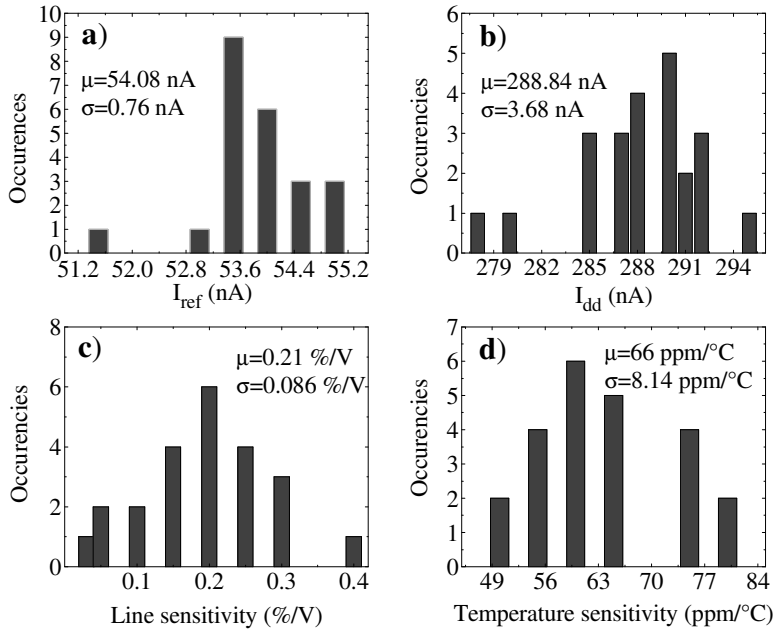

Fig. 3.8 distribution over 23 samples of: a) reference current value. b) total current consumption. c) line sensitivity of the reference current. d) temperature sensitivity of the reference current.

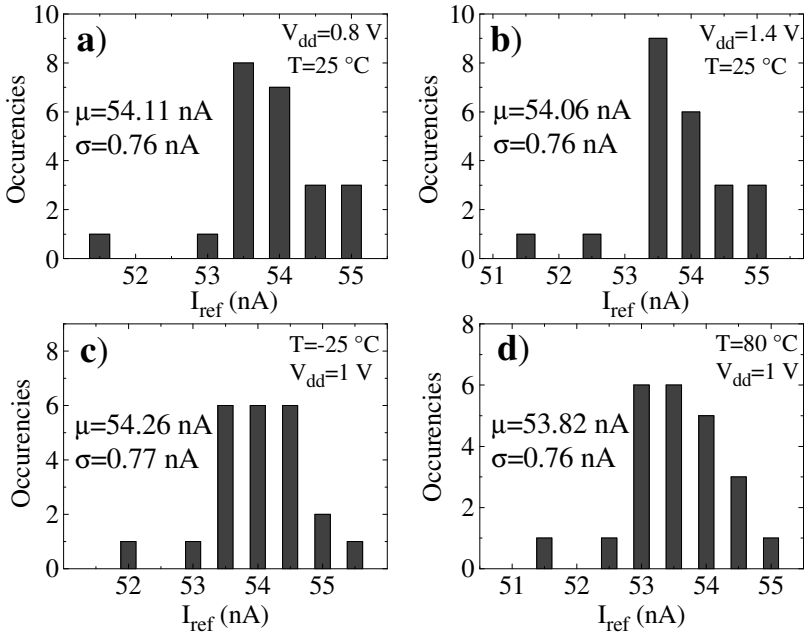

Fig. 3.9 distribution over 23 samples of the reference current value: a) at V_{dd}=0.8 V T=25 °C. b) at V_{dd}=1.4 V, T=25 °C. c) at V_{dd}=1 V, T=-25 °C. d) at V_{dd}=1 V, T=80 °C.

We do not have results from multiple batches, and also all results in the literature are from a single batch. However, while our scheme is based on intrinsically more robust quantities, we expect advantages in accuracy even if multiple batches were considered. The main drawback is the very large area occupation.

3.4 Conclusion

We propose a variability-aware reference current generator with low power consumption. In order to obtain low process sensitivity, we use devices which are intrinsically more stable towards process and which are available also in standard CMOS technologies, such as diffusion resistors, better than poly resistors or MOSFETs for this purpose. We also consider a "classic" bipolar-based bandgap architecture, which allows to obtain a reference voltage "anchored" to the silicon energy gap. An accurate analysis for the reduction of offset and mismatch has been considered together with a careful design for power consumption reduction, allowing the achievement of a very good trade off between robustness to process variability and power consumption. The main drawback of our design is a large area occupation, mainly due to the large resistances, needed to reduce the power consumption of the circuit.

4. DESIGN OF A LOW POWER, LOW PROCESS SENSITIVE RELAXATION OSCILLATOR

4.1 Introduction

The development of biomedical applications for electronic circuits poses several problems, related to reproducibility of electrical quantities and low power consumption. For these applications, a clock source is usually required for the use of clock-based signal-processing techniques. Furthermore, an on-chip clock generator is also important for portable applications such as passive transponders. In the latter case a low frequency is required, with reasonable accuracy and very low power consumption. A clock reference is also important for wireless sensor networks: networks nodes can be put in a sleep mode when they are not used in the communication, in order to reduce power consumption, but this requires a clock for synchronization to ensure the simultaneous sleep and wake-up times for all nodes [113]. Also for this application it is important to obtain a precise clock frequency (with accuracy of few percent) with a low power consumption.
Crystal oscillators are the most accurate clock generators, because they provide a very stable oscillation frequency against supply voltage, temperature and process variations, but the lack of on-chip integration increases area and cost of the system. In recent years, several architectures have been proposed in order to obtain a completely integrated frequency reference, because the presence of external (eventually trimmable) components adds to area and cost. LC oscillators can provide reasonable accuracy [114], but with a high power consumption of more than 100 µW due to the limited Q of integrated inductors. Therefore the most common architecture in order to realize an on-chip (low frequency) oscillator is the one which bases the oscillation frequency on a RC product and which can be implemented with the use of a ring oscillator or with the use of a relaxation oscillator architecture. Ring oscillators are constituted by a loop of an even number of inverters, whose bias current (inversely proportional to a resistance R) and output capacitance C determine the oscillation frequency. Relaxation oscillators are based on the charging and discharging of a capacitance C between two reference voltages by means of a (usually constant) current source inversely proportional to a resistance R. The typical performance of the two kinds of RC oscillators is presented in [70]: relaxation oscillators usually allow to obtain a lower frequency with respect to ring oscillators (in the range 1-100 kHz instead of 10-1000 MHz), with better precision (±1% instead of ±5%) and a lower power consumption (in the range 1-100 µA instead of 10 µA - 100 mA). However relaxation oscillators are usually one order of magnitude larger than ring oscillators. Furthermore, a lower power consumption implies a larger jitter, and in literature some architectures [116] have been presented in order to combine a low jitter with a high control linearity. However, jitter is not an issue if the clock is used as a wake-up timing source only. We can conclude that the relaxation oscillator seems the most useful architecture in order to obtain precise low frequency clock generators with low power

consumption. We should consider that the applications we are dealing with may require a better accuracy and lower power consumption with respect to what is typically obtained.
In order to improve the precision of clock generators, various techniques have been proposed: we can use feedback compensation circuits, or trimming, or we can use components which are intrinsically more stable towards process. Concerning feedback compensation circuits, several examples have been presented of compensated ring-oscillators, which obtain high accuracy at the expense of a mW power consumption. For example, [117] stabilizes the bias current of a ring oscillator towards process and temperature with the use of a control voltage. This is generated using a threshold voltage sensing scheme and a temperature compensation circuit. Authors obtain a 7.03 MHz oscillation frequency with large power consumption (1.5 mW) and a large area occupation (1.6 mm^2), but with a high accuracy of the so obtained frequency: its relative standard deviation (RSD) is 0.13%. A feedback loop can also be used to lock the output frequency of a VCO to an external time constant RC [118], [119], [120]: this implies good frequency accuracy with large power consumption. In [118], authors obtained an accuracy of 0.7%, including power supply, temperature and chip-to-chip variations, for a power consumption of 100 µW. In [119] they obtain an accuracy of 1.67% with a power consumption of 20 µW for [119]. Moreover these solutions are obviously not completely integrated.
An alternative way to obtain accuracy of the clock frequency is with the use of trimming, for example by using a trimmed RC oscillator [118], [121]. In [121] trimming is required only in order to compensate for inter die and inter wafer variability, while the effect of intra die variability is low, with the proper choice of components (for example long poly resistors and MIM capacitors). Also ref. [122] proposes a relaxation oscillator frequency based on an RC product with integrated resistors and capacitors, and uses trimming on these components in order to obtain a good precision. The use of feedback loops in order to lock the frequency to an external RC product can also be combined with the possibility of internal trimming and programmability, as in [120].
The third approach to improve accuracy is with the use of components intrinsically more stable towards process. For example, Ref. [113] proposes a MOSFET-mobility based clock generator in order to obtain high accuracy with a low oscillation frequency (100 kHz), because mobility is less sensitive to process variations than other parameters, such as polysilicon resistances and oxide capacitances. The problem of this solution is the high power consumption of about 40 µW and the high temperature sensitivity of the reference frequency, which can be reduced only with a corresponding reduction of the accuracy. In addition, the obtained frequency RSD is 2.3% and is futher improved with single-trim calibration (which also reduced the temperature sensitivity of the oscillation frequency). Ref. [123] proposes a relaxation oscillator which bases the time reference on MOSFET mobility, reaching a very low power consumption (11 nW), but again with a high temperature sensitivity, due to the large temperature sensitivity of mobility.
In conclusion, in order to generate a low clock frequency for implantable applications and passive transponders, the relaxation architecture is the most convenient choice, because it allows to obtain both high accuracy and low power consumption. We aim to reach this result without trimming, that can be a very expensive operation, and without the use of complex compensation loops, but with

the use of devices which are intrinsically more stable towards process and which are available also in standard CMOS technologies.

4.2 Description of the relaxation oscillator topology

The operating principle of relaxation oscillators is based on the continuous application of disturbances: each one starts a transient in the system which dissipates energy to return to the equilibrium state, until a threshold near the equilibrium point is reached and a new disturbance (with additional energy) is applied. The oscillation period is determined by the time span between two consecutive disturbances. The oscillator behavior is characterized by long periods of dissipation followed by short impulses. A very common way to implement it is with the use of capacitances, which can store energy and then dissipate it to set up the oscillations. In particular, we used the topology shown in Fig. 1. Two capacitors are alternatively charged to a reference voltage value V_{ref} and completely discharged. The charge is made with the use of a constant current generator I_{ref}, while the discharging is obtained by activating an nMOSFET switch Mb connected in parallel to each of the capacitors themselves in such a way to short them to ground. Two comparators compare the voltage across the two capacitors with the reference voltage; their outputs are connected to the input of an SR latch that controls the switches Ma and Mb for charging and discharging the capacitors.

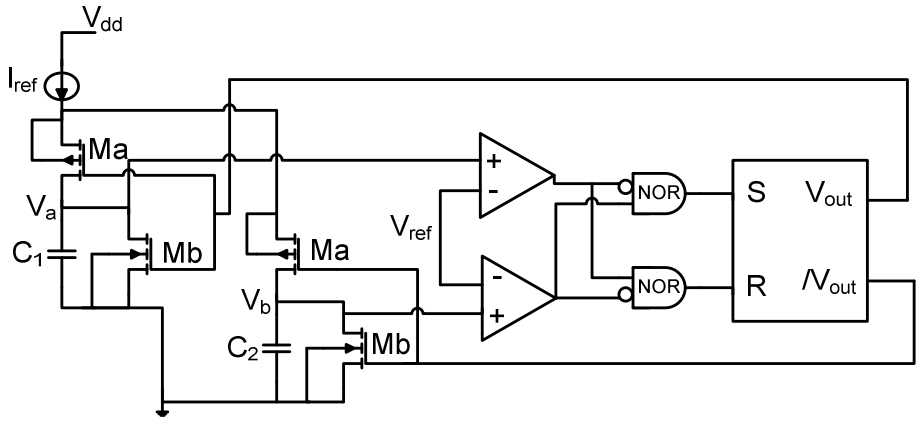

Fig. 4.1: schematic of the proposed relaxation oscillator

The two NOR gates are used in order to prevent the status S=R=1 at the input of the SR latch in the transient when the circuit is powered on. Indeed in this transient, before the start up time of the reference voltage generator, the outputs of both comparators could be high, with the consequent bias of the SR latch in a metastable state and its incorrect behavior.

The oscillator operation is as follows: when voltage V_a is larger than V_{ref}, the latch is set and V_{out} goes high, thus turning on the MOSFET Mb that discharges capacitor C_1, while / V_{out} is low and this enables the charge of C_2, until V_b reaches V_{ref}. When this condition occurs, the latch is reset, V_{out} goes low, thus enabling the charge of C1, and / V_{out} goes high, with the consequent discharge of C_2.

The oscillation frequency can be determined starting from the relation for the charge and discharge of a capacitance C:

$$I_{ref} = C\frac{\partial V}{\partial t}. \tag{4.1}$$

If we apply this relation to the circuit of Fig. 4.1, neglecting the effect of mismatch and the delay introduced by comparators and latch, we obtain the following expression for the oscillation frequency (considering C_1=C_2=C):

$$f = \frac{I_{ref}}{2CV_{ref}}. \tag{4.2}$$

From this expression we can note that if I_{ref} and V_{ref} are stable with respect to process variations, therefore the spread of the frequency depends on variations of C. In order to implement C we consider the capacitance between gate and source-drain (which are connected together to ground) of two native nMOSFET transistors. We use transistors instead of integrated capacitors in order to reduce the area occupation and in order to reduce the sensitivity of capacitance to the process (the gate oxide capacitance has a lower process sensitivity with respect to integrated MIM capacitance: for example in the used UMC 0.18 µm technology the native MOSFETs Cox variation between fast and slow corners is of about 12%, while MIM capacitance variation between the two corners is of about 30%). Moreover we use native transistors to have a stable capacitance for all the voltage drops (between zero and V_{ref}), across the capacitor itself: in fact a native transistor is biased in the same region (saturation) for all these voltage drops, so its capacitance is stable during all the oscillation period.
We can note that in order to have a stable frequency, it is important to have a precise ratio of the reference current to the reference voltage, while the individual precision of each of the two quantities is not necessary. In this case, however, we could not compensate the temperature and process variations of R, if we did not use external or trimmable components. This is why we choose to start from two reference quantities, that are intrinsically robust towards process in such a way as to obtain also a stable ratio and with some degree of freedom to adjust temperature sensitivity.
We must also consider that the relaxation architecture can be implemented with two topologies:

- by using only a capacitor which is charged and discharged between two reference voltage values;
- by using two capacitors which are alternatively charged to a reference voltage value and discharged to ground (this is the implemented topology).

In both topologies we must consider the delay introduced by the comparator with the upper threshold voltage, which can be an important source of inaccuracy for the oscillation frequency. Furthermore, if we use the first topology we must also consider the delay of the comparator with the lower threshold voltage (which is not present in the second topology). In the second topology, instead, we must consider

the mismatch between the two capacitances, which adds to the frequency process variability. If we can assume a 1% mismatch between the two capacitances, its effect on the oscillation period is half the effect of the delay of the lower threshold comparator, used in the first topology. This delay has been approximately evaluated considering the typical comparator output capacitance and resistance and considering the time constant RC. Moreover if we use the first topology, we must have two threshold voltages available. This can be a problem if, for example, we use the proposed voltage generator, which has a reference voltage smaller than 250 mV. In fact low reference voltages pose problems to the design of comparator, which is slower and less accurate. Finally the second topology is more symmetric, because we have two nominally identical capacitances and comparators.

4.3 Design of the oscillator blocks

4.3.1 Voltage reference

The reference voltage is obtained with the circuit previously described in Chapter 2, based on the use of a bipolar bandgap topology, with a careful design in order to obtain a good trade off between power consumption and process sensitivity of the reference voltage. In particular, from simulation, we obtain a reference voltage of 241.7 mV at 1 V, with a power consumption of 68.3 nW, a temperature sensitivity of 66.9 ppm/°C (from 0 °C to 100 °C) a line sensitivity of 0.25 %/V varying line voltage between 0.8 V and 1.4 V, and an area occupation of 0.28 mm^2.

4.3.2 Current reference

The reference current is obtained with the circuit previously described in Chapter 3, based on the use of devices which are intrinsically more stable towards process, such as bipolar transistor in a "classical" bandgap topology and diffusion resistors for the transconductive factor. The reference current value is 55 nA at 1 V, with a power consumption of less than 290 nW, a temperature sensitivity of 63 ppm/°C (from 0 °C to 100 °C), a line sensitivity of 0.21 %/V (varying line voltage between 0.8 V and 1.4 V) and an area occupation of 0.245 mm^2.

4.3.3 Comparators

Comparators are based on a very simple topology [122] which is shown in Fig. 4.2. This topology reduces to minimum the number of stacked transistors between V_{dd} and ground. The reference voltage is very low (about 242 mV), so MOSFETs M1 and M2 are biased in subthreshold region, which helps reduce the total power consumption (when both comparator input voltages are equal to V_{ref}, the current consumption for each comparator branch is about 105 nA in nominal conditions).
The comparator operation is as follows: if V+ is larger than V-, the current mirrored by M4 to M3 is larger than the quiescent current of M1 and M3, so the output node voltage rises in order to reduce the mirrored current. On the contrary, if V+ is smaller than V-, the current mirrored by M4 to M3 is lower than the quiescent current of M1 and M3 and so the output node voltage decreases in order to increase the mirrored current. We choose M1 equal to M2, and a consequent unitary current mirror ratio between M3 and M4. We also consider $V-=V_{ref}$ and $V+=V_{cap}=V_{in}$, where V_{cap} is the voltage drop across capacitor C.

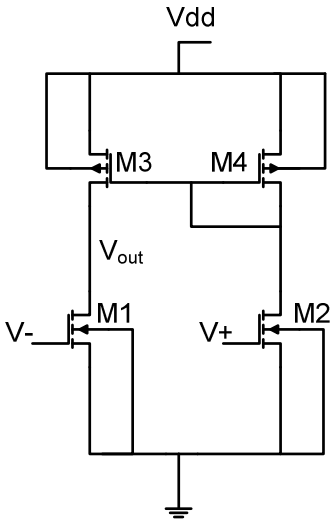

Fig. 4.2: schematic of the proposed comparator

The design of this block has been made in such a way as to obtain fast transitions of the output node voltage V_{out}, because this implies a smaller effect of the comparator delay on the oscillation period. For this purpose we must take into account the trade off between the current charging the output node (which is increased by increasing MOSFETs *W*) and the parasitic capacitance associated with output node, which also increases with larger MOSFET *W*. A small *L*, instead, is useful both for the increase of the charge/discharge current and for the reduction of the load capacitance. However, it is important to choose a larger-than-minimum length for pMOSFETs M3 and M4, in such a way as to reduce the channel length modulation effect. On the basis of these considerations, and on the basis of parametric simulations in order to obtain the optimum sizes for speed requirements, we choose the MOSFETs sizes summarized in Table I.

MOSFET parameter	Value (μm)
W_p	2
L_p	10
W_n	15
L_n	5

Table I: comparators MOSFETs sizes

With these sizes and considering a comparator bias point with both inputs equal to the reference voltage, we obtain an output capacitance of 5.83 fF and an output resistance of 71.12 MΩ.
Fig. 4.3 shows the typical waveforms for the comparator inputs and output when capacitor voltage linearly increases and exceeds the reference voltage (we consider a constant charging current equal to I_{ref}). The black line represents the reference voltage, the red line represents the capacitor voltage and the blue line represents the comparator output voltage.

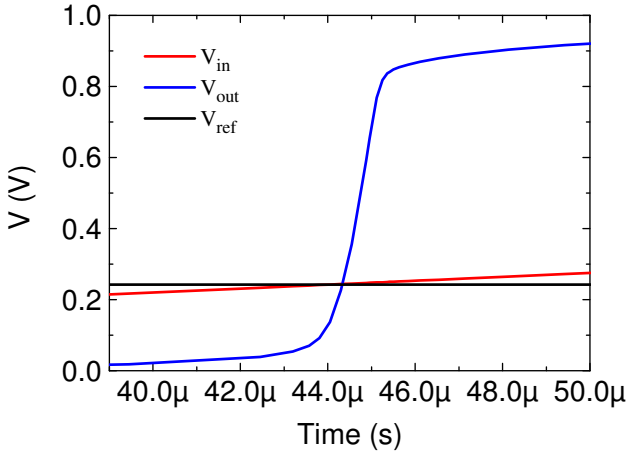

Fig. 3: comparator input and output waveforms

The typical delay introduced by the comparator and its variations considering MOSFETs process corners are shown in Table II.

Process Corner	Comparator Delay
Typical	0.9 µs
Fast	0.6 µs
Slow	1.4 µs

Table II: delay introduced by comparator

The comparator delay variation with corners is very large (nearly 90%), so it is important to obtain a small nominal delay with respect to the oscillation period. In this case the impact of the comparator delay on the oscillation period is of 2% in nominal conditions, and it varies between 1.36% and 3.2% considering fast and slow corners respectively. The time difference between fast and slow corner is 0.8 µs, which is quite negligible with respect to the nominal semi-period of 44 µs: its impact on the oscillation frequency is of 1.8%.

4.3.3.1 Effect of comparators offset

If we consider an offset V_{os} at the two comparators' input, the relative error for the oscillation period, ΔT/T, is equal to:

$$\frac{\Delta T}{T} = \frac{V_{io}}{V_{ref}}. \tag{4.3}$$

This can be reduced by reducing in turn the input offset voltage of the two comparators (or by increasing the reference voltage V_{ref}). We must however consider that the use of large sizes in order to reduce the offset voltage implies a larger capacitance associated with the node and so a slower commutation, which

also affects the oscillation period. With the chosen sizes, we obtain a standard deviation for the nMOSFETs threshold voltage of 1.08 mV, based on the empirical relation provided by the UMC 0.18 µm design kit. This implies a relative standard deviation of the oscillation period of 0.45%.
If we consider also a mismatch ΔC between the two capacitors C_1 and C_2 and if we consider a mismatch ΔI between the two currents charging the two capacitances, the relative error for the oscillation period can be approximated as:

$$\frac{\Delta T}{T} \approx \frac{1}{2}\left(\frac{\Delta C}{C} + \frac{V_{io}}{V_{ref}} - \frac{\Delta I}{I}\right). \tag{4.4}$$

This relation implies that it is important to reduce the mismatch between capacitances (with the use of a common centroid layout) and between charging currents. This is obtained by connecting the current generator output branch to both capacitances. In the literature, in order to reduce the effect of comparators offset, some offset-cancellation schemes have been proposed, for example the self-clocked scheme proposed by [124], applied to the design of a relaxation oscillator with a few tenth of µW power consumption.

4.3.4 Latch SR

The SR latch has the classical architecture shown in Fig. 4.4.

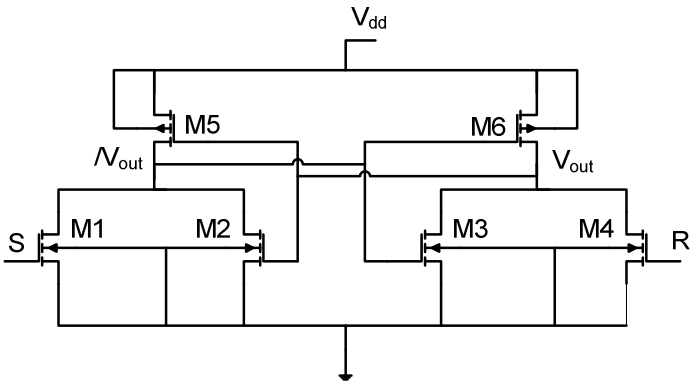

Fig. 4.4: SR latch

All transistors are of minimum size, except for M1 and M4, which are critical in order to have a fast output node transient and therefore require a larger beta. This must be obtained without increasing too much area in order not to increase too much the parasitic capacitance associated with the output node itself. We choose $W_{1,4}$=20 µm and $L_{1,4}$=1.45 µm on the basis of parametric simulations. We also need the logic gates indicated in Fig. 1 because at the power on there is a start up time for the reference voltage generator and for the reference current generator to reach the static condition. Before this condition is reached, the inputs of the SR latch could be both high, which is not allowed, so the two NOR gates prevent the input voltages of the latch to be simultaneously high when the output of the two

comparators are both high (for example because the reference voltage has not reached the static value). The NOR gates and the inverters are CMOS and their MOSFETs are of minimum size.
The typical waveforms for the set transient are shown in Fig. 4.5.

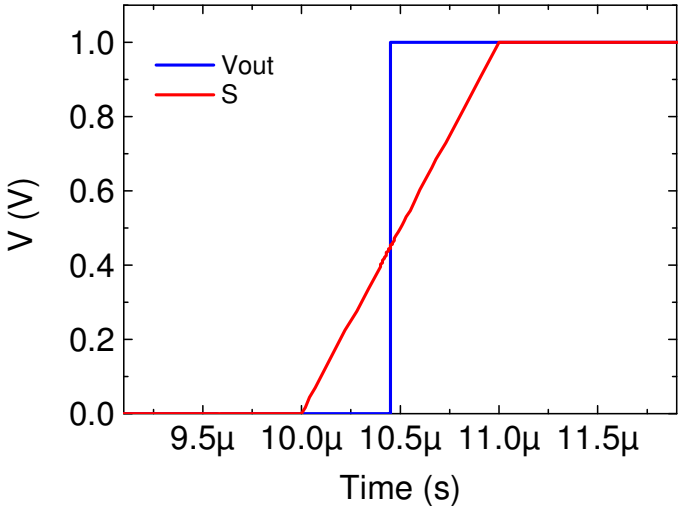

Fig. 4.5: set transient for the SR latch

4.3.5 Switches and capacitors

The two switches are constituted by two MOSFETs, a pMOSFET and a nMOSFET, as shown in Fig. 4.6.

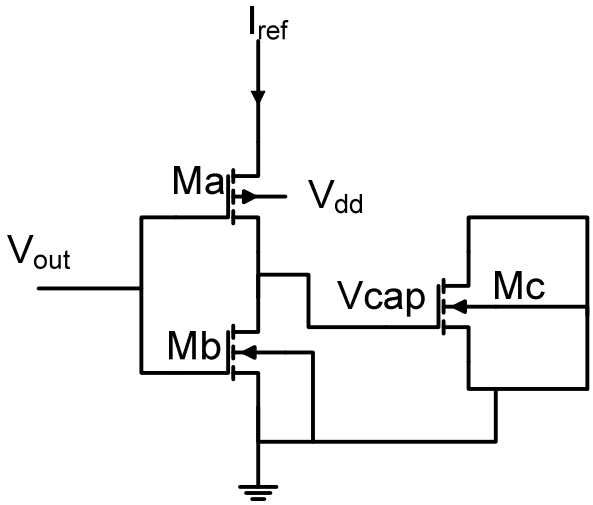

Fig. 4.6: schematic of the switches

The source of the two pMOSFETs Ma is connected to the output of the same current generator, because alternatively only one of the two is shorted and it allows the bias current to flow through the gate-substrate capacitance of only one native MOSFET Mc. When MOSFET Mb is active, it shorts the capacitance voltage to ground. The native MOSFET Mc is constituted by three transistors in parallel with W=L=20 µm: the equivalent gate-substrate capacitance is 10 pF. Transistors Ma have W=1 µm and L=20 µm, while transistors Mb have W=2 µm and L=190 nm. It is important for the discharging transient to be faster than the charging transient, this is why we chose Mb length lower than Ma length.

4.4 Simulation results

The proposed oscillator has been designed in a UMC 0.18 µm CMOS technology and simulations have been performed with CADENCE Spectre. We obtain a nominal oscillation frequency of 10.11 kHz at 1 V, with a power consumption of about 318 nW at 1 V. The temperature sensitivity of the reference frequency is 144 ppm/°C in a temperature range from 0°C to 100 °C, while the line sensitivity is 1.35 %/V by varying line voltage from 0.8 V to 1.4 V. The area occupation is of about 0.5 mm^2. The circuit layout is shown in Fig. 4.7.

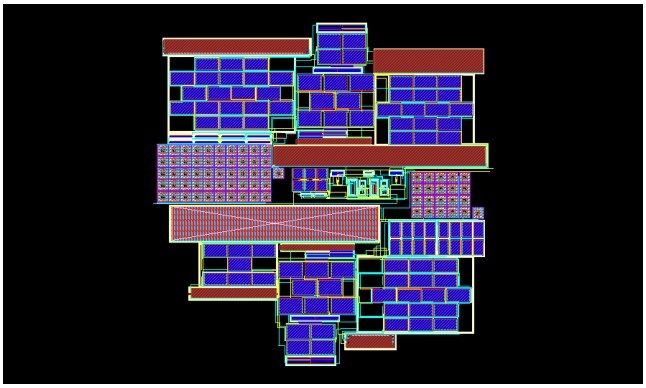

Fig. 4.7: oscillator layout

The obtained results are compared with the ones of other relaxation oscillators proposed in the literature with an oscillation frequency lower than 1 MHz (see Table III).

Ref.	Proposed work	[125]	[123]	[122]	[113]
Techn. (CMOS)	0.18 µm	0.35 µm	0.35 µm	0.13 µm	65 nm
Freq.	10.11 kHz	50 kHz	3.3 kHz	1.52 MHz	100 kHz
V_{dd} (V)	1	0.8	1	0.8	1.2
Power	318 nW	616 nW	11 nW	320 nW	41 µW
TC (ppm/°C) (°C)	144 (0÷100)	3000	500 (-20÷80)		
LS (%/V)	1.35	-	3.5	2.5	0.37
Process sens. (σ/µ)	1.8%	10%	6.9%		2.3%
Area (mm^2)	0.5	0.24	0.1	0.0134	0.11

Table III: comparison between the proposed oscillator performance and results from literature

Ref [123] obtains a very low power relaxation oscillator: it uses only one capacitor and one comparator in order to reduce power consumption and it employs also a digital section in order to obtain the 50% duty cycle clock signal. It tries to obtain a low process sensitivity of the oscillation frequency with the use of a current source which contains in its expression the same type of capacitor which is charged and discharged, and with the use of a PTAT process insensitive reference voltage, in such a way that the oscillation frequency process sensitivity is only related to MOSFET mobility process variability. The main drawback of this solution is the high temperature sensitivity. Ref [122] uses the same architecture of the proposed oscillator, based on the alternative charging and discharging of two capacitors between ground and a reference voltage. In this case, however, the reference voltage is proportional to the reference current through a factor R, in such a way that the oscillation frequency is only dependent on the product RC. This choice allows to reduce power consumption and to significantly reduce area occupation, but the process variability is not compensated and can be adjusted, until reaching the desired performance, with the use of trimming, which can be expensive. Ref [113], already discussed, proposes an oscillation frequency based on intrinsically more stable devices, in particular MOSFET mobility, which has a standard deviation smaller than 2% in the used process. Authors proposed a current-controlled relaxation oscillator, in which the current is proportional to mobility. However the current consumption of the proposed architecture is very high, of the order of some tenth of µW. Moreover the temperature sensitivity of the clock frequency is high. The proposed oscillator (on the basis of simulation results), has the lowest process sensitivity of the clock frequency because it is based on devices which are intrinsically more stable than standard MOSFETs and which however are available also in standard CMOS technologies.
Similar results in terms of process variability of the oscillation frequency have also been obtained by considering the design approach proposed in [55], based on the

generation of process insensitive reference currents based on the sum of two reverse correlated currents. These currents stable towards process have been used in the design of a process-insensitive 3 GHz ring oscillator [68], [129], which obtains a 2.7% process variation but with a power consumption greater than 1 mW, because it uses a control loop for the oscillation frequency similar to the one of a PLL.
In conclusion the proposed oscillator provides the best trade off between power consumption and process sensitivity of the oscillation frequency, as indicated by simulations. A silicon batch of the proposed circuit is under fabrication.

5. A METHOD FOR THE ANALYSIS OF THE NUMBERS AND STABILITY OF NONLINEAR LOOP CIRCUITS OPERATING POINTS

5.1 Introduction

If we consider electronic circuits containing active devices, the determination of the operating point is the first and fundamental step of the design process. A non-linear time-independent circuit (i.e., without capacitors and inductors) can be described with a system of equations $F(x)=0$, where the vector x is composed by node voltages and/or branch currents. The system can have an unknown number of solutions x_i. This problem usually requires the solution of an inherently non-linear physical system, and, since non-linear systems cannot generally be solved in closed form, the electronic designer must approximate solutions with numerical analysis tools or, sometimes, clever ad hoc tricks. In fact, this intrinsic non-linearity can be a problem, since most circuits are designed to have an operating point that can be easily determined.

However, some applications demand the use of circuits for which the computation of the operating point is non trivial. Moreover, many circuits have only one solution, but circuits with more than one solution are well known. Eccles-Jordan flip flops are positive feedback circuits which generally have three solutions, one of which is "unstable". We must note that even the "stability" of the solution is not a well-defined concept. Solutions of time-independent circuits cannot be "stable" or "unstable". Indeed, unstable solution are not solutions at all. A formal definition of "stable solution" can be found in [130]: a solution of $F(x)=0$ is potentially stable if it is possible to build (adding capacitors between nodes and inductors in series to the branches of the given circuit) an augmented circuit which is robustly stable in the time domain. Robustly stable means that the stability is not compromised by the addition of another set of sufficiently small capacitors and inductors to the given circuits (i.e. the values of the first set of capacitors and inductors must not be critical). Solutions which are not potentially stable are unstable. In most practical cases, the stability of a solution can be assessed by examining the impedances seen from circuit nodes to ground. If they are positive, the examined solution is stable (with regard to the above definition, the solution is potentially stable with a set of null capacitors and inductors). The analysis of these positive feedback non linear circuits can be challenging; furthermore, commonly used circuit simulators may provide unreliable information, since they can converge to an "unstable" solution.

General methods have been developed for the non-linear analysis of active circuits [131], but they are generally too abstract, they do not provide information about the circuit operation, and so they are not useful for the circuit designer. Such methods can be implemented in a circuit simulator [132] [133], but are not widely used. As a consequence, non-linear circuits are often analyzed with simple pencil and paper methods [134]. These calculations are performed with crude first-level device

models, which can lead to grossly approximated solutions, missed solutions and sometimes even to spurious solutions. Another common way to investigate the stability properties of circuits is with the use of transient simulations, but these are time consuming and can also provide unreliable information in case of circuits with widely separated time constants. In order to overcome these shortcomings, we propose a method [135] that is able to find the operating points and the stability properties of many commonly used non-linear feedback circuits and which was also applied to the designed generators in order to verify their correct design in terms of stability. We will show that this technique is also able to determine if a start up circuit is necessary, and we will show that it allows us to obtain results with high accuracy and easiness because it makes use of the standard EDA tools used for the circuit design. This also implies that accurate device models provided by design kits are available.
The proposed method was successfully applied also for the design of a self-biased current generator [136], for which MOSFETs size are critical for the determination and the stability of the circuit operating point. This circuit can be represented in a simplified view as indicated in Fig.5.1. In practical implementations, the reference voltages can be replaced by resistors [137] or triode MOSFETs [138]. Referring to Fig. 5.1, the self-biased current generator is composed by two current-controlled current generators connected back-to-back in a positive-feedback loop. Transistors M3 and M4 form a linear current mirror, duplicating the drain current of M4 in M3. This current mirror provides a linear relationship between its input and output, which can be indicated by relation (5.1):

$$I_{out\_um} = k_{um} I_{in\_um}, \tag{5.1}$$

where k_{um} is a linear proportionality constant which depends on the geometry of M3 and M4.

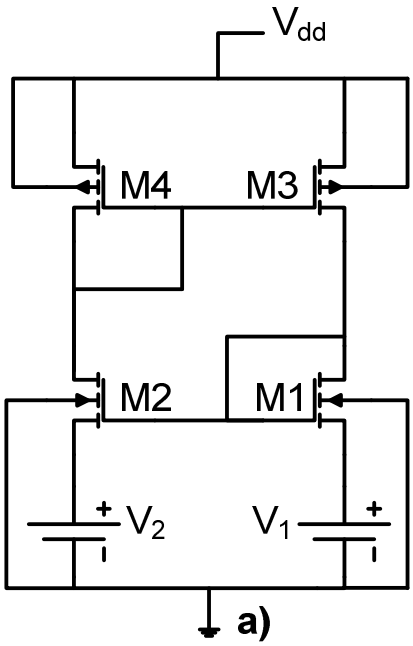

Fig.5.1: self biased current generator

The lower mirror, instead, (which is composed by transistors M1 M2, and by voltage sources V_1 and V_2) provides a nonlinear relationship between the input current and the output, which can be represented by relation (2) where f is a non linear function of its input I_{in_um}:

$$I_{out\_um} = f\left(I_{in\_um}\right) \tag{5.2}$$

This implies that he ratio of the input to the output current k_{im} depends on the input current. At equilibrium, however, we must have:

$$k_{um} = 1/k_{im} \tag{5.3}$$

If k_{im} is a monotonic function of the input, relation (5.3) can be satisfied for a single set of currents. However, as [134] points out, both mirrors of the circuit provide zero current when fed with a zero input and hence another equilibrium point exists, with all null currents (where k_{im} is undefined). For this reason most designers of self-biased current generators include a startup circuit which forces the circuit to the desired non-zero solution.

However, the above discussion is oversimplified. We simulate the circuit with a UMC 0.18µm CMOS technology, and with identically sized M3 and M4: we indicate with $\beta_i=\mu C_{OX}W_i/L_i$, where W_i and L_i are respectively transistors width and length, µ is the carrier mobility and C_{OX} is the gate oxide capacitance per unit area, the beta of transistors M_i. We find that if $\beta_1>\beta_2$ and $V_1>V_2$, the circuit always settles to non-zero currents. Hence, no startup circuit seems to be required. Instead, if $\beta_1<\beta_2$ and

$V_1<V_2$ the circuit never settles in the equilibrium point suggested by relation (3), and no startup circuit helps. For the other possible configurations, relation (3) is never verified and no equilibrium point exists. This simple circuit shows that MOSFETs size are really important for the position and the stability of the circuit operating point, and that a simple but effective method for the study of this problem is very important for the circuit designer.

5.2 The proposed method for the determination of the number and stability of circuit operating points

To solve this problem we developed a technique that provides valuable information on the equilibrium points of nonlinear circuit. We describe this technique in a general way, and then we will show the provided results in the case of the self-biased current generator.

We consider a nonlinear circuit as a closed loop of nonlinear blocks, as shown in Fig. 5.2a. Starting from this simplified view, we can cut open the loop and insert the circuitry shown in Fig. 5.2b. The independent current source sends a test current I_t in the circuit which gives rise to a voltage V_p across its terminals. The voltage-controlled generator imposes the same voltage V_p to node B, the other end of the cut loop. Obviously, when the current I_v sinked by the voltage generator is equal to I_t, the circuit is in equilibrium. The two sides of the cut could be directly connected without altering the branch currents and the node voltages. Hence if we plot I_v vs. I_t, equilibrium points can be identified as the intersections between the $I_v(I_t)$ curve and the $I_v=I_t$ line. In addition, the derivative $\partial I_v / \partial I_t = \lambda$ at the equilibrium point enables us to determine the stability of the equilibrium point.

Let us call R_t the differential resistance seen by the I_t generator: if the test current increases by ΔI_t, the voltage V_p increases by $\Delta V_p= R_t \Delta I_t$. The current I_v, instead, increases by $\Delta I_v = \lambda \Delta I_t$. Since the nodes A and B are at the same voltage, we connect them and redraw the circuit as in Fig. 2c. The total differential resistance seen between nodes A=B and ground (as shown in Fig. 5.2c) can be written as:

$$R_d = \frac{\Delta V_p}{\Delta I_{tot}} = \frac{R_t \Delta I_t}{\Delta I_t - \lambda \Delta I_t} = \frac{R_t}{1-\lambda} \tag{5.4}$$

where ΔI_{tot} is indicated in Fig. 5.2c.

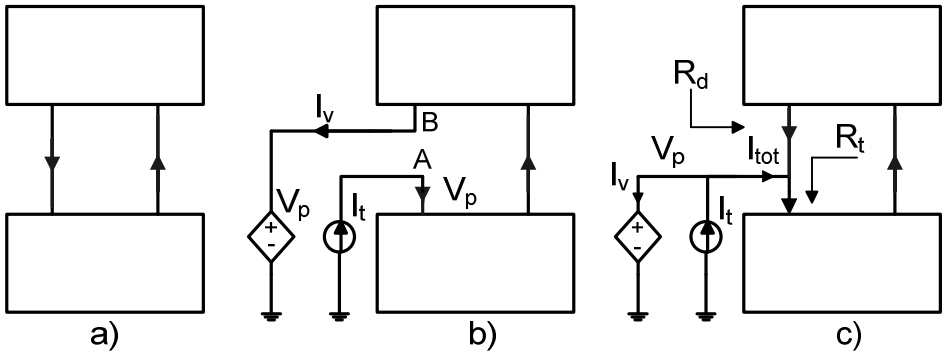

Fig.5.2: principle of the proposed method

Considering relation (5.4) we can conclude that, if $\lambda>1$, the solution is unstable. Let us underline that we assumed $R_t>0$, as is typical in practical circuits, but the method can be generalized to any initial sign of R_t.
Furthermore, λ is the small-signal DC loop gain, and hence the fact that values in excess of 1 lead to instability is well known.
Hence, the practical application of the method consists of cutting open a loop, inserting the proper generators and performing a DC simulation of the circuit with an input current sweep. This implies that the proposed method is very easy and familiar for the circuit designer, and also accurate, because it is based on the accurate devices models used for the circuit simulation.
This method was successfully applied to the circuit of Fig. 5.1. We made a cut on the drain of M3 and we inserted the current generator on the gate drain of M1 and the voltage generator on the M3 drain. We then plot the current of the current generator and the current which flows in the voltage generator. The graph of Fig. 5.3 and Fig. 5.4 are related to the case of $\beta_1>\beta_2$ and $V_1>V_2$.The black line represent the current of the current generator, while the red line represents the current of the voltage generator. These results show that a single and stable operating point is obtained. It is worth noticing that in this case no equilibrium point exist at $I_t=0$ (it is clearly evident from Fig. 5.4) and hence no startup circuitry is needed.

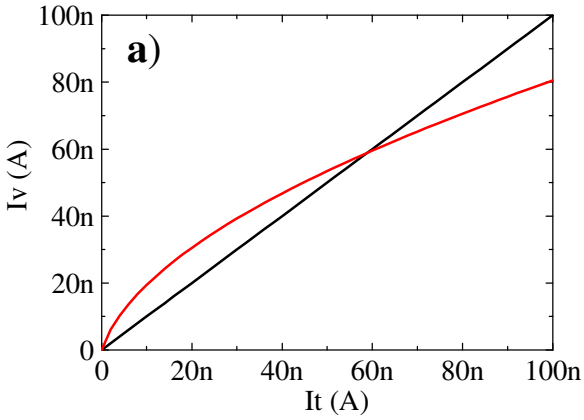

Fig.5.3: SPECTRE dc sweep of circuit of Fig. 5.1 in the case $\beta_1>\beta_2$ and $V_1>V_2$

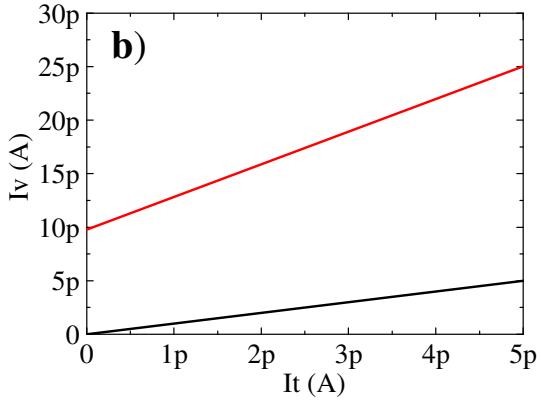

Fig.5.4: particular of the SPECTRE dc sweep of circuit of Fig .5.1 in the case $\beta_1>\beta_2$ and $V_1>V_2$

If we consider, instead, the case $\beta_1<\beta_2$ and $V_1<V_2$, we obtain the results shown in Fig. 5.5 and Fig. 5.6, which indicates that the desired solution is unstable, and another stable solution is present for very small currents (this solution is evident from Fig. 5.6).

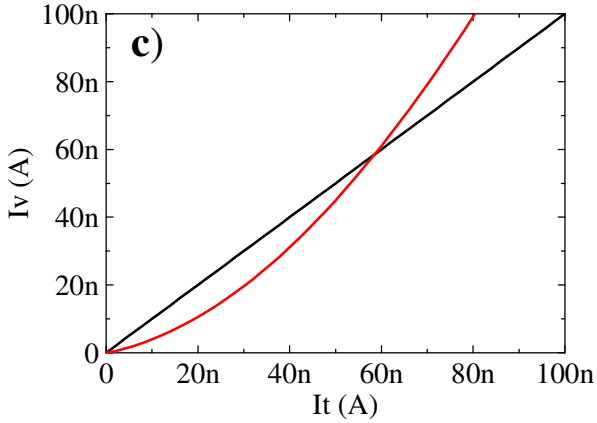

Fig.5. 5: SPECTRE dc sweep of circuit of Fig.5.1 in the case $\beta_1<\beta_2$ and $V_1<V_2$

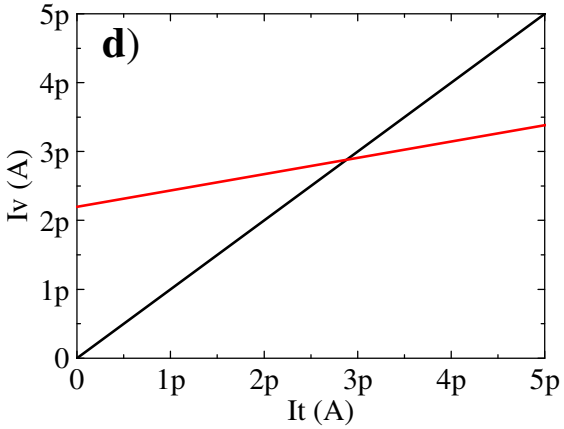

Fig.5.6: particular of the SPECTRE dc sweep of circuit of Fig.5.1 in the case $\beta_1>\beta_2$ and $V_1>V_2$

Therefore, with the use of a circuit simulator equipped with accurate device models we can learn that often some pencil-and-paper results, such as the zero-current stable solution, can indeed be artifacts due to the use of too simplistic device models. Furthermore, differently from other methods based on the insertion of test generators [139], this approach provides valuable physical insights on the circuit. Since the $I_v(I_t)$ relationship provided by the simulations can be interpreted as the input-output characteristic of an amplifier, a designer can usually devise modifications to the circuit which can modify it in a foreseeable manner.
Hence, the above analysis not only can provide evidence of bias or stability problems, but is also a valuable tool for their solution.

This method was successfully applied also for the design of the reference voltage generator, the reference current generator and the low frequency oscillator. It allows us to verify the stability of the desired operating points and the absence of other operating points (in particular we verify that the zero currents operating point is not a solution of our circuits). This is why we do not insert start up circuits in our generators. This method was also applied to the design of a self biased native MOSFETs based current generator, whose design will be described in more details in the next paragraph.

5.3 Design of a self-biased native-MOSFETs based current generator

The current generator has been designed based on the circuit schematic of Fig. 5.1, where we consider V_2=0 and where we consider V_1 as a reference voltage stable with power supply, temperature and process variations, obtained with the principle shown in Chapter 2. With these considerations, we can write:

$$V_1 + V_{GS1} = V_{GS2} \tag{5.5}$$

Neglecting the difference between the threshold voltages of M1 and M2, if we consider M1 and M2 biased in saturation and we consider a current mirror ratio α between M3 and M4 (so $I_1 = \alpha I_2$), we obtain:

$$I_2 = \frac{\beta_2 V_1^2}{2\left(1 - \sqrt{\frac{\alpha\beta_2}{\beta_1}}\right)^2} \tag{5.6}$$

From this expression we can note that the major source of variability for the current is the variability of MOSFET beta, which is mainly related to the variability of carriers mobility and oxide thickness. Applying our approach for the reduction of process variability we considered the beta of native MOSFET instead of the one of standard MOSFETs, because we thought that the mobility of native transistors should have a lower process sensitivity due to the lower doping of native transistors with respect to the standard ones. In reality, we already showed in Chapter 3 that the mobility of native transistors is not less process sensitive than the one of standard transistors. Moreover, we verified from experimental results that the native MOSFETs models are not accurate and so the measured operating point is different from the simulated point. We characterize some native MOSFETs and we correct their models, in such a way to obtain results more adherent to measurements. Even if the design of this block presents these problems, it will be presented as an example of a native MOSFETs self-biased generator design and as an application example of the proposed method. In fact the use of native MOSFETs for the implementation of M1 and M2 poses some problems related to the bias circuit, because the diode connection for M1 implies that it is not biased in the saturation region, but in the linear region, so the scheme of Fig. 5.1 can not be used if we want M1 and M2 to be native MOSFETs biased in saturation. The circuit

has so been modified as in Fig. 5.7, in which the circuit M1 was not diode-connected and a proper bias circuit (M6 M7 M8) was added in order to bias M1 in saturation.

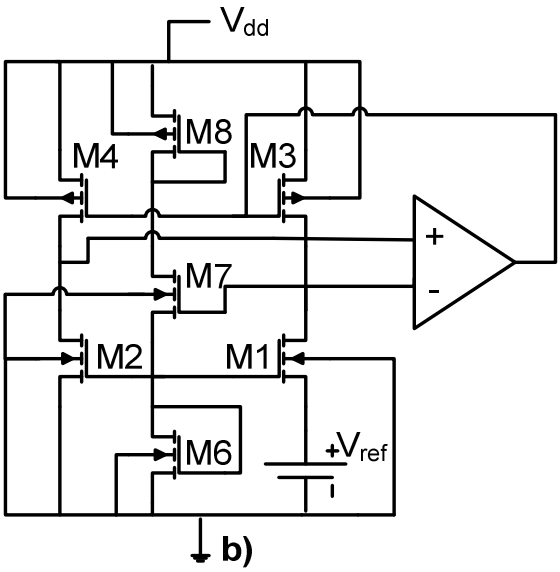

Fig.5.7: proposed current generator with native MOSFETs

We consider the relation which must be satisfied by M1 in order to be in saturation:

$$V_{DS1} > V_{GS1} - V_{th1} \tag{5.7}$$

where V_{th1} is equal to -190 mV in the nominal condition (as indicated by the provided design kit) and varies between -90 mV and -290 mV by considering fast and slow corners for native transistors. While:

$$V_{D1} = V_{G1} + V_{GS7} \tag{5.8}$$

the saturation condition for M1 is verified if

$$V_{GS7} > |V_{th1}|. \tag{5.9}$$

In order for this relation to be verified in all corners, we choose V_{GS7} equal to 370 mV. Transistors M6 and M8 are necessary for the biasing of M7, so they provide its current in such a way that the gate to source voltage of M7 reaches the desired value. In order to reduce power consumption, transistor M6 is replaced by five diode connected native diode-connected transistors in series, in such a way that the total gate-to-source voltage of M1 is distributed to the five gate-to-source voltages of the five transistors.

For what concerns the choice of M1 and M2 sizes, for both we choose a large L in order to reduce power consumption and the effect of mismatch between the two transistors. This is obtained by connecting a certain number (in this case, 30) of transistors of the maximum length (50 μm) in series, with common gate terminals. For what concerns the choice of M1 and M2 width W, we choose small values in order to reduce power consumption and also in such a way to reduce process variability of the reference current. In fact, considering relation (6), the use of small W implies a dependence of the relative beta process sensitivity to width W, so by properly choosing W_1 and W_2 (with the aid of simulations) we can obtain a partial compensation effect of the numerator β_2 process variability with the beta ratio $\alpha\beta_2 / \beta_1$ process variability. In particular we made simulations considering various values of α and W_1 and finding the W_2 value which minimize process sensitivity of the reference current (by considering corners). We choose α=16, W_1=4.6 μm e W_2= 1 μm. A larger α "amplifies" the effect of beta ratio compensation on the numerator beta variability, but it also implies a larger power consumption. The current mirror MOSFETs are constituted by MOSFETs of W=30 μm and L=50 μm, with the proper ratio (so M4 is constituted by 16 transistors like M3 connected in parallel). In order to reduce current mirror mismatch, MOSFETs have large sizes and we also add source degeneration resistors of 4.746 MΩ for M3 and 0.297 MΩ for M4. The reference current is mirrored to an output pMOSFET M5 with a current mirror ratio 1:4 between M3 and M5.
We choose a high value of V_1, V_1=335 mV, in order to reduce the effect of mismatch between M1 and M2, which directly adds to voltage V_1.
The voltage V_{ref} is imposed on a standard nMOSFET with W=30 μm and L=49 μm with the use of a single stage operational amplifier.
We applied the proposed method to the designed current generator, with the modified native MOSFETs models on the basis of an experimental characterization, obtaining the results of Fig. 5.8. The current in M1 is about 7 nA, and the operating point is stable. This was confirmed by measurements on 20 samples realized in a 0.18 um UMC CMOS technology, from which we obtained a mean current value of 6 nA and no problem of start up.

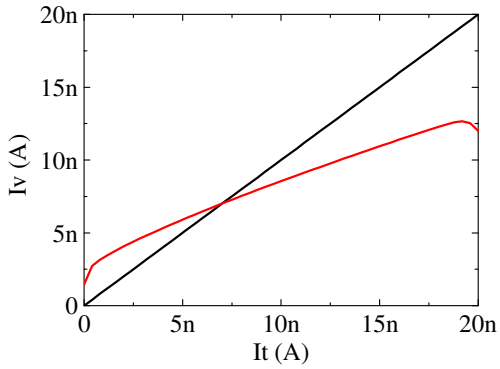

Fig. 5.8: I_v versus I_t for the complete circuit of Fig. 5.1

CONCLUSION

This work proposes a series of solutions for the problem of process variability, whose effectiveness is especially related to analog basic building blocks for very low power applications such as passive RFID transponders and implantable medical devices.
Very precise reference quantities, such as voltages, currents, or frequencies, are important because they are used in a broad variety of analog and digital systems and they constitute the basis for the implementation of more complex feedback process compensation systems. The coupling of very low process sensitivity together with very low power consumption is an essential need for the mentioned applications, where the standard solutions to the process variability problem (trimming or compensation loop with external components) can become very expensive and not feasible.
After a brief description of the techniques proposed in the literature for the reduction of process variability, both in digital and in analog systems, we focused on our "variability-aware" approach, whose effectiveness has been proven by experimental results on two reference generators.
This approach is related to the use of devices which are "intrinsically" more stable towards process and which are available also in standard CMOS technologies. It is extremely useful especially for the design of low process sensitive, low power analog basic building blocks such as reference generators.
We applied this approach to the design of a 70 nW, 0.18% reference voltage generator based on the use of a "classical" bandgap architecture with the use of substrate bipolar transistors. We made an accurate statistical characterization of this generator by considering 20 samples from a single batch in a UMC 0.18 µm technology. We also obtain statistical results related to inter-batch variability (a 0.35% cumulative relative standard deviation of the reference voltage by considering two batches), even if on a slightly different voltage generator, with a larger temperature sensitivity. This result confirms the effectiveness of the proposed approach also for the reduction of inter-batch variability, because it is based on the use of "intrinsically" robust quantities with respect to process variations.
We have also shown that the good trade-off between process sensitivity of the reference voltage and low power consumption is not achievable with conventional solutions such as trimming, which has a large power consumption, and which implies the presence of a digital section, which can be very expensive especially considering very simple analog systems like passive RFID transponders and implantable medical devices.
We have also proposed the design of a current generator based on this approach in a UMC 0.18 µm CMOS technology, showing the difficulty in the design of a stable reference current because of the presence of a transconductive factor very sensitive to process variations. We showed that the use of a bipolar bandgap architecture and of diffusion resistors is a very effective solution in order to obtain a current with a small process sensitivity (0.14% based on measurements on 20 samples from a single batch) with a very low power consumption (less than 300 nW).
The proposed current and voltage generators are then used in the implementation of a low-frequency, low-power oscillator with a very stable frequency towards

process. We do not have experimental results related to this oscillator because it is currently under fabrication, but simulation results show a very low power consumption of about 300 nW with a process sensitivity of the oscillation frequency of 1.8%.

The main drawback of generators based on this approach is a large increase in area occupation, which however can be acceptable when one uses aggressively scaled CMOS technologies, which provide abundant margins in terms of die area.

Finally, we show the effectiveness of our proposed method for the determination of the number and stability of circuit operating points by successfully applying it in the design of our reference generators. We also used this method in order to establish that start up circuits are very often not needed in our generators, with possible savings in terms of die area and power consumption.

We can conclude that our study proposes important considerations on analog design and on the problem of process variability in analog basic building blocks.

BIBLIOGRAPHY

[1] B. Daoyle at al, "Transistor elements for 30 nm physical gate length and beyond", Intel Technol. J. vol. 6, pag. 42, 2002.

[2] K. A. Bowman, S. G. Duvall and J. D. Meindl, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration", IEEE J. Solid-State Circuits, vol. 37, no. 2, pp. 183-190, Feb. 2002.

[3] N. Drego, A. Chandrakasan and D. Boning, "An all-digital, highly scalable architecture for measurement of spatial variation in digital circuits", in Proc. IEEE ASSCC, Nov. 2008, pp. 393-396.

[4] G. Gielen et al., "Analog and digital circuit design in 65 nm CMOS: end of the road?" in proc. of DATE 2005.

[5] D. J. Frank, Y. Taur, M. Ieong and H.-S. P. Wong, "Monte Carlo modeling of threshold variation due to dopant fluctuations", in Tech. Dig. VLSI Symposium, pp. 169-171, Jun 1999.

[6] K. Nishinohara, N. Shigyo and T. Wada, "Effects of microscopic fluctuations in dopant distributions on MOSFET threshold voltage", IEEE Trans. Electron Devices, vol. 39, no. 3, pp. 634-639, March 1992.

[7] A. Asenov, S. Kaya and A. Brown, "Intrinsic parameter fluctuations in decananometer MOSFETs introduces by gate line edge roughness", IEEE Trans. Electron. Devices, vol. 50, pp. 1254-1260, May 2003.

[8] P. Oldiges et al., "Modeling line edge roughness effects in sub 100 nm gate length devices", in Proc. SISPAD, 2000, pp. 131-134.

[9] P. A. Stolk, F. P. Widdershoven and D. B. M. Klalssen, "Modeling statistical dopant fluctuations in MOS transistors", IEEE Trans. Electron Devices, vol. 45, no. 9, pp. 1960-1971, Sept. 1998.

[10] A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub 0.1 micron MOSFETs: a 3D "atomistic" simulation study", IEEE Trans., Electron Devices, vol. 45, pp. 2505-2513, Dec. 1998.

[11] A. Asenov and S. Saini, "Polysilicon gate enhancement of the random dopant induced threshold voltage fluctuations in sub 100 nm MOSFET's with tunneling oxide", IEEE Trans. Electron Devices, vol. 47, no. 12, pp. 805-812, April 2000.

[12] A. Asenov, S. Kaya and J. H. Davies, "Intrinsic threshold voltage fluctuations in decanano MOSFETs due to local oxide thickness variations", IEEE Trans. Electron. Devices, vol. 49, no. 1, pp. 112-119, Jan. 2002.

[13] Jan M. Rabey, A. P. Chandrakasan and B. Nicolić, Digital Integrated Circuits- A design Perspective (2nd Edition), Prentice Hall, 2003.

[14] B. H. Calhoun, A. Wang and A. Chandrakasan, "Modeling and sizing for minimum energy operation in subthreshold circuits", IEEE J. Solid-State Circuits, vol. 40, no.9, pp. 1778-1786, Sept. 2005.

[15] B. H. Calhoun, F. A. Honorè and A. P. Chandrakasan, "A leakage reduction methodology for distributed MTCMOS", IEEE J. Solid-State Circuits, vol. 39, no. 5, pp. 818-826, Feb. 2004.

[16] J. W. Tschanz et al, "Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and

leakage", IEEE J. Solid-State Circuits, vol.37, no. 11, pp. 1396-1402, Nov. 2002.
[17] A. AbdelHamid et al., "A standardized knobs and monitors RTL2RTL insertion methodology for fine grain SoC tuning", in IEEE Euromicro Conf. on DSD/Arch., methods and tools, 2009, pp. 401-408.
[18] H. Wang et al. "Variable tapered pareto buffer design and implementation allowing run-time configuration for low-power embedded SRAMs", IEEE Trans. Very Large Integration Systems, vol. 13, no. 10 Oct. 2005.
[19] A. Papanikolaou et al, "A system-level methodology for fully compensating process variability impact of memory organizations in periodic applications", in IEEE/ACM/IFIP Int. Conf. on Hardware/Software Codesign and System Synthesis, 2005, pp. 117-122.
[20] F. Klass, A. Jain, G. Hess and B. Park, "An all-digital on-chip process-control monitor for process-variability measurements", in Dig. Tech. Papers ISSCC, 2008, pp. 408-409, 623.
[21] A. Drake, "A distributed critical-path timing monitor for a 65 nm high-performance microprocessor", IEEE ISSCC, 2007, pp. 398-399.
[22] A. Bassi et al., "Measuring the effects of process variations on circuit performance by means of digitally-controllable ring oscillators", in Proc. Int. Conf. Micro Electronic Test Structures, 2003, pp. 214-217.
[23] C. Kim et al., "Self calibrating circuit design for variation tolerant VLSI systems", in Proc. Int. On-line Test Symp., 2005, pp. 100-105.
[24] R. Rao, K. A. Jenkins and J-J. Kim, "A local random variability detector with complete digital on-chip measurement circuitry", IEEE J. Solid-State Circuits, vol. 44, no. 9, pp. 2616-2623, Sept. 2009.
[25] T. Kuroda et al., "A 0.9 V 150 MHz, 10 mW 4 mm^2 2-D discrete cosine transform core processor with variable threshold-voltage (VT) scheme", IEEE J. Solid-State Circuits, vol. 31, pp. 1770-1779, Nov. 1996.
[26] C. H. Kim et al. "A process variation compensating technique with an on-die leakage current sensor for nanometer scale dynamic circuits", IEEE Trans. VLSI Syst., vol. 14, no. 6, pp. 646-649, Jun. 2006.
[27] S. Mukhopadhyay et al., "Self-reparing SRAM for reducing parametric failures in nanoscaled memory", in Proc. Symp. VLSI Circuits, Dig. Tech. Papers, 2006, pp. 132-133.
[28] M. Griffin et al., "A process-independent, 800-MB/s, DRAM byte-wide interface featuring command interleaving and concurrent memory operation", IEEE J. Solid-State circuits, vol. 33, no. 11, pp. 1741-1751, Nov. 1998.
[29] J. W. Tschanz, S. Narendra, R. Nair and V. De, "Effectiveness of adaptive supply voltage and body bias for reducing impact of parameter variations in low power and high performance microprocessors", IEEE J. Soli-State Circuits, vol. 38, no. 5, pp. 826-829, May 2003.
[30] M. Miyazaki, G. Ono and K. Ishibashi, "A 1.2-GIPS/W microprocessor using speed-adaptive threshold-voltage CMOS with forward bias", IEEE J. Solid-State Circuits, vol. 37, no. 2, pp. 210-217, Feb. 2002
[31] N. Drego, A. Chandrakasan and D. Boning, "All-digital circuits for measurement of spatial variation in digital circuits", IEEE J. Solid-State Circuits, vol. 45, no. 3, pp. 640-651, March 2010.

[32] M. Eirener et al., "In-situ delay characterization and local supply voltage adjustment for compensation of local parametric variations", IEEE J. Solid-State Circuits, vol. 42, no. 7, pp. 1583-1592, Jul. 2007.
[33] M. Eireiner, S. Henzler, G. Georgakos, J. Berthold and D. Schmitt-Landsiedel, "Local supply voltage adjustment for low power parametric yield increase", Proc. of ESSCIRC, 2006, pp. 98-101.
[34] A.-J- Annema, "Analog circuit performance and process scaling", IEEE Trans. Circuits Syst. II- An. and Dig. Signal Proc., vol. 46, no. 6, pp. 711-725, June 1999.
[35] V. Potdar, A. Sharif and E. Chang, "Wireless sensor networks: a survey", Proc. of Int. Conf. Advanced Inf. Networking and Appl. Workshops (WAINA), 2009, pp. 636-641.
[36] J. Yick, B. Mukherjee and D. Ghosal, "Wireless sensor network survey", Computer Networks, vol. 52, no. 12, pp. 2292-2330, Aug. 2008.
[37] I. F. Akyldiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless sensor networks: a survey", Computer Networks, vol. 38, no.4, pp. 393-422, March 2002.
[38] J. Yick, B. Mukherjee and D. Ghosal, "Analysis of a prediction-based mobility adaptive tracking algorithm", Proc. of the IEEE Int. Conf. on Broadband Networks (BROADNETS), pp. 753-760, 2005.
[39] G. Wener-Allen et al., "Deploying a wireless sensor network on an active volcano" IEEE Internet Comupting, vol. 10, no. 2, pp. 18-25, March-April 2006.
[40] M. Becker et al., "Logistic applications with wireless sensor networks", in Proc. of HotEmNets, 2010.
[41] L. Evers, P. Havinga and J. Kuper, "Flexible sensor network reprogramming for logistics", IEEE Int. Conf. on Mobile Adhoc and Sensor Systems (MASS), 2007, pp. 1-4.
[42] E. Petriu et al., "Sensor-based information appliances", IEEE Instr. and Meas. Magazine, vol. 3, no. 4, pp. 31-35, Dec. 2000.
[43] F. Boekhorst, "Ambient intelligence, the next paradigm for consumer electronics: how will it affect silicon?" Dig. of Tech. Papers Int. Solid-State Circuits Conf. (ISSCC), 2002, pp. 28-31.
[44] J. Ko et al., "Wireless sensor networks for healthcare", Proc. of the IEEE, vol. 98, no. 11, Nov. 2010.
[45] Y. Zhang et al., "Energy efficient design for body sensor nodes", J. Low power Electron. Appl. vol. 1, pp. 109-130, April 2011.
[46] A. Wong et al. "A 1V, micropower system-on-chip for vital-sign monitoring in wireless body sensor networks", in Dig. of Tech. papers IEEE Int. Solid-State Circuits Conf. (ISSCC), 2008, pp. 138, 602.
[47] A. T. Bart et al., "TEMPO 3.1: a body area sensor network platform for continuous movement assessment", in Int. Workshop on Wearable and Implantable Body Sensor Networks, 2009, pp. 71-76.
[48] D. Halperin et al., "Security and privacy for implantable medical devices", IEEE Pervasive Computing, vol. 7, no. 1, pp. 30-39, Jan-March 2008.
[49] P. Qiao, H. Corporaal and M. Lindwer, "A 0.964 mW digital hearing aid system", in DATE 2011, pp. 1-4.

[50] K. Daniluk, E. Niewiadomska-Szynkiewicz, "Energy-efficient security in implantable medical devices", in Proc. Federated Conf. on computer Science and Information Systems, 2012, pp. 773-778.
[51] R. Pathak and S. Joshi, "Recent trends in RFID and a Java based software framework for its integration in mobile phones", in Asian Himalayas Int. Conf. on Internet (AH-ICI), 2009, pp. 1-5.
[52] Y. Gao, Z. Zhang, H. Wang and H. Lu, "Research on supply chain optimal control based on RFID integrated system", in IEEE Int. Conf. on Information Management, innovation Management and Industrial Engineering (ICIII), 2012, pp. 407-409.
[53] Z. Pala and N. Inanc, "Smart parking applications using RFID technology", in IEEE RFID Eurasia, 2007, pp. 1-3.
[54] U. Karthaus and M. Fischer, "Fully integrated passive UHF RFID transponder IC with 16.7 μW minimum RF input power", IEEE J. Solid State Circuits, vol. 38, no. 10, pp. 1602-1608, Oct. 2003.
[55] A. M. Pappu, X. Zhang, A. V. Harrison and A. B. Apsel, "Process-invariant current source design: methodlogy and examples", IEEE J. Solid-State Circuits, vol. 42, no. 10, pp. 2293-2302, Oct. 2007.
[56] A. P. Brokaw, "A simple three-terminal IC bandgap reference", IEEE J. Solid-State Circuits, vol. SC-9, no. 6, pp.388-393, Dec. 1974.
[57] P. R. Wilson and R. Wilcock, "Yield improvement using configurable analogue transistors", IEEE Electronics Letters, vol. 44, no. 19, Sept. 2008.
[58] B. Murmann, C. Vogel and H. Koeppl, "Digitally enhanced analog circuits: system aspects" in Proc. Int. Symp. Circuits and Systems (ISCAS), 2008, pp. 560-563.
[59] M. Pastre and M Kayal, "Methodology for the digital calibration of analog circuits and systems using sub-binary radix DACs", in Proc. MIXDES 2009, pp. 456-461.
[60] R. Serrano-Gotarredona et al., "The stochastic I-pot: a circuit block for programming bias currents", IEEE Trans. Circuits Systems II-Express Briefs, vol. 54, no. 9, pp. 760-764, Sept. 2007.
[61] T. Delbruck and P. Lichtsteiner, "Fully programmable bias current generator with 24-bit resolution per bias" in Proc. IEEE 2006 Int. Symp. Circuits Syst. (ISCAS), May 2006, pp. 2849-2852.
[62] M. Frey and H.-A. Loeliger, "On the static resolution of digitally-corrected analog-to-digital and digital-to-analog converters with low-precision components", IEEE Trans. Circuits and Systems I, vol. 54, no.1, pp. 229-237, Jan. 2007.
[63] E. Iroaga and B. Murmann, "A 12-bit 75-MS/s pipelined ADC using incomplete settling", IEEE J. Solid-State Circuits, vol. 42, no. 4, pp. 748-756, April 2007.
[64] A. Panigada and I. Galton, "Digital background correction of harmonic distortion in pipelined ADCs" IEEE trans. Circuits Syst. I, vol. 53, no. 9, pp. 1885-1895, Sept. 2006.
[65] Y. Oh and B. Murmann, "System embedded ADC calibration for OFDM receivers", IEEE Trans. Circuits and Syst. I, vol. 53, pp. 1693-1703, 2006.
[66] R. B. Staszewski et al., "All-digital TX frequency synthesizer and discrete-time receiver for Bluetooth radio in 130-nm CMOS", IEEE J. Solid-State Circuits, vol. 39, no. 12, pp. 2278-2291, Dec. 2004.

[67] B. Murmann and B. E. Boser, "A 12-b, 75MS/spipelined ADC using open-loop residue amplification" IEEE J. Solid-State Circuits, vol. 38, no. 12, pp. 2040-2050, Dec. 2003.
[68] X. Zhang and A. B. Apsel, "A low-power, process-and-temperature-compensated ring oscillator with addition-based current source", IEEE Trans. Circuits and Systems I, vol. 58, no. 5, pp. 868-878, May 2011.
[69] K. Sundaresan, P. E. Allen and F. Ayazi, "Process and temperature compensation in a 7 MHz CMOS clock oscillator", IEEE J. Solid-State Circuits, vol. 41, no. 2, pp. 433-442, Feb. 2006.
[70] R. J. Widlar, “New developments in IC voltage regulators,” IEEE J. Solid-State Circuits, vol. 6, no. 1, pp. 2-7, Feb. 1971.
[71] K. Sanborn, D. Ma and V. Ivanov, "A sub-1-V low-noise bandgap voltage reference", IEEE J. Solid-State Circuits, vol. 42, no. 11, pp. 2466-2481, Nov. 2007.
[72] G. C. M. Meijer, P. C. Schmale and K. Van Zalinge, "A new curvature-corrected bandgap reference", IEEE J. Solid-State Circuits, vol. 17, no. 6, pp. 1139-1143, Dec. 1982.
[73] A. Annema and G. Goksun, "A 0.0025 mm2 bandgap voltage reference for 1.1 V supply in standard 0.16 µm CMOS", in Dig. Tech. Papers IEEE Int. Solid-State Circuits Conf. (ISSCC), Feb. 2012, pp. 364-366.
[74] C. J. Fayomi, G. I. Wirth, H. F. Achigui and A. Matsuzawa, "Sub 1 V CMOS bandgap reference design techniques: a survey", Analog Integr. Circ. Sig. Process, vol. 62, pp. 141-157, Aug. 2009.
[75] H. Banba, H. Shiga, A. Umezawa, T. Miyaba, T. Tanzawa, S. Atsumi and K. Sakui, “A CMOS bandgap reference circuit with sub-1-V operation,” IEEE J. Solid-State Circuits, vol. 34, no. 5, pp. 670-674, May 1999.
[76] G. Giustolisi, G. Palumbo, M. Criscione and F. Cutrì, “A low-voltage low-power voltage reference based on subthreshold MOSFETs,” IEEE J. Solid-State Circuits, vol. 38, no. 1, pp. 151-154, Jan. 2003.
[77] K. Ueno, T. Hirose, T. Asai and Y. Amemiya, “A 300 nW, 15 ppm/°C, 20 ppm/V CMOS voltage reference circuit consisting of subthreshold MOSFETs,” IEEE J. Solid-State Circuits, vol. 44 no. 7, pp. 2047-2054, Jul. 2009.
[78] G. De Vita and G. Iannaccone, “A sub-1-V, 10 ppm/°C, nanopower voltage reference generator,” IEEE J. Solid-State Circuits, vol. 42, no. 7, pp. 1536-1542, Jul. 2007.
[79] K. N. Leung and P. K. T. Mok, “A CMOS voltage reference based on weighted ΔVGS for CMOS low-dropout linear regulators,” IEEE J. Solid-State Circuits, vol. 38, no. 1, pp. 146-150, Jan. 2003.
[80] W. Yan, W. Li and R. Liu, “Nanopower CMOS sub-bandgap reference with 11 ppm/°C temperature coefficient,” Electron. Lett., vol. 45, no. 12, pp. 627-629, Jun. 2009.
[81] L. Magnelli, F. Crupi, P. Corsonello, C. Pace and G. Iannaccone, “A 2.6 nW, 0.45 V temperature-compensated subthreshold CMOS voltage reference,” IEEE J. Solid-State Circuits, vol. 46, no. 2, pp. 465-474, Feb. 2011.
[82] M. Seok, G. Kim, D. Blaauw and D. Sylvester, "A portable 2-transistor picowatt temperature-compensated voltage reference operating at 0.5 V", IEEE J. Solid-State Circuits, vol. 47, no. 10, pp. 2534-2545, Oct. 2012.

[83] V. Ivanov, R. Brederlow and J. Gerber, "An ultra low power bandgap operational at supply from 0.75 V", IEEE J. Solid-State Circuits, vol. 47, no. 7, pp. 1515-1523, Jul. 2012.
[84] R. S. Muller and T. I. Kamins, Device electronics for integrated circuits, 2nd ed, Singapore: Wiley, 1986, pp. 236-240.
[85] F. Cucchi, S. Di Pascoli and G. Iannaccone, "Variability-aware design of a bandgap voltage reference with 0.18% standard deviation and 68 nW power consumption", submitted to IEEE Trans. Circ. Syst. I.
[86] T. Hirose, K. Ueno, N. Kuroki and M. Numa, “A CMOS bandgap and sub-bandgap voltage reference circuits for nanowatt power LSIs,” in Proc. IEEE Asian Solid-State Circuits Conf. (A-SSCC), Nov. 2010, pp. 1-4.
[87] P. Tadeparthy, "A CMOS bandgap reference with correction for device-to-device variation", in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS), May 2004, pp. I- 397-400.
[88] A-J. Annema, “Low-power bandgap references featuring DTMOST’s,” IEEE J. Solid-State Circuits, vol. 34, no. 7, pp. 949-955, Jul. 1999.
[89] K. N. Leung, P. K. T. Mok and C. Y. Leung, "A 2-V 23 µA 5.3-ppm/°C curvature-compensated CMOS bandgap voltage reference", IEEE J. Solid-State Circuits, vol. 38, no. 3, pp. 561-564, Mar. 2003.
[90] G. Ge, C. Zhang, G. Hoogzad and K. A. A. Makinwa, "A single-trim CMOS bandgap reference with a 3σ inaccuracy of ±0.15% from -40 °C to 125 °C", IEEE J. Solid-State Circuits, vol. 46, no. 11, pp. 2693-2701, Nov. 2011
[91] R. T. Perry, S. H. Lewis, A. P. Brokaw and T. R. Viswanathan, "A 1.4 V supply CMOS fractional bandgap reference", IEEE J. Solid-State Circuits, vol. 42, no. 10, pp. 2180-2186, Oct. 2007.
[92] R. J. Widlar, “New developments in IC voltage regulators,” IEEE J. Solid-State Circuits, vol. 6, no. 1, pp. 2-7, Feb. 1971.
[93] H. Banba et al., "A CMOS bandgap reference circuit with sub-1-V operation," IEEE J. Solid-State-Circuits, vol. 34, no. 5, pp.670-674, May 1999.
[94] E. Vittoz, "The design of high performance analog circuits on digital CMOS chips", IEEE J. Solid-State-Circuits, vol. SC-20, no. 3, pp. 657-665, June 1985.
[95] J. Chen and B. Shi, “1 V CMOS current reference with 50 ppm/°C temperature coefficient”, Electron. Lett., vol. 39, no. 2, pp. 209-210, Jan. 2003.
[96] M. Radecker, A. Knoll, R. Kocaman, V. Buguszewicz and R. Rudolf, "A wide range temperature stable integrated current reference", in Proc. IEEE ESSCIRC, Sept. 2003, pp. 583-586.
[97] B.-D. Yang, Y.-K. Shin, J.-S. Lee, Y.-K. Lee and K.-C. Ryu, “An accurate current reference using temperature and process compensation current mirror”, in Proc. IEEE A-SSCC, Nov. 2009, pp. 241-244.
[98] J. Lee and S. Cho, "A 1.4-µW 24.9-ppm/°C current reference with process-insensitive temperature compensation in 0.18-µm CMOS", IEEE J. Solid-State-Circuits, to be published.
[99] G. De Vita and G. Iannaccone, "A 109 nW, 44 ppm/°C CMOS current reference with low sensitivity to process variations", in Proc. IEEE ISCAS, May 2007, pp. 3804-3807.

[100] F. Maloberti, Analog design for CMOS VLSI systems, New York: Kluwer Academic Publisher, 2003, pp. 59-72.
[101] L. Magnelli, F. Crupi, P. Corsonello, C. Pace and G. Iannaccone, “A 2.6 nW, 0.45 V temperature-compensated subthreshold CMOS voltage reference,” IEEE J. Solid-State Circuits, vol. 46, no. 2, pp. 465-474, Feb. 2011.
[102] A. Bendali and Y. Audet, “A 1-V CMOS current reference with temperature and process compensation”, IEEE Trans. Circuits Syst. I, vol. 54, no. 7, pp. 1424-1429, July 2007.
[103] E. M. Camacho-Galeano, C. Galup-Montoro and M. C. Schneider, "A 2-nW 1.1-V self-biased current reference in CMOS technology", IEE Trans. Circuits Syst. II, vol. 52, no. 2, pp. 61-65, Feb. 2005.
[104] R. Dehghani and S. M. Atarodi, "A new low voltage precision CMOS current reference with no external components", IEEE Trans. circuits and Syst. II, An. and Dig. Signal Proc., vol. 50, no. 12, pp. 928-932, Dec. 2003.
[105] B. Razavi, Design of analog CMOS integrated circuits, Singapore: McGraw-Hill, 2001, pp. 377-381.
[106] F. Fiori and P. S. Crovetti, "A new compact temperature-compensated CMOS current reference", IEEE Trans. Circuits Syst. II, vol. 52, no. 11, pp. 724-728, Nov. 2005.
[107] C.-H. Lee and H.-J. Park, "All-CMOS temperature independent current reference", Electron. Lett., vol. 32, no. 14, pp. 1280-1281, July 1996.
[108] H. J. Oguey and D. Aebischer, "CMOS current reference without resistance", IEEE J. Solid-State-Circuits, vol. 32, no. 7, pp. 1132-1135, July 1997.
[109] R. A. Blaushild, "An integrated time reference", in Proc. IEEE Int.Solid-State Circuits (ISSCC), pp. 56-57, 1994.
[110] R. S. Muller and T. I. Kamins, "Device electronics for integrated circuits", Singapore, Wiley.
[111] W. M. Sansen, F. O. Eynde and M. Steyaert, "A CMOS temperature-compensated current reference", IEEE J. Solid-State Circuits, vol. 23, no. 3, pp. 821-824, June 1988.
[112] G. Serrano and P. Hasler, "A precision low-TC wide-range CMOS current reference", IEEE J. Solid-State-Circuits, vol. 43, no. 2, pp. 558-565, Feb. 2008.
[113] F. Sebastiano at al., "A low-voltage mobility-based frequency reference for crystal-less ULP radios", *IEEE J. Solid-State Circ.*, vol. 44, no. 7, pp. 2002-2009, Jul. 2009.
[114] M. S. McCorquodale et al., "A 0.5-to-480 MHz self-references CMOS clock generator with 90 ppm total frequency error and spread-spectrum capability", in *IEEE ISSCC Dig. Tech. Papers*, 2008, pp. 524-525.
[115] Y. Tokunaga, S Sakiyama, A. Matsumoto and S. Dosho, "An on-chip CMOS relaxation oscillator with voltage averaging fedback", IEEE J. Solid-State Circuits, vol. 45, no. 6, pp. 1150-1158, Jun. 2010.

[116] S. Gierkink and E. van Tuijl, "A coupled sawtooth oscillator combining low jitter with high control linearity", IEEE J. Solid-State Circ. vol. 37, no. 6, pp.702-710, Jun. 2002.
[117] K. Sundaresan, P. Allen and F. Ayazi, "Process and temperature compensation in a 7 MHz CMOS clock oscillator", *IEEE JSSC*, vol. 41, no. 2, pp. 433-442, Feb. 2006.
[118] K. Lasanen and J. Lostamovaara, "A 1.2-V CMOS RC oscillator for capacitive and resistive sensor applications", *IEEE Trans. Instrum. Measur.*, vol. 57, no. 12, pp. 2792-2800, Dec. 2008.
[119] K. Lasanen, E. Räisänen-Ruotsalainen, J. Kostamovaara, "A 1-V, Self Adjusting, 5-MHz CMOS RC-Oscillator," *IEEE-ISCAS*, vol. 4, pp. IV-377- IV-380, May, 2002.
[120] F. Bala and T. Nandy, "Programmable high frequency RC oscillator", *IEEE VLSID* 2005.
[121] V. De Smedt, P. De Wit, W. Vereecken and M. Steyaert, "A 66 µW 86 ppm/°C fully-integrated 6 MHz wienbridge oscillator with a 172 dB phase noise FOM", *IEEE J. Solid-State Circuits*, vol. 44, no. 7, pp. 1990-2001, July 2009.
[122] R. Barnett and J. Liu, "A 0.8 V 1.52 MHz MSVC relaxation oscillator with inverted mirror feedback reference for UHF RFID", *IEEE CICC* 2006, pp. 769-772.
[123] U. Denier, "Analysis and design of an ultralow-power CMOS relaxation oscillator", *IEEE Trans. Circuits Syst. I*, vol. 57, no. 8, pp. 1973-1982, Aug. 2010.
[124] K. Choe, O. Bernal, D. Nuttman and M. Je, "A precision relaxation oscillator with a self-clocked offset-cancellation scheme for implantable biomedical SoCs", *IEEE ISSCC* 2009, pp. 402-403.
[125] F. Marraccini, G. De Vita, S. Di Pascoli, G. Iannaccone, "Low Voltage nanopower clock generator for RFID applications," *Microelectron. J.* (2008).
[126] M. Flynn and S. Lidholm, "A 1.2 µm CMOS current-controlled oscillator", *IEEE J. Solid-State Circuits*, vol. 27, no. 7, pp. 982-987, Jul. 1992.
[127] A. M. Pappu, X. Zhang, A. V. Harrison and A. B. Apsel, "Process-invariant current source design: methodlogy and examples", IEEE J. Solid-State Circuits, vol. 42, no. 10, pp. 2293-2302, Oct. 2007.
[128] X. Zhang and A. B. Apsel, "A low-power, process-and-temperature-compensated ring oscillator with addition-based current source", IEEE Trans. Circuits and Systems I, vol. 58, no. 5, pp. 868-878, May 2011.
[129] X: Zhang, M. Mukadam, I. Mukhopadhyay and A. Apsel, "Process compensation loops for high speed ring oscillators in sub-micron CMOS", IEEE Trans. on Emerging and Selected Topics in Circ. and Syst., vol. 1, no. 1, pp. 59-70, March 2011.
[130] M. Green and A. Willson, "How to identify unstable dc operating point", IEEE Trans. Circuits Syst., vol. 39, pp. 820-832, 1992.
[131] L. Chua and D. Green, "A qualitative analysis of the behavior of dynamic nonlinear networks: stability of autonomous networks", IEEE Trans. Circuits Syst., vol. 23, pp. 355-379, 1976.

[132] M. Green and A. Willson, "An algorithm for identifying unstable operating points using SPICE", IEEE Trans. on Computer Aided design of integrated circuits and systems, vol. 14, pp. 360-370, 1995.
[133] G. Storti Gajani, A. Brambilla and A. Premoli, "Numerical determination of possible multiplie DC solutions of nonlinear circuits", IEEE Trans. Circ. Syst., vol. 55, pp. 1074-1083, 2008.
[134] A. Sedra and K. Smith, "Microelectronics Circuits", USA, 1997, pp. 1360.
[135] F. Cucchi, S. Di Pascoli and G. Iannaccone, "Stability and startup of nonlinear loop circuits", submitted to Electronics Letters
[136] A Sedra and K. Smith, "Microelectronics Circuits", USA, 1997, pp. 1360.
[137] F. Fiori and P. S. Crovetti, "A new compact temperature-compensated CMOS current reference", IEEE Trans. Circuits Syst. II, vol. 52, no. 11, pp. 724-728, Nov. 2005.
[138] H. J. Oguey and D. Aebischer, "CMOS current reference without resistance", IEEE J. Solid-State-Circuits, vol. 32, no. 7, pp. 1132-1135, July 1997.
[139] W. Hou and M. Green, "Use of a continuation method for analyzing startup circuits" Proc. of IEEE ISCAS, pp. 1527-1530, May 2010.