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

Subthreshold design of ultra low-power analog modules

Autore:

Luca Magnelli *Firma*_____

Relatori:

Prof. Bruno Neri *Firma*_____

Prof. Giuseppe Iannaccone *Firma*_____

Prof. Felice Crupi *Firma*_____

Anno 2012

SOMMARIO

Il consumo di potenza rappresenta l'indicatore chiave delle performance di recenti applicazioni portatili, come dispositivi medici impiantabili o tag RFID passivi, allo scopo di aumentare, rispettivamente, i tempi di funzionamento o i range operativi. La riduzione della tensione di alimentazione si è dimostrata l'approccio migliore per ridurre il consumo di potenza dei sistemi digitali integrati. Al fine di tenere il passo con la riduzione delle tensioni di alimentazione, anche le sezioni analogiche dei sistemi mixed signal devono essere in grado di funzionare con livelli di tensione molto bassi. Di conseguenza, sono richieste nuove metodologie di progettazione analogica e configurazioni circuitali innovative in grado di lavorare con tensioni di alimentazioni bassissime, dissipando una potenza estremamente bassa. Il regime di funzionamento sottosoglia consente di ridurre notevolmente le tensioni applicabili ai dispositivi ed si contraddistingue per i livelli di corrente molto bassi, rispetto al ben noto funzionamento in forte inversione. Queste due caratteristiche sono state sfruttate nella realizzazione di moduli analogici di base ultra low voltage, low power.

Tre nuove architetture di riferimenti di tensione, che lavorano con tutti i transistor polarizzati in regime sottosoglia, sono stati fabbricati in tecnologia CMOS 0.18 μm. I tre circuiti si basano sullo stesso principio di funzionamento per compensare gli effetti della variazione della temperatura sulla tensione di riferimento generata. Tramite il principio di funzionamento proposto, la tensione di riferimento può essere approssimata con la differenza delle tensioni di soglia, a temperatura ambiente, dei transistor. Misure sperimentali sono state effettuate su set con più di 30 campioni per ogni configurazione circuitale. Una dettagliata analisi statistica ha dimostrato un consumo medio di potenza che va da pochi nano watt a poche decine di nano watt, mentre la minima tensione di alimentazione, raggiunta da una delle tre configurazioni, è di soli 0.45 V. Le tensioni di riferimento generate sono molto precise rispetto alle variazioni della temperatura e della tensione di alimentazione, infatti sono stati ottenuti coefficienti di temperatura e line sensitivity medi a partire rispettivamente da 165 ppm/°C e 0.065 %/V.

Inoltre, è stata trattata anche la progettazione di amplificatori ultra low voltage, low power. Sono state illustrate linee guida dettagliate per la progettazione di amplificatori sottosoglia e le stesse sono state applicate per la realizzazione di un amplificatore a due stadi, con compensazione di Miller, funzionante con una tensione di alimentazione di 0.5 V. I risultati sperimentali dell'op-amp proposto, fabbricato in tecnologia CMOS 0.18 μm, hanno mostrato un guadagno DC ad anello aperto di 70 dB, un prodotto banda-guadagno di 18 kHz ed un consumo di potenza di soli 75 nW. I risultati delle misure sperimentali dimostrano che gli amplificatori operazionali in sottosoglia rappresentano una soluzione molto interessante nella realizzazione di applicazioni efficienti in termini energetici per gli attuali sistemi elettronici portatili. Dal confronto con amplificatori ultra low power, low voltage presenti in letteratura, si evince che la soluzione proposta offre un miglior compromesso tra velocità, potenza dissipata e capacità di carico.

ABSTRACT

Power consumption is the key performance indicator of emerging portable applications such as implantable medical devices or passive RFIDs, in order to increase their battery autonomy or operating range, respectively. Supply voltage reduction has proven to be the best approach to reduce power consumption in integrated digital systems. In order to keep pace with supply voltage reduction, also analog sections of mixed signal systems should be able to operate with very low voltage levels. Therefore, innovative analog design methodologies and circuit configurations to be operated at very low supply voltages, while dissipating extremely low power, are required. Subthreshold regime allows a remarkable reduction of the supply voltages and is characterized by very low current levels, when compared to the well known strong inversion operation. Those two features have been exploited in the implementation of ultra low power, low voltage analog subsystems.

Three novel voltage reference architectures, operating with all transistors biased in weak inversion, were fabricated with a 0.18 µm CMOS technology. The three circuits are based on the same operating principle to compensate temperature variation effects on the generated voltage. By means of the proposed operating principle, the reference voltage can be approximated by the difference of transistor threshold voltages at room temperature. Measurements were carried out over sets of more than 30 samples per each circuit configuration. Detailed statistical analysis showed an average power consumption ranging from few nanowatts to few tens of nanowatts, while the minimum supply voltage, achieved by one of the three configurations, is just 0.45 V. The generated reference voltages are very accurate versus temperature and supply voltage variations, indeed, mean temperature coefficients as low as 165 ppm/°C and line sensitivities as low as 0.065 %/V were obtained.

Furthermore, the design of ultra low voltage, low power amplifier was also covered. Detailed subthreshold design guidelines of op-amps were provided and applied to the implementation of a 0.5-V two-stage Miller-compensated amplifier. Experimental results of the proposed op-amp fabricated with a 0.18 µm CMOS process showed a DC open loop gain of 70dB, a gain-bandwidth product of 18 kHz and a power dissipation of just 75 nW. Such results demonstrate that subthreshold op-amps are a very attractive solution to implement sub-1 V energy-efficient applications for modern portable electronic systems. A comparative analysis with low-voltage, low-power op-amp designs available in the literature highlighted that the proposed solution represents a better trade-off among speed, power and load capacitance.

INDICE

SOMMARIO 2

ABSTRACT 3

1. INTRODUCTION 6

1.1. Low-power electronics 6

1.2. Energy-constrained systems 7

1.2.1. Radio Frequency IDentification (RFID) transponders 7
1.2.2. Wireless Sensor Networks (WSNs) 9
1.2.3. Implantable medical devices 11
1.2.4. Ultra low-power VLSI systems 12

1.3. Low-power design 12

1.3.1. Sources of energy dissipation in VLSI CMOS circuits 13
1.3.2. Subthreshold CMOS circuits 14
1.3.3. Low-power, low-voltage analog design. 14

1.4. Organization of the work 16

2. WEAK INVERSION REGIME 18

2.1. Introduction 18

2.2. I-V characteristic 18

2.3. Threshold Voltage 22

2.4. Small signal equivalent circuit 23

3. LOW-VOLTAGE, LOW-POWER VOLTAGE REFERENCES 28

3.1. Introduction 28

3.2. Performance specifications 29

3.2.1. Line sensitivity 30
3.2.2. Temperature coefficient 30
3.2.3. Power Supply Rejection Ratio 30
3.2.4. Other specifications 30

3.3. Bandgap References (BGRs) 30

3.3.1. Basic principle 31
3.3.2. Sub-bandgap reference circuits 33

3.4. Subthreshold CMOS voltage references 37

3.4.1. BGR-principle and weak inversion 37
3.4.2. Not BGR-based architectures 40

3.5. Proposed voltage references *45*
3.5.1. Operating principle 45
3.5.2. Electrical configurations 50
3.5.3. Experimental results 67
4. SUBTHRESHOLD OPERATIONAL AMPLIFIER DESIGN 77
4.1. Introduction *77*
4.2. Low-voltage, low-power CMOS op-amps *77*
4.2.1. Bulk-driven MOSFETs 77
4.2.2. Forward body bias 79
4.2.3. Subthreshold operation 80
4.3. Design of ultra low-power, low-voltage subthreshold opamps *82*
4.3.1. Design guidelines 83
4.3.2. Two-stage op-amp design 87
4.4. Experimental results *90*
CONCLUSION 95
REFERENCES 96

1. INTRODUCTION

1.1. Low-power electronics

In recent years the Electronics Market has been experiencing an impressive increase in demand for battery operated equipments. Moreover, very low-power applications, like portable medical devices or microsensor nodes, represent a rapidly growing niche of such market segment. For this category of products, low power is the main requirement, and other key parameters, such as speed and/or dynamic range, might have to be sacrificed. Indeed, power consumption reduction is the most common way to extend battery lifetime, in order to reduce recharging or replacements, that are often costly and/or difficult to perform. For example, sensors for flow-rates monitoring in oil pipelines [1], sensors used in heating, ventilation and air conditioning (HVAC) systems [2] or those used to measure a building's structural integrity during an earthquake [3], are often inaccessible and battery replacement could require infrastructure disassembling. Moreover, considering implanted medical devices, the battery replacement is not only costly but requires invasive surgery and remote charging would require high-power density that could cause tissue heating and damage [4]. Very often, in such applications, energy harvesting systems, that can further extend battery lifetime, are provided [5].

Nevertheless, power reduction issues became also crucial in applications like mobile phones and PDAs, that require high performance for limited periods of time and then spend most of the remaining time doing non-performance-critical tasks. The growing relative costs of power supplies are leading to an increasing demand for low power design techniques. In [6] the Ultra-Dynamic Voltage Scaling (UDVS) [7] was suggested as power management solution in mobile phones, since they often operate in a near-idle fashion waiting for input from the user or the wireless link. Thanks to the UDVS technique, a circuit can operate at high-voltage/high-frequency during performance-critical applications and at subthreshold/low-frequency during energy-constrained applications.

Moreover, relentless technology scaling recently gave rise to a set of reliability problems in circuits that traditionally are driven almost solely by performance constraints. These problems made power consumption an important design issue. Excessive power consumption in high performance VLSI-based systems causes self-heating, i.e. the chip temperature increases and device reliability and system performance are consequently degraded. Therefore, expensive cooling packages are required for safe operation and more and more expensive batteries are needed in order to ensure enough lifetime of high performance portable systems, such as laptops.

The growing interest towards low-power electronics is also testified by the ORTC's (Overall Roadmap Technology Characteristics) tables [8] of the International Technology Roadmap fo Semiconductor, where specific power supply and power dissipation requirements for low-operating-power systems are specified. Moreover,

in the “ITRS Design and System Drivers” of the ITRS 2010 update, a key message is:

“… power consumption has become the key technical parameter that controls feasible semiconductor scaling. As a result, the device roadmap has become power-driven, and frequency is being pushed to a flat trend.”

As a consequence, energy-constrained design methodologies are gaining more and more attention in the electronic research field. Subthreshold (or weak inversion) design has become very popular nowadays in Ultra Low-Power design techniques. Although subthreshold operation was first applied more than 30 years ago to limit the power consumption of integrated circuits within electronic watches, it is recently extending to a larger market, mostly because of the relentless supply voltage scaling of integrated circuits. Indeed, supply voltage reduction is the main means for power reduction in digital circuits, leading to operate devices with voltages well below their threshold (V_{th}).
However, voltage downscaling and weak inversion operation pose several challenges to designers. Indeed, apart from the expected performance degradation, the most important problem they have to face is the high sensitivity to temperature variations and process variability. Moreover, considering the ever increasing need for high integration that is pushing towards mixed-signal SoCs, analog subsections are also supposed to operate, in the near future, with decreasing supply voltages, e.g. under 1 V. In that case, only few procedures for subthreshold analog design have been proposed [9].
Nevertheless, weak inversion operation allows power savings around 10x or 100x compared with strong inversion operation. In the following paragraphs a survey of applications for which energy consumption is the key performance indicator is presented.

1.2. Energy-constrained systems

1.2.1. Radio Frequency IDentification (RFID) transponders

Radio Frequency IDentification is a technology that exploits RF signals to exchange data between a reader and an electronic tag, for purpose of identification and tracking. RFID tags are relatively small and cheap. They are, usually, attached to the objects to be automatically identified. The RFID technology was originally proposed as a replacement of barcodes and smart cards whereas, nowadays, its area of application is growing increasingly: supply chain management, access control to buildings, public transportation and open-air events, airport baggage, animal identification, express parcel logistics, and many more. To meet the requirements of such applications an RFID systems must have several specific features, among them the most important are extreme low power and low cost.
Identification information is stored within the silicon chip included in the RFID transponder. Therefore, it can be identified in a unique way by an identification code. Basically, the silicon chip can store a large amount of information that can be read and written wirelessly at a distance of up to several meters. Typically, tags are grouped in two main categories, depending on the presence or the absence of an autonomous power source on board. They are defined “active” if they require an on-board battery to supply all or part (in that case they are called semi-passive) of

the power required by the transponder for transmission and data processing. A "passive" RFID tag, instead, does not have an on-board battery, thus the transponder is supplied by rectifying the RF power transmitted by the reader. In particular, a fraction of such power is employed by the tag to communicate with the reader by modulating the impedance that the reader sees. It follows that passive devices are usually smaller and their lifetime is not limited by energy, although their available power budget is very limited.
RFID transponders design is definitely power-driven. Reducing power consumption would benefit both types of tags. For active tags, minimizing the power dissipation leads to longer battery lifetimes. For passive tags, instead, the power is constrained by the ability to utilize the converted energy from the built-in antenna. If the power dissipation can be reduced, then the distance between the reader and the tag can increase since less transmitted power has to reach the tag. As an example of an RFID transponder architecture, the one of the passive transponder IC fabricated in [10] is shown in Fig. 1.1,

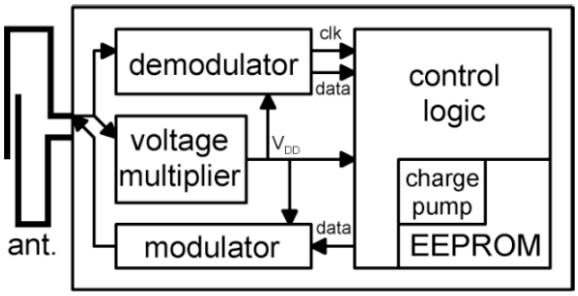


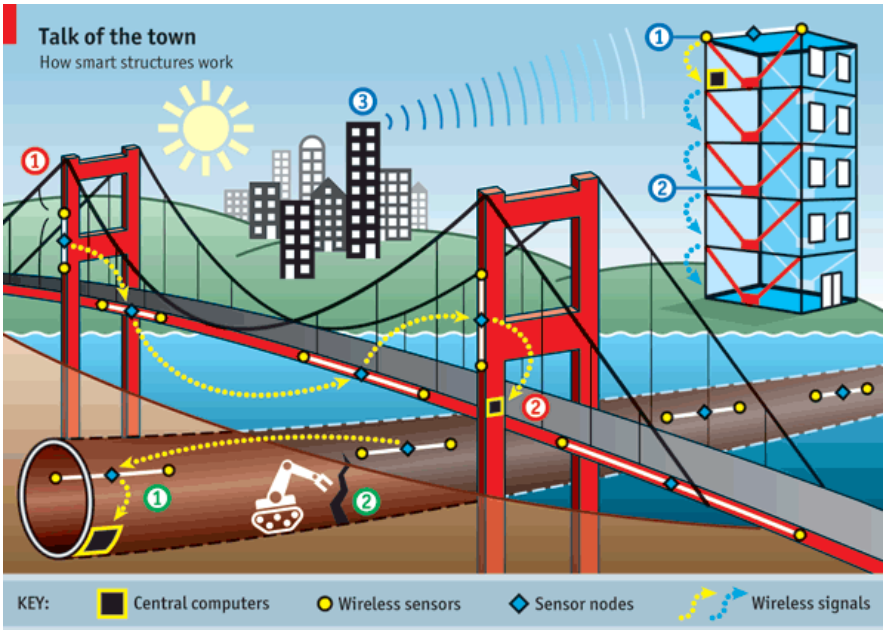
Fig. 1.1 – Block diagram of a passive RFID tag

The nature of the antenna varies accordingly to the coupling method adopted in the RFID system. In the case of inductive coupling the antenna is a coil, inductive coupling systems operate at frequencies around 135 kHz or 13.56 MHz achieving operating ranges smaller than 1 m. In electromagnetic coupled RFID systems, the coupling element is an antenna, which typically is a dipole or a patch antenna. They usually exploit an electromagnetic coupling in the UHF (868 MHz in Europe and 916 MHz in USA) or μwave range (2.45 GHz or 5.8 GHz). Such systems can reach an operating range of few meters, in the case of passive transponder, and larger than 15 m, in the case of active transponder. The voltage multiplier converts a part of the incoming RF signal power to DC for a voltage regulator (e.g. [11]) which provide a stable power supply for all active circuits on the chip. The average input impedance of the voltage multiplier should be power matched with the antenna in order to reduce losses. The demodulator converts the pulse-width-modulated input signal to digital data and generates a synchronous system clock. The control logic handles the received data that can be written or read. A charge pump circuit is provided to obtain suitable voltage levels for EEPROM programming. The modulator then converts data from the control logic to changes in the input impedance using a MOS varactor.. The logic circuitry handles the protocol, including anticollision features, cyclic redundancy checks (CRC) checks,

error handling, enabling and disabling of analog circuits (power down, standby, power up), etc..
Further improvements in RFID systems in order to increase the operating range and/or battery lifetime can be achieved both at system level, e.g. by a proper choice of modulation technique, and at circuit level, by developing extreme low voltage, low power design methodologies.

1.2.2. Wireless Sensor Networks (WSNs)

A WSN consists of a large number (from tens to thousands) of spatially distributed autonomous sensors (nodes) that monitor physical quantities and/or environmental conditions, like temperature, sound, acceleration, vibration and strain, pressure, pollutants, etc., and cooperatively through the network to a main location. Such sensors have existed for decades but they have never been widely deployed as in recent years, because they needed wires to provide power and carry data. Nowadays, thanks to the miniaturization of electronics and improvements in battery and radio technology, each node in a WSN is self powered (i.e. battery operated) and can share the collected and/or processed information through the a series of wireless links between nodes. Specific features of interest to the end-user can be extracted from the collection of data gathered by several nodes.
First applications of WSNs were proposed in the military field, such as battlefield surveillance. Today more civilian applications are envisaged ranging from industrial to consumer applications such as: environment and habitat monitoring, industrial process monitoring and control, healthcare applications, home automation, traffic control, structural monitoring and automotive sensing.
In [12] a survey of WSNs for structural monitoring is presented, evidencing how wireless sensors can be added to existing structures much more easily, quickly and cheaply than wired sensors. A network of 64 wireless sensors has been installed on the Golden Gate Bridge in San Francisco to monitor vibrations. In the anchorage chamber of the Humber Bridge in East Yorkshire (England), humidity sensors helped maintenance, since steel struts must be kept dry. A straightforward example of how WSNs works in structural monitoring in constructions like buildings, tunnels or bridges, can be found in Fig. 1.2 ([12]).
Getting WSNs to work reliably and for a long time in similar environments presents several challenges. Generally, wireless sensors benefit from long device lifetimes, reduced costs and small size. Low power and energy harvesting from the environment increase the time before energy stored on-sensor is depleted. Performance requirements of sensor nodes are quite low, the rate at which data change, for example, for environmental or health monitoring, is of the order or seconds to minutes.
The block diagram of a wireless sensor [13] is shown in Fig. 1.3, where power consumption is the key parameter of each component. The analog front end interfaces the sensors (of temperature, pressure, humidity, etc.) with the internal blocks. Usually this first stage consists of an LNA providing adequate gain and drive strength for subsequent circuits while maintaining a high SNR. The low output rates of sensors allowed ultra low power amplifier design with devices operated in weak inversion. In [14] a 3.5 μW LNA for amplifying low frequency (<300 Hz) electroencephalographic signals from electrodes placed on the scalp, is presented. After the amplification stage the sensor signals can be converted to digital,



SMART BUILDING

1. Sensors in a building monitor the building's movement in response to strong winds or earthquake tremors.
2. Shock absorbers (hydraulic dampers) can then be made to stiffen or relax and heavy weights (mass dampers) can be moved to reduce oscillations in strong winds, or minimise damage in the event of an earthquake.
3. Buildings that detect an earthquake tremor could even warn other buildings nearby of the approach of a shockwave, so they could sound an alarm and prepare themselves accordingly.

SMART BRIDGE

1. Wireless sensors mounted on the bridge monitor vibrations, displacement and temperature. This information then "hops" across the network of sensor nodes to a central computer for analysis.
2. If a problem is detected, such as a loose bolt or cable, or the beginning of a crack, a warning can be sent by SMS.

SMART TUNNEL

1. Wireless sensors mounted on the walls of a tunnel monitor displacement, temperature and humidity. This information then "hops" across the network of sensor nodes to a central computer for analysis.
2. If a problem with the tunnel lining is detected, appropriate maintenance can be carried out. In future, a smart tunnel could even use robots to perform some maintenance tasks automatically.

Fig. 1.2 – Example of WSNs for structural monitoring

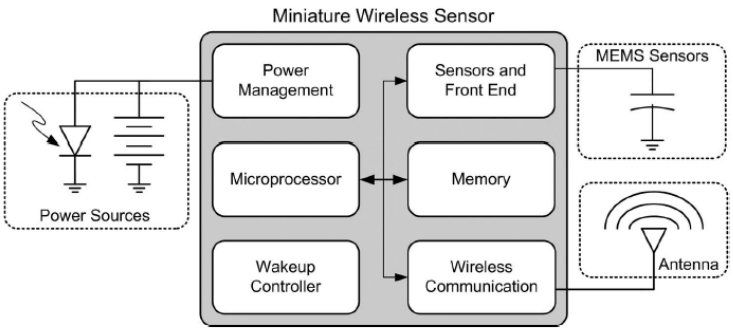


Fig. 1.3 – Architecture of a wireless node

so they can be processed, stored and transmitted. Microwatts successive approximations register (SAR) ADCs have been designed for sensor environments [15], [16]. Voltage scaling is the most effective way of reducing energy consumption in processors as well as in memories. Notwithstanding, designers must face to increased latency and susceptibility to process variations that, for example, are much more critical in subthreshold operated devices than in those operated in strong inversion. Wireless communication is achieved by means of RF transmission. Nowadays, even the most energy efficient transceiver standards in commerce, i.e. ZigBee, WiFi or Bluetooth, have power consumptions in the order of tens of mW that is not acceptable for a typical wireless sensor. Again, due to low bandwidth, weak inversion operation and voltage scaling can help to dramatically reduce power consumption as in the case of the 0.5 V receiver presented in [17]. Network nodes require supply voltage ranging from 400 mV (for subthreshold operated modules) to 1.2 V (transmission modules). Since Lithium (3.3 V – 4.2 V), zinc-air (1.5 V) and alkaline (1.5 V) batteries are very common, dc-to-dc converting power electronics are needed. Linear regulators are often employed to create an output supply in ultra low-power applications [11]. Linear regulators buffer the output of a voltage reference that, generally, consists of a Bandgap Reference (BGR). Unfortunately BGRs tend to have μW power consumption that could exceed the power budget of low-power sensor nodes. Nevertheless, CMOS voltage reference circuits with nA current levels and sub-1 V supply voltages have been developed [18], [19]. Switched capacitor networks or buck down-converters might also be used, compared to linear regulators they offer better power efficiencies but are noisier, unless the switching frequency of the network is very high. Finally, energy-harvesting systems, such as photovoltaic solar diodes or thermocouples, can be employed to recharge batteries for longer sensor lifetime.
It follows that each component of the architecture should be designed to achieve robust ultra low power operation in order to allow long battery lifetime (e.g. > 1 year).

1.2.3. Implantable medical devices

Implanted Medical Devices (IMDs) are an emerging class of medical devices for continuously measuring internal health status and physiological signals [20]. Generally their purpose is to constantly monitor health parameters impossible to

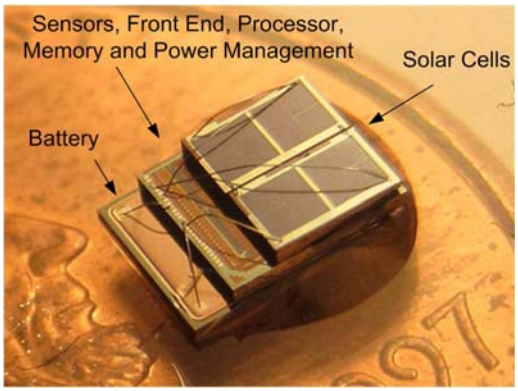


Fig. 1.4 – Millimeter scale ultra low power sensor

monitor externally, like intraocular pressure in glaucoma patients [21]. Others IMDs uses the collected measurements as trigger for physiological interventions that prevent impending adverse events (e.g. epileptic seizures [22]) and for physical assistance (e.g. brain-controlled motor prosthetics [23]). The latter is also the case of the well known systems for arrhythmias detection and heartbeat correction, such as peacemakers and Implanted Cardioverter Defribillators.
Given their implantable nature these devices face severe size constraints and need to communicate and receive power wirelessly. Moreover, power consumption reduction is also crucial: early hearing aid systems dissipated few tens of mW [24], nowadays the power budget of an IMD does not exceed 500 nW [13] as in the case of the intraocular pressure monitoring sensor [13], shown in Fig. 1.4, assembled with energy harvesting elements and an RF antenna for data transmission.
Thus it is evident that IMDs are another compelling platform showing the need for innovative low-energy techniques in order to extend battery lifetime, thus reducing interventions for their replacement.

1.2.4. Ultra low-power VLSI systems

Most portable electronics used for consumer applications require a low power DSP or MCU. In that case adequate performance during active functioning is as important as low power consumption. In a variety of applications, for example, the Texas Instruments (TI) MSP430 family has been used for portable measurement, metering and instrumentation. Such MCU has a current consumption in active mode of 330 μA at 1 MHz with a 2.2 V supply voltage, while in standby mode it only dissipates 1.1 μA. It has been used in the SHIMMER wearable sensor platform [25] along with a CC2420 IEEE 802.15.4 radio, a triaxial accelerometer and a rechargeable Li-polymer battery. The total device power budget is 60 mW when active and a few μW in sleep mode, whereas the device lifetime varies from hours to days, depending on the application duty cycle.
There are many other applications that take advantage of a wide range of performance needs to reduce energy consumption and extend the system lifetime (mobile phones, PDAs, laptops, etc.). They should minimize energy when in standby or sleep modes and maximize performance when in high activity modes. Ultra low power techniques, many of them based on the subthreshold operation mode, are also used in such applications [7].

1.3. Low-power design

Even though with the advent of the transistor and, later on, the coming of the integrated circuits, greater emphasis was placed on performance and miniaturization issues, nonetheless power dissipation was not completely ignored. Actually, battery operated equipments, such as pocket calculators, implantable peacemakers, hearing aids and, most importantly, wristwatches, drove low power electronics. As a consequence, power requirements become one of the most critical concerns in the evolution of microelectronic technology. In order to continue to improve the performance and to integrate more and more functionalities within chips, feature size had to shrink accordingly. As a result, the magnitude of power per unit area has kept growing and the associated problem of heat removal and cooling has kept getting worse. Even with the scaling down of supply voltage,

power dissipation has not decreased since higher levels of integration also allows for more functionalities. Moreover, in the last years, energy-constrained applications, like those shown in the previous paragraphs, have been emerging which require new design techniques to achieve longer lifetimes.

1.3.1. Sources of energy dissipation in VLSI CMOS circuits

As long as power consumption of VLSI systems is concerned, it is well known that the largest part is due to digital sections. The power consumption in digital VLSI circuits can be written as [26], [27]:

$$P = \alpha f C_L V_{DD}^2 + I_{SC} V_{DD} + I_{leak} V_{DD}, \tag{1.1}$$

where the first addend accounts for the switching power, the second one for the short circuit power and the last one for the static leakage power. In (1.1) V_{DD} is the supply voltage, C_L is the load capacitance, f is the switching frequency, α is the switching activity (i.e. the fraction of f representing the commutations in which energy is absorbed from the power supply), I_{SC} is the short circuit current and I_{leak} is the static leakage current.

The switching power and the short circuit power form the dynamic power fraction of power dissipation in digital circuits. Thus far, the successful approach for reducing energy consumption has been V_{DD}-scaling. Such approach has been also driven by reliability and technological issues associated to feature size scaling as depicted in Fig. 1.5 [26],

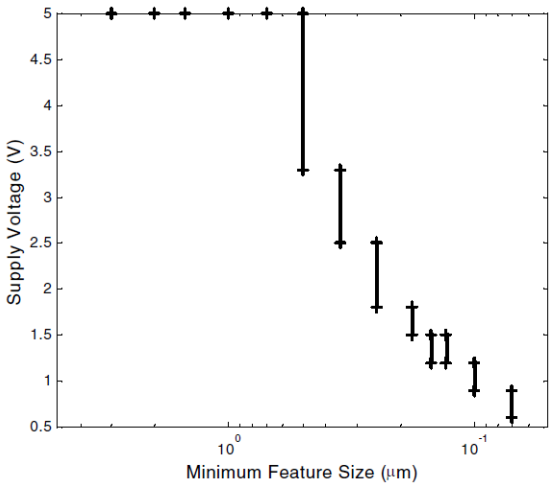


Fig. 1.5 – Max and min V_{DD} values as a function of feature size.

To preserve performance and, therefore, sufficient MOSFETs' current drive, threshold voltages (V_{th}) has been consequently decreased. Unfortunately V_{th}-scaling leads to a leakage current (I_{leak}) increase. In particular, decreasing V_{DD} and V_{th}, among all the leakage factors within a MOSFET [27] (pn-junction and punchthrough leakage, gate oxide leakage, hot-carrier injection, subthreshold leakage, gate induced drain leakage, etc.), the subthreshold current becomes the prevailing one, since it changes exponentially with V_{th} variations.

Several design techniques have been developed to face with dynamic and static energy dissipation: device sizing for power efficiency [28], multithreshold design techniques [29], power gating, variable body biasing [30], [31] and subthreshold design [7], [32].
Process technology has a large effect on energy consumption of analog and digital circuits. Newer processes have smaller device capacitances and lower dynamic energy, while older technologies exhibit lower standby power (higher V_{th}s). CMOS processes could be optimized specifically for low power operation modes, such as subthreshold or near-threshold. For example, recent CMOS technologies include halo implants at the edge of the conducting channel to reduce the effect of the drain-induced barrier lowering (DIBL). For subthreshold operation, DIBL is not a relevant issue since the involved drain-to-source voltages are low. This, therefore, allows the removal of halo implants, which improves subthreshold swing and overall subthreshold device performance [33].

1.3.2. Subthreshold CMOS circuits

Subthreshold circuits are ideal for application where performance is not critical but minimizing energy consumption is the key. The very first employment of subthreshold conduction in the CMOS technology (1954) was inside electronic watches [34], because they needed micropower level to limit power consumption. Subsequently, in 1972, a model for subthreshold conduction was proposed in [35], in which the authors showed how applying it to find the transfer characteristic of an inverter, CMOS logic circuits can operate at a V_{DD} as low as $8k_{\mathrm{B}}T/q$ (i.e. about 200 mV at room temperature).
At that time most subthreshold circuit design were focused on analog circuits, in [36] Vittoz and Fellrath demonstrated subthreshold circuits such as an amplitude detector, a quartz ring oscillator, a bandpass amplifier and a Proportional To Absolute Temperature (PTAT) current reference that represented the starting point of the well known Oguey's current reference [37] capable to supply a stable current as low as 1 nA. Subthreshold transconductance amplifiers were designed for cochlea applications [38].
Subthreshold digital circuits appeared later, with the newly recognized need of limiting power consumption [39], in particular inside portable systems. With modern submicron processes, a clock frequency beyond 10 MHz can be achieved with V_{DD}s of just few hundred millivolts, dramatically reducing the power-delay product [40].
As a result, the reduction of supply voltages imposed by the process of scaling down feature size imposes a reduction of the saturation voltage of transistors used in analog circuits. This is only possible to achieve by reducing the amount of MOSFET channel inversion, thus entering in weak inversion region for $V_{DD} \leq 0.5$ V.

1.3.3. Low-power, low-voltage analog design.

Even though power consumption in digital circuits is dramatically reduced by voltage-scaling, low power approaches are quite different in the analog domain. Indeed, power consumption is mainly set by the required Signal-to-Noise-Ratio (SNR) and the required bandwidth in analog circuits, while decreasing V_{DD} does not necessarily reduce it. In analog signal processing, power is absorbed from power supply to keep the signal energy well above the fundamental thermal noise, in order to achieve a satisfactory SNR.

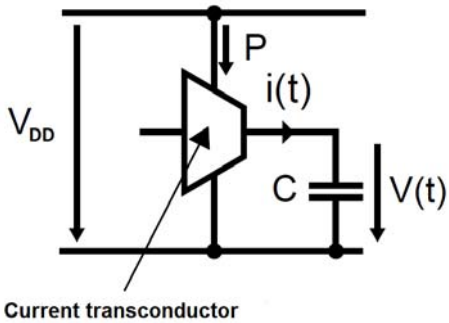


Fig. 1.6 – Basic scheme of an integrator

As a clarifying example let us consider the power consumed to realize a single pole characterized by the integrator shown in Fig. 1.6 [41]. The integrator consists of an integrating capacitor and a 100% efficient current transconductor (i.e. all the current absorbed from the power supply is employed to charge the capacitor). The power absorbed from the supply voltage V_{DD} by the integrator to have a sinusoidal voltage $V(t)$ with a peak-to-peak amplitude V_{pp} and a frequency f across the capacitor C, is $P = fCV_{pp}^2\left(V_{DD}/V_{pp}\right)$. Since the output noise power is k_BT/C, where k_B is the Boltzmann constant and T is the absolute temperature, the SNR can be expressed as,

$$SNR = \frac{V_{pp}^2/8}{K_BT/C}. \tag{1.2}$$

Therefore, the power necessary to realize a single pole, in order to fulfill a given SNR, is,

$$P = 8k_BTf\frac{V_B}{V_{pp}}SNR. \tag{1.3}$$

As a result, the minimum power consumption at a specific T, is fixed by the SNR and the required bandwidth. It is clear that power efficient analog circuits should thus be designed to maximize the voltage swing and therefore should handle rail-to-rail signal voltages (i.e. $V_{pp}=V_{DD}$). This power limit is very severe, since it requires an increase of 10x in power for every 10 dB of SNR improvement. It follows that analog circuits become power inefficient in the case of systems requiring high values of SNR.

Moreover, practical circuits present further technological limitations that hinder the achievement of the discussed theoretical minimum power consumption. For example, bias circuitry (voltage and current references, current mirror, etc.) wastes energy, they should be designed in order to reduce as much as possible power consumption. Furthermore, they can increase the noise and therefore a proportional increase of power to maintain a given SNR. Since minimum power consumption is achieved for $V_{pp}=V_{DD}$, the signal should be amplified at the early stages of the analog systems and maintained along the processing path. Additional noise sources coming from active and passive components or from the power supply push the power to increase. Another aspect that leads to an increase in power consumption is the need for precision, which imposes the use of larger

dimensions for active and passive components, with a resulting increase in parasitic capacitances and, therefore, power.
Even though the minimum power consumption for a given SNR from (1.3) yields to no fundamental reasons to reduce the supply voltage, because of the technological limits described above, the power dissipated in analog circuits also depends on the V_{DD} value. Moreover, nowadays the trend in CMOS technologies is the V_{DD}-scaling: in [8] the supply voltages of low operating power systems are expected to reach 0.6 V within 2021. Furthermore, in order to avoid power consuming voltage boosting circuits, such as charge pumps, analog circuits within mixed-signal SoCs should be redesigned for ultra low-voltage.
In micropower analog circuits, MOSFETs operated in weak inversion regime provide several advantages for low-power and low-voltage purposes. The most important aspect for low-voltage systems is that the drain-source saturation voltage in weak inversion is much smaller than the one needed in strong inversion, since it is sufficient that it is larger than some thermal voltages, as will be explained in the following chapter. Moreover, another advantage strongly related to the analog design is that the transconductance-to-current ratio reaches its maximum in weak inversion. Nevertheless, weak inversion operation has drawbacks too, such as the intrinsic slowness (the transition frequency of subthreshold devices does not exceed few hundred MHz) and the higher sensitivity to process and temperature variations, compared to the strong inversion operation.

1.4. Organization of the work

In this chapter the motivations for innovative design approaches in future energy-constrained applications have been discussed. In particular, portable applications have stringent power requirements. Nowadays, one way to extend battery lifetimes is to run at microwatt power levels. Supply voltage reduction has been proved as the best approach to reduce power consumption in digital systems. In order to keep the pace with supply voltage reduction, analog sections of mixed signal systems should be able to operate with very low voltage. Moreover their power consumption should not represent the overwhelming part of the total energy budget. Therefore, novel analog design methodologies and circuit configurations to be included in battery-operated applications, where power consumption is the key performance indicator, are required. Subthreshold operation holds the most promise for ultra low power, low voltage design in analog as well as in digital domain.
In Chapter 2, subthreshold conduction in MOSFETs is examined: basic DC relations and AC small signal equivalent circuit parameters are given and discussed. Moreover a brief introduction to the EKV model is mentioned.
Chapter 3 concerns low voltage, low power voltage reference circuits. After illustrating the performance specifications of such circuits, the basic bandgap principle for temperature compensation is presented and low voltage BGR circuits are shown. Since the V_{GS} of MOSFETs in weak inversion region exhibits a thermal behavior (CTAT) similar to those of BJTs' V_{BE}, several voltage reference architectures exploiting subthreshold operating MOSFETs have been proposed in the scientific literature. Furthermore, CMOS voltage references based on different temperature compensation techniques are also shown. Subthreshold operation

allowed them to work with sub-microwatt power consumption and sub-1-V supply voltages.
Subsequently, in the same chapter, the proposed voltage reference configurations are presented and discussed. They exploit an innovative temperature compensation principle based on the subthreshold operation of MOSFETs. Three electrical configurations have been implemented and fabricated in a 0.18 μm CMOS process. A detailed statistical analysis of the measurement results carried out over sets of almost 40 samples for each voltage reference architecture, is reported. Moreover, the measured performance of the best solution proposed, in terms of power consumption, minimum supply voltage and accuracy, are compared with those of low voltage, low power competitors.
Finally, Chapter 4 deals with the design of low power, low voltage op-amps. Ultra low voltage solutions exploiting techniques such as bulk-driven MOSFETs, forward body biasing and weak inversion operation are shown. The design guidelines of subthreshold op-amps are given and a design example of a two-stage Miller compensated op-amp working with just 0.5 V is presented and compared with other low power, low voltage CMOS op-amps.

2. WEAK INVERSION REGIME

2.1. Introduction

The state of weak inversion (or subthreshold conduction) at the silicon surface in a metal-oxide-silcon (MOS) structure is characterized by the fact that majority carriers have been pushed away from the silicon surface, thus leaving a depletion charge of fixed atoms. The density of minority carriers is increased with respect to the distant bulk, but it is still negligible in the overall charge balance, and does therefore not affect the capacitance-voltage curves of the MOS diode. However, these minority carriers represent the only mobile charge available at the surface. Hence, as soon as some voltage is applied between the source (S) and the drain (D) of a MOS transistor structure (see Fig. 2.1), they move by diffusion, unlike strong inversion, in which the drift current dominates.
In the following the subthreshold regime in MOSFETs is described and basic expressions for the design of the ultra low-voltage low-power circuits presented in the next chapters are shown. Therefore, particular attention has been given to the impact of the subthreshold operation in analog design.

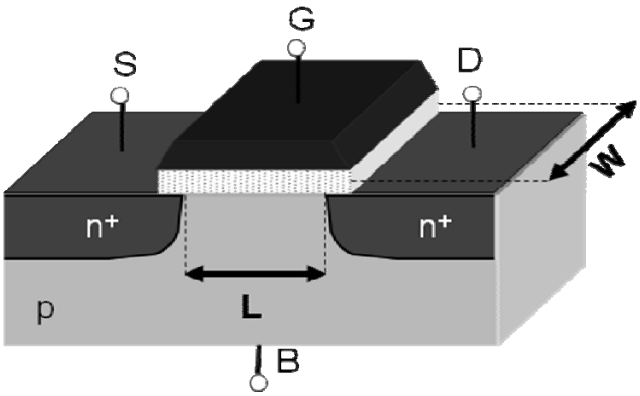


Fig. 2.1 – n-channel MOSFET structure

2.2. I-V characteristic

Depending on the voltages applied to the four terminals (Source-S, Gate-G, Drain-D and Body-B) the transistor can be biased in one of the three regions shown in Fig. 2.2. For gate-to-source voltages (V_{GS}s) higher than V_{th} (i.e. strong inversion region), the MOSFET operates in saturation or linear region if the drain-to-source (V_{DS}) voltage is higher or lower than the overdrive voltage (V_{OV}=V_{GS}-V_{th}), respectively. Subthreshold conduction is established when V_{GS}<V_{th}. The effects of the body terminal have been neglected in such description since it affects channel conduction in a similar way as the gate terminal through V_{th}-variations (frequently it is also identified as second gate or backgate).

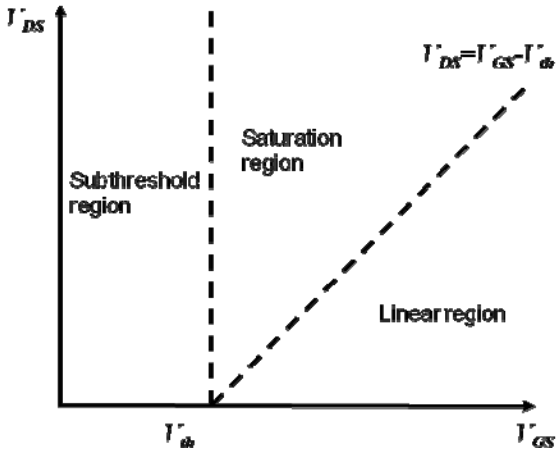


Fig. 2.2 – MOSFET operating regions

When V_{GS} is lower than V_{th}, the drain current does not abruptly approach zero but follows an exponential dependence on V_{GS} (Fig. 2.3). Fig. 2.3 also highlights the dominant component of the channel current, i.e. the diffusion component in weak inversion and the drift component in strong inversion [42]. The transition region between the previous ones is also called moderate inversion region, where drift and diffusion currents are both significant.

The source-channel-drain area forms two back to back pn diodes, they are usually reversely biased to isolate drain and source wells from the substrate. The band profile within the channel for V_{GS}=V_{DS}=0 V is shown in Fig. 2.4 (a) [43]. Obviously,

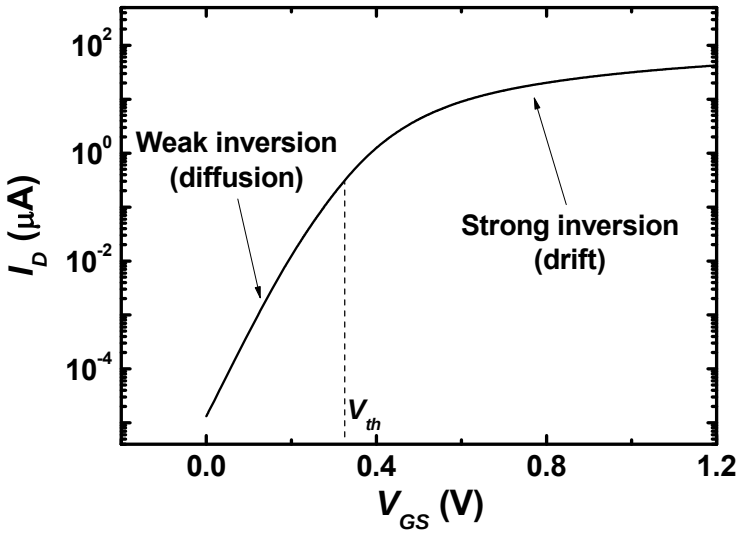
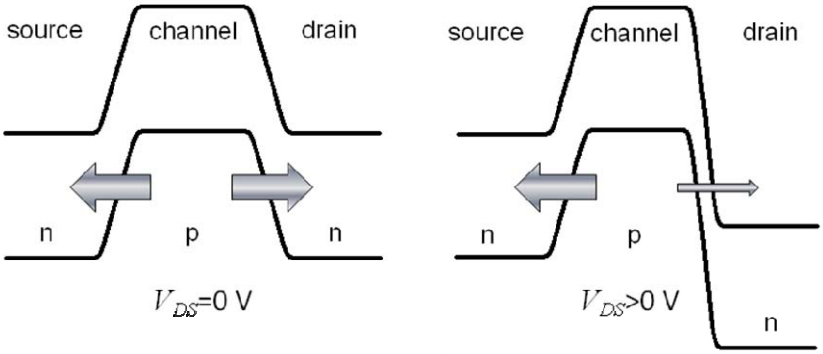


Fig. 2.3 – NMOS I_D-V_{GS} characteristic



(a) *(b)*

Fig. 2.4 – Band diagram of the longitudinal cross section of a NMOS for different values of V_{DS}.

in that case the net flux of charge through the channel is null. However, a V_{DS} increase will entirely drop across the drain-body junction, reverse biasing it more than the source-body junction, thus leading to a net current flux into the channel, Fig. 2.4 (b). Moreover, as V_{GS} increases, a fraction of it falls across the gate oxide while the remaining fraction lowers the potential barrier, i.e. increases the silicon surface potential φ_S. Variation of V_{GS} are shared between the oxide capacitance per unit area C_{OX} and the semiconductor total surface capacitance per unit area C_S [36],

$$\frac{\partial \varphi_S}{\partial V_{GS}} = \frac{C_{OX}}{C_{OX} + C_S} = \frac{1}{m}, \tag{2.1}$$

where m is the subthreshold slope parameter. C_S can be approximated with the depletion layer capacitance, $C_d = \varepsilon_{Si}/W_d$ where ε_{Si} is the Si dielectric constant and W_d is the depletion layer width. Given that $C_{OX} = \varepsilon_{OX}/t_{OX}$, where ε_{OX} is the gate oxide dielectric constant and t_{OX} is the oxide thickness, and $\varepsilon_{Si}/\varepsilon_{OX} \approx 3$, therefore m can be rewritten as,

$$m = 1 + \frac{3W_d}{t_{OX}} = 1 + \frac{3}{t_{OX}} \sqrt{\frac{2\varepsilon_{Si}\varphi_S}{qN_{ch}}}, \tag{2.2}$$

in which q is the elementary charge and N_{ch} is the channel doping concentration. In [42] a theoretical estimation of m, considering the W_d value at the onset of strong inversion, is given: it typically lies between 1.1 and 1.6.

The well known expression for subthreshold conduction, which also takes into account the V_{GS} and V_{DS} effects discussed, is [46],

$$I_D = \frac{W}{L} I_0 e^{\frac{V_{GS}-V_{th}-V_{OFF}}{mV_T}} \left(1 - e^{-\frac{V_{DS}}{V_T}}\right), \tag{2.3}$$

$$I_0 = \mu_0 V_T^2 \sqrt{\frac{q\varepsilon_{Si} N_{ch}}{2\varphi_S}}, \tag{2.4}$$

where W/L is the aspect ratio of the transistor (W and L are the channel width and length, respectively), $V_T=k_BT/q$ is the thermal voltage (k_B is the Boltzmann constant and T is the absolute temperature), I_0 is a characteristic current, it depends on temperature and mobility (μ_0) as indicated in (2.4), and V_{OFF} is the offset voltage, it determines the drain current at V_{GS}=0 V.

Similarly to a bipolar transistor, I_D changes exponentially with the control voltage V_{GS}. The exponential term in round brackets takes into account the change in current due to the reverse biasing of the drain-body junction. However, it is clear that, if V_{DS}≥3÷4V_T (75÷100 mV at room temperature), V_{DS} effect on I_D becomes negligible with respect to the current coming from the source, and the subthreshold biased MOS transistor behaves as a V_{GS}-controlled current source, as shown in Fig. 2.5. Generally, the exponential dependence of I_D on V_{GS} represents a figure of merit of MOSFETs, the subthreshold slope S. It specifies the required V_{GS} reduction to obtain an I_D decrease of a factor 10,

$$S = \left(\frac{d\left(\log_{10} I_D\right)}{dV_{GS}}\right)^{-1} \approx 2.3 \cdot V_T m, \tag{2.5}$$

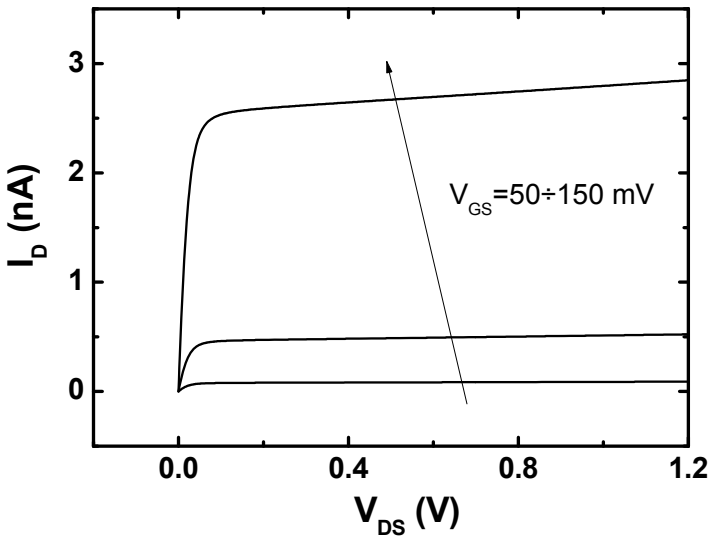


Fig. 2.5 – NMOS I_D-V_{DS} curves for different V_{GS}s

According to the possible variations of m, S can assume values in the range 60÷100 mV/decade.
Moreover, the drain current also depends on V_{DS} and the body-to-source voltage (V_{BS}) mainly because of Drain Induced Barrier Lowering (DIBL) and body effects, respectively. Such dependence of I_D on V_{DS} and V_{BS} can be taken into account by considering the following expression for V_{th} [45],

$$V_{th} = V_{th0} - \lambda_D V_{DS} - \lambda_B V_{BS} \,, \tag{2.6}$$

in which V_{th0} is the threshold voltage at room temperature extrapolated for V_{BS}=0V, λ_D and λ_B are the DIBL effect coefficient and the body effect coefficient, respectively.

2.3. Threshold Voltage

The threshold voltage is the most critical device parameter in subthreshold design. Its variation from the expected value affects exponentially the transistor I_D, unlike strong inversion region where its influence is quadratic or linear. Primarily, V_{th} deviations are due to temperature and process variations.
In a MOS system (p-type semiconductor), V_{th} is the voltage applied between the metal and the semiconductor at which the electron concentration at the semiconductor surface (i.e. the conducting channel) equals the hole concentration in the bulk semiconductor, that is when φ_S reaches the strong inversion value $2\varphi_B = V_T ln(N_{ch}/n_i)$ (ni is the intrinsic carrier density). Indicating with Q_D the charge per unit area stored in the depletion region, at the onset of strong inversion and with V_{fb} the flat band voltage to align the gate potential and the body potential, V_{th} can be written as [42],

$$V_{th} = V_{fb} + 2\varphi_B + \frac{Q_D}{C_{OX}} = -\frac{E_g}{2q} + \frac{Q_{OX}}{C_{OX}} + \varphi_B + \frac{\sqrt{4\varepsilon_{Si} q N_{ch} \varphi_B}}{C_{OX}} \,, \tag{2.7}$$

where E_g is the energy bandgap of the semiconductor and Q_{OX} the equivalent charge per unit area stored at the interface oxide-semiconductor.
Process variations affect the absolute (among different dies) and relative (within the same die) V_{th} precision. Apart from the relative uncertainties, that can be reduced by means of careful circuit layout techniques, Die-to-Die variations, that cause V_{th} dispersion because of changes in parameters such as Q_{OX}, N_{ch} and φ_B, are difficult to compensate and commonly represent the main cause of inaccuracy in analog circuits [19].
Temperature has a remarkable influence on V_{th} too, mainly through E_g and φ_B variations,

$$\begin{aligned} \frac{\partial E_g}{\partial T} &\approx -2.73\times10^{-4} \text{ eV/°K for Si near 300°K} \\ \frac{\partial \varphi_B}{\partial T} &= \frac{\partial}{\partial T}\left[V_T \ln\left(\frac{N_{ch}}{n_i} \right) \right] \end{aligned} \,, \tag{2.8}$$

Since $n_i(T)$ in Si increases exponentially with temperature for T>300 °K, φ_B also decreases when temperature increases. For example, considering a process with N_{ch}~10^{16} cm^{-3} and m=1.1, $\partial V_{th}/\partial T$=1 mV/°K, that is a ΔV_{th}=150 mV in a temperature range -25÷125 °C, very common in modern CMOS technologies.
Although subthreshold operation offers several advantages in the ultra low-power, low-voltage scenario, it also poses challenges for the designers to devise and employ circuit compensation techniques.

2.4. Small signal equivalent circuit

In order to start a low power low voltage analog design activity with circuits employing subthreshold operated transistors a good understanding of DC parameters, frequency response and noise performances of MOSFET in subthreshold region is needed. Moreover, in this paragraph a comparison between the behavior of saturated MOSFETs in weak (V_{DS}>3÷4V_T) and strong inversion is presented.
To this aim, the small-signal equivalent MOSFET model shown in Fig. 2.6 is considered. In the model, g_m is the transconductance, r_d is the drain-source resistance, g_{mb} is the substrate transconductance. C_{gs}, C_{gd}, C_{gb}, C_{sb}, C_{db} are the gate-to-source, gate-to-drain, gate-to-bulk, source-to-bulk and drain-to-bulk capacitances, respectively. Aiming to calculate the small-signal MOSFET parameters (g_m, r_d and g_{mb}), I_D can be conveniently rewritten as follows according to (2.3) and (2.6), considering a V_{DS}>3÷4V_T, i.e. neglecting the term in round brackets in (2.3):

$$I_D = \beta e^{\frac{V_{GS}}{mV_T}} e^{\frac{\lambda_D V_{DS}}{mV_T}} e^{\frac{\lambda_B V_{BS}}{mV_T}}$$
$$\beta = \frac{W}{L} I_0 e^{-\frac{V_{th0}+V_{OFF}}{mV_T}} \quad , \qquad (2.9)$$

From (2.9), g_m, r_d, g_{mb} and the intrinsic MOSFET DC gain (A_{V0}= $g_m r_d$) can be derived as reported in TABLE 2.1. The table also summarizes the values of g_m, r_d, g_{mb} and A_{V0} in saturation region (in the expressions shown in the table, K=$\mu_n C_{OX}$ is the process transconductance and λ is the channel length modulation coefficient).
In subthreshold regime g_m linearly depends on I_D, while it increase with $\sqrt{I_D}$ in saturation region. This leads to a greater g_m/I_D ratio in subthreshold than in

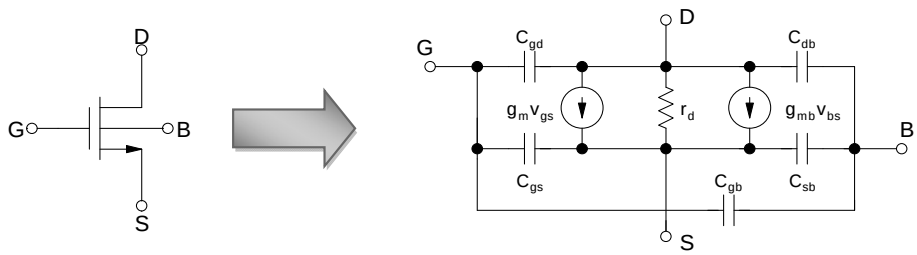


Fig. 2.6 - Small signal equivalent circuit for a NMOSFET.

TABLE 2.1
SMALL-SIGNAL MOSFET PARAMETERS

	Subthreshold	Strong inversion
$g_m = \dfrac{\partial I_D}{\partial v_{GS}}$	$\dfrac{I_D}{mV_T}$ (2.10a)	$\sqrt{\dfrac{2KWI_D}{L}}$ (2.10b)
$r_d = \left[\dfrac{\partial I_D}{\partial v_{DS}}\right]^{-1}$	$\dfrac{mV_T}{\lambda_D I_D}$ (2.10c)	$\dfrac{1}{\lambda I_D}$ (2.10d)
$g_{mb} = \dfrac{\partial I_D}{\partial v_{BS}}$	$\dfrac{\lambda_B I_D}{mV_T}$ (2.10e)	$g_m \dfrac{\gamma}{2\sqrt{2\lvert\varphi_F\rvert + \lvert V_{SB}\rvert}}$ (2.10f)
$A_{V0} = g_m r_d$	$\dfrac{1}{\lambda_D}$ (2.10g)	$\dfrac{1}{\lambda}\sqrt{\dfrac{2KW}{I_D L}}$ (2.10h)

saturation region or, similarly, the g_m value is greater in a subthreshold biased MOSFET for a fixed I_D (see Fig. 2.7 [44]). Another important difference between weak and strong inversion operation is related to the intrinsic MOSFET DC gain. In saturation region, A_{V0} depends on the drain current, the transistor aspect ratio, the process transconductance and the channel length modulation parameter λ, whereas it only depends on the DIBL effect coefficient λ_D in subthreshold region

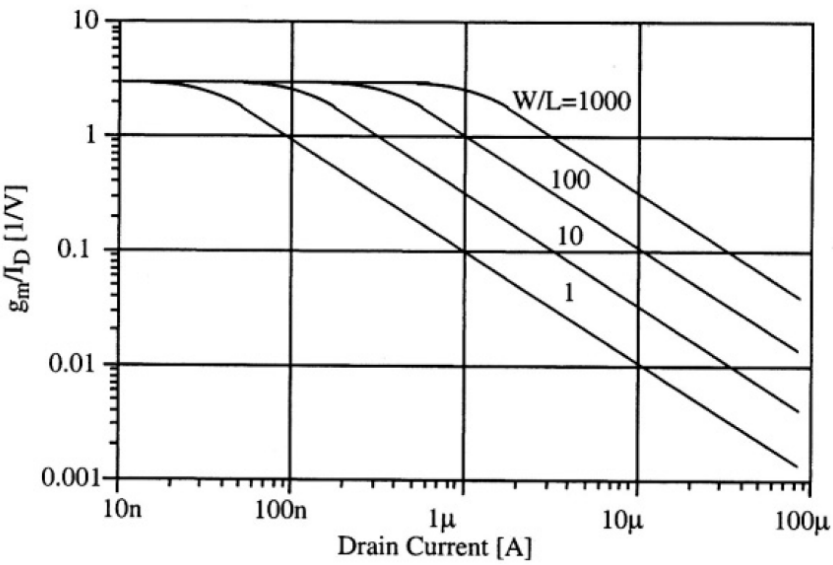


Fig. 2.7 – g_m versus I_D in weak and strong inversion.

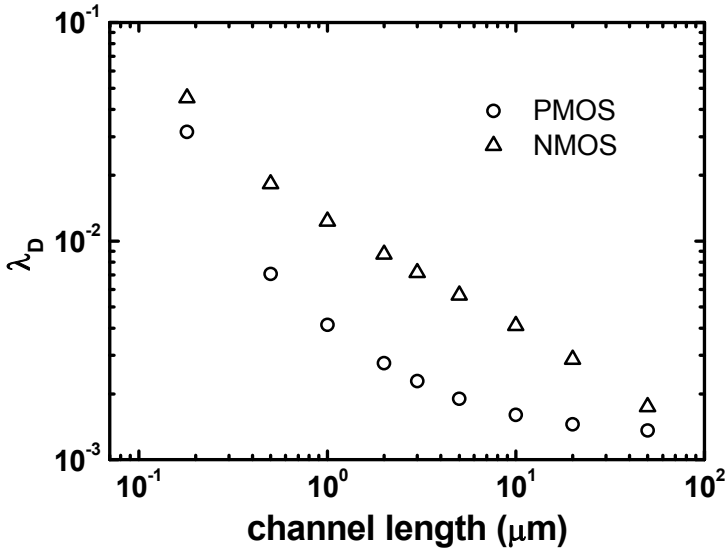


Fig. 2.5 –NMOS and PMOS DIBL coefficients of the UMC 0.18 μm technology

(TABLE 2.1). Since DIBL effect is strongly related to the distance between drain and source wells, λ_D (and consequently A_{V0}) is mainly controlled by MOSFET channel length. In order to have a first estimation of the minimum transistor lengths required to achieve a desired DC gain, the λ_D-L characteristic shown in Fig. 2.8 was extracted for the technology UMC RF/MIXED-MODE 0.18 μm. To this aim, $\lambda_D = \left[\frac{\partial I_D}{\partial V_{GS}}\right]^{-1} \frac{mV_T I_D}{V_{DS}}\Bigg|_{V_{BS},V_{GS},V_{DS}}$ was derived from (2.10c), where the subthreshold slope parameter was evaluated as $m = \frac{\partial I_D}{\partial V_{GS}} \frac{V_{GS}}{V_T I_D}\Bigg|_{V_{BS},V_{DS},V_{GS}}$, from (2.10a). The λ_D and m values were extracted from SPECTRE simulations for devices with W=10 μm, L= 2 μm biased at V_{BS}=0 V and V_{GS}=V_{DS}=200 mV. It is worth noting that the threshold voltages of the CMOS process considered are 320 mV and 456 mV for NMOS and PMOS transistors, respectively.

To have a quantitative idea of dynamic performances of subthreshold operated MOSFETs, the intrinsic transition frequency f_{Ti} is considered.

$$f_{Ti} = \frac{g_m}{2\pi C_X} = \begin{cases} \frac{1}{2\pi}\frac{I_D}{mV_T}\frac{C_d + C_{OX}}{WL \cdot C_d C_{OX}}, & \text{weak inversion} \\ \frac{1}{2\pi}\sqrt{\frac{2KWI_D}{L}}\frac{3/2}{WL \cdot C_{OX}}, & \text{strong inversion} \end{cases}, \qquad (2.11)$$

In (2.11) the generic expressions for the f_{Ti} of saturated MOSFET in weak and strong inversion are given [43], where C_X is the gate input capacitance (C_X=C_{GS}+C_{GD}+C_{GB} (Fig. 2)). The values of C_X in subthreshold (C_{XSUB}) and strong inversion (C_{XSAT}) regions are given by [43],

$$\begin{cases} C_{X\,SUB} = \left(C_{GSO} + C_{GDO}\right) + WL\dfrac{C_d C_{OX}}{C_d + C_{OX}} \simeq WL\dfrac{C_d C_{OX}}{C_d + C_{OX}} \\ C_{X\,SAT} = \left(C_{GSO} + C_{GDO}\right) + \dfrac{2}{3}WLC_{OX} \simeq \dfrac{2}{3}WLC_{OX} \end{cases}, \qquad (2.12)$$

where C_{GSO} and C_{GDO} are the gate-source and gate-drain overlap capacitances, respectively, and C_d is the depletion region capacitance between gate and body per unit area. It follows C_{XSUB}=1.5 C_{XSAT} at the most (for $C_d \rightarrow \infty$).

From the comparison between f_{Ti}s in weak and strong inversion regions it clearly follows that, because of the different g_m/I_D ratios in the two regions, f_{Ti} linearly depends on I_D in subthreshold, whereas it depends on $\sqrt{I_D}$ in strong inversion. As it appears from (2.11), the transistor speed is strongly degraded by the channel area (WL) in both operating regions. Therefore, considering the correlation between A_{V0} and L in weak inversion, a trade-off should occur in the choice of transistors channel length.

Finally, the noise performances are considered. In subthreshold region, the MOSFET drain current exhibits a white noise part (S_{iw}=$2qI_D$) caused by shot noise [43]. The power spectral density of the input-referred noise voltage is given by the following expressions,

$$S_{vw} = \frac{S_{iw}}{g_m^2} = \begin{cases} \dfrac{2qI_D}{g_m^2} = \dfrac{2q}{I_D}\left(mV_T\right)^2, & \text{weak inversion} \\ \dfrac{8k_BT}{3g_m} = \dfrac{8k_BT}{3}\sqrt{\dfrac{L}{2KWI_D}}, & \text{strong inversion} \end{cases}, \qquad (2.13)$$

As shown, S_{vw} depends on $1/I_D$ and $1/\sqrt{I_D}$ in weak and strong inversion regions, respectively. However, considering a single MOSFET, the output noise voltage amplitude is given by the product of the drain noise current amplitude ($i_{Dn\_flicker}$) and the output resistance (r_D). At low frequencies $i_{Dn\_flicker}$ ($\mathrm{A}/\sqrt{\mathrm{Hz}}$) in subthreshold regime is proportional to the drain current level (I_D) while the output resistance (r_D) of a MOSFET is proportional to $1/I_D$. Therefore the low current level does not intrinsically increases the low frequency noise. At high frequencies the shot noise current is proportional to $\sqrt{I_D}$, in that case reducing I_D the output noise voltage will increase with $\sqrt{I_D}$. However at high frequencies it becomes easy to filter noise effects.

From the above discussion it results that owing to the different g_m/I_D behavior in the two operating regions, subthreshold operation allows one to achieve a better trade-

off among frequency response, noise performances and bias current of transistors compared with strong inversion.

3. LOW-VOLTAGE, LOW-POWER VOLTAGE REFERENCES

3.1. Introduction

Ultra low power, low voltage design requirements are crucial in emerging applications such as portable medical devices, microsensor nodes and passive RFIDs. For such portable systems, power consumption reduction becomes essential to extend the battery lifetime and/or the communication range. This leads to a strong demand for circuit building blocks operating with low supply voltages and low power consumption.

Among them, Voltage Reference circuits are ubiquitous, they are broadly used in analog and digital systems to generate a DC voltage independent of Process, supply Voltage and Temperature (PVT) variations. Regarding low power analog design, as stated in chapter 1, the power dissipated by bias circuitry, such as voltage references, is not useful to maintain neither an adequate bandwidth nor a required SNR, in an analog system. Therefore several solutions have been presented so far to achieve sub-microwatt power consumption and/or sub-1-V supply voltage.

Voltage reference circuits are widely used in Analog to Digital Converters (ADCs) [58], see Fig. 3.1. A voltage regulator locks the output voltage (V_{OUT}) to the reference voltage generated by means of a resistor ratio,

$$V_{OUT} = \frac{R_2 + R_1}{R_2} V_{REF} \,. \tag{3.1}$$

Actually it is a two-stage feedback amplifier with a reference voltage as an input. The first stage is an op-amp and the second stage is a source follower. Depending on its aspect ratio, this source follower can deliver a large current to the ADC. The supply voltage, V_{DD}, usually contains a ripple that is suppressed by the regulator. The accuracy of the output voltage, sent to the ADC, depends on both accuracies of the resistor ratio and the reference voltage (V_{REF}). The resistor ratio can have a smaller error than 0.1% [59]. As a result, the final accuracy will depend on the absolute error of V_{REF}.

The most common solution for on-chip integration is the Bandgap Voltage Reference (BGR), which can be implemented in standard CMOS technology exploiting parasitic vertical BJTs. Conventional BGRs generate a nearly temperature independent reference, of about 1.25 V, and therefore they require a higher supply voltage, which might not meet the low-voltage constraints for low-power applications. However, several solutions, exploiting the BGR principle, have been implemented that ensure sub-1V operation [47]. In [24, 48, 49] the reference voltage is lowered by means of resistive subdivision. As an alternative approach, floating gate structures have been used to realize high precision programmable voltage references, as in the case of [50], where subthreshold operation allowed a minimum power consumption of 1.35 μW. Most often forward biased PN-junctions of BGRs are substituted with MOSFETs biased in the

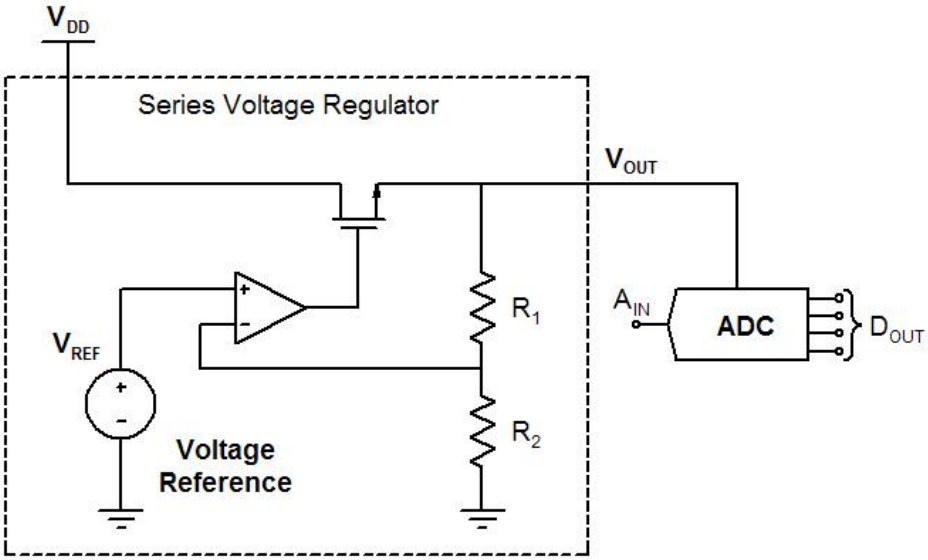


Fig. 3.1 – Voltage reference system for an ADC.

subthreshold region [51-53]. Thus, a supply voltage (V_{DD}) of 0.6 V and a power consumption of 9 μW are achieved. Other solutions have been implemented in standard CMOS technology, without exploiting the traditional BGR principle [18, 54-57]. Subthreshold operation allowed sub-1μW power consumption [18, 55-57] and sub-1V operation [18,57].

In this section the typical parameters of performance and figures of merit of voltage reference architectures are described. The operating principles of BGR and sub-1-V configurations are shown. Subsequently it is presented how subthreshold operation has been successfully exploited in low voltage, low power voltage reference circuits. At first, circuit solutions exploiting the BGR principle by means of subthreshold operated MOSFETs, instead of bipolar transistor, are presented. Then ultra low power, low voltage architectures working in weak inversion regime, without exploiting the BGR principle, are shown. Afterwards the proposed CMOS voltage reference configurations operating with all transistors biased in subthreshold region are introduced. Their novel operating principle and circuit architectures are described in detail. Then design considerations are discussed and, finally, experimental results are shown and compared with low voltage, low power competitors found in the scientific literature.

3.2. Performance specifications

A voltage reference is an analog block that provides a stable DC voltage, V_{REF}, starting from a less stable supply voltage V_{DD}. Apart from general key performance indicators of integrated circuits such as power consumption and Si area occupation, voltage reference circuits typically have stringent accuracy requirements. Usually, deviations from the expected value of the generated reference voltage are mainly due to temperature and supply voltage variations, process variations, spurious signal from power lines and noise generated by active devices within the voltage reference circuit.

3.2.1. Line sensitivity

The line sensitivity (LS) is the capability of the circuit to maintain a stable output despite V_{DD} variations. A ΔV_{REF} change resulting from a variation of the supply voltage of ΔV_{DD} gives a line sensitivity of,

$$LS = \frac{\Delta V_{REF}}{\Delta V_{DD}}. \tag{3.2}$$

It is usually expressed as mV/V or μV/V. An alternative definition is based on the percentage change of V_{REF},

$$LS(\%) = 100 \frac{\Delta V_{REF} / V_{REF}}{\Delta V_{DD}}, \tag{3.3}$$

expressed as %/V.

3.2.2. Temperature coefficient

The temperature coefficient (TC) gives a measure of the ability of the voltage reference to maintain the expected V_{REF} under varying thermal conditions. Similarly to LS, it is defined in two forms,

$$TC = \frac{\Delta V_{REF}}{\Delta T} \quad [\text{mV/°C or } \mu\text{V/°C}], \tag{3.4}$$

$$TC(\%) = 10^6 \frac{\Delta V_{REF} / V_{REF}}{\Delta T} \quad [\text{ppm/°C}]. \tag{3.5}$$

Where ΔT express the change in temperature which has the consequent ΔV_{REF} change.

3.2.3. Power Supply Rejection Ratio

The Power Supply Rejection Ratio expresses the capability of the circuit to reject noise coming from power supply. It is evaluated as the small signal ratio between the output voltage and the power supply,

$$PSRR = 20 \log\left(\frac{v_{ref}}{v_{dd}}\right) \quad [\text{dB}], \tag{3.6}$$

3.2.4. Other specifications

Other performance indicators are the equivalent output noise generated by the circuit over frequency, the process variation effects that cause a dispersion of the V_{REF} value and the long term stability drift. Namely the latter measures the V_{REF} variation over a long period of time at some specified condition of steady state operation.

3.3. Bandgap References (BGRs)

Bandgap reference voltage circuits are one of the most widely used analog circuits in memories, ADCs and power management circuits. They exhibit very little dependence on temperature and, since most process parameters vary with

temperature, they also show a remarkable independence from process variations. In order to generate a voltage independent of temperature variations, a BGR adds two quantities having opposite TCs with proper weighting. For example, if voltages V_1 and V_2 vary in opposite directions with temperature, α_1 and α_2 can be chosen such that,

$$\alpha_1 \frac{\partial V_1}{\partial T} + \alpha_2 \frac{\partial V_2}{\partial T} = 0 , \tag{3.7}$$

and, therefore, the resulting reference voltage

$$V_{REF} = \alpha_1 V_1 + \alpha_2 V_2 , \tag{3.8}$$

will show a zero TC. The following paragraph describes the principle of operation of BGR.

3.3.1. Basic principle

Complementary To Absolute Temperature (CTAT) and Proportional To Absolute Temperature (PTAT) behaviors were firstly observed in bipolar transistors.
For a bipolar devices the collector current can be written as,

$$I_C = I_S e^{V_{BE}/V_T} , \tag{3.9}$$

where V_{BE} is the base-emitter voltage and the saturation current, I_S, is proportional to $\mu k_B T n_i^2$, in which μ is the mobility of minority carriers and n_i the intrinsic minority carrier concentration in Si. The temperature dependence of such quantities is,

$$\begin{aligned} &\mu \propto \mu_0 T^m \quad \text{with} \quad m \approx -3.2 \\ &n_i^2 \propto T^3 e^{-E_g/k_B T} \end{aligned} , \tag{3.10}$$

Thus, I_S can be rewritten as,

$$I_S = b T^{4+m} e^{-E_g/k_B T} , \tag{3.11}$$

where b is a proportionality factor. Evaluating V_{BE} from (3.9) and assuming I_C constant with temperature, it is possible to compute the TC of V_{BE} as,

$$\frac{\partial V_{BE}}{\partial T} = \frac{\partial V_T}{\partial T} \ln \frac{I_C}{I_S} - \frac{V_T}{I_S} \frac{\partial I_S}{\partial T} . \tag{3.12}$$

Combining (3.11) and (3.12) it is possible to obtain,

$$\begin{aligned} \frac{\partial V_{BE}}{\partial T} &= \frac{V_T}{T} \ln \frac{I_C}{I_S} - (4+m) \frac{V_T}{T} - \frac{E_g}{k_B T^2} V_T = \\ &= \frac{V_{BE} - (4+m) V_T - E_g/q}{T} \end{aligned} . \tag{3.13}$$

Considering that $E_g/q \approx 1.12$ V, with $V_{BE} \approx 750$ mV and T=300 °K, the TC of V_{BE} is -1.5 mV/°K (i.e. CTAT).

A PTAT voltage, instead, can be generated by the difference between two V_{BE}s of BJTs operated at different currents. If two identical BJTs are biased with collector currents I_1 and I_2=nI_1, then, from (3.9), it is possible to write,

$$\Delta V_{BE} = V_{BE2} - V_{BE1} = V_T \left(\ln \frac{nI_1}{I_S} - \ln \frac{I_1}{I_S} \right) = V_T \ln(n) . \tag{3.14}$$

The resulting positive temperature coefficient is,

$$\frac{\partial \Delta V_{BE}}{\partial T} = \frac{k_B}{q} \ln(n) \approx \ln(n) 0.087 \text{ mV/°K} . \tag{3.15}$$

Therefore a reference voltage with nominally zero temperature coefficient can be written as, recalling (3.7),

$$V_{REF} = \alpha_1 V_{BE} + \alpha_2 V_T \ln(n) , \tag{3.16}$$

Since $\partial V_{BE}/\partial T$ and k_B/q are fixed quantities, it follows that the α_1 and α_2 values for zero TC, lead to a V_{REF} value of about 1.25 V.
A simple circuit that adds the PTAT and CTAT voltages discussed previously is shown in Fig. 3.2 [59]. BJTs $Q2$ and $Q1$*(pnp)* have an emitter area ratio of n, PMOS mirror current imposes the same current in the two branches of the voltage reference, while the op-amp guarantees same voltages at its positive and negative inputs. The current flowing in the two resistors is PTAT, since it depends on the difference between the base emitter voltages of $Q2$ and $Q1$: I=$\Delta V_{EB}/R_{PT}$. As a result, V_{REF} can be easily expressed as,
xxx
BGR circuits can be integrated in CMOS technology by exploiting parasitic vertical BJTs that form themselves naturally on the substrate. In Fig. 3.3 it shown a pnp transistor in a n-well CMOS technology. A p^+ region (the same as the S/D region of

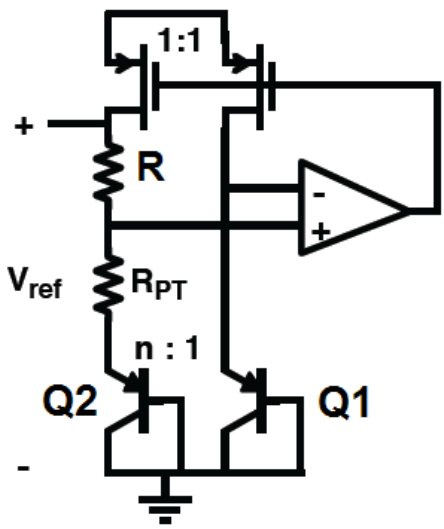


Fig. 3.2 – Example of Bandgap Reference circuit.

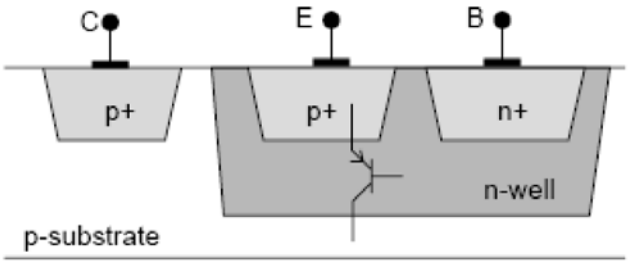


Fig. 3.3 – Realization of a pnp transistor in CMOS technology.

PMOSs) inside a n-well serves as the emitter and the n-well as the base. The

$$V_{REF} = V_{EB1} + V_T \ln(n) \frac{R}{R_{PT}} \quad (3.17)$$

p-type substrate acts as the collector.

3.3.2. Sub-bandgap reference circuits

As process technologies go into deep sub-micron eras and the demand for battery-operated portable equipment increases, voltage supply has to be scaled down. Low V_{DD} requires new circuit design technique for designing BGRs because of the high traditional bandgap reference voltage of 1.25 V.

One of the first portable applications requiring sub-1-V operation for enabling significant energy minimization was the digital hearing IC presented in [24]. The complete IC consumes 9 mW and is supplied with a single cell zinc-air or silver-oxide battery. A representative graph of the open circuit battery voltage (or electromotive force, emf) versus time is shown in Fig. 3.4. The unloaded terminal voltage reaches 1.1 V at the end of lifetime. The minimum battery lifetime required for hearing aid equipments is a one week, 16-hours-a-day usage. For most batteries this limits the supply current to a maximum of 1.5 mA. The internal battery

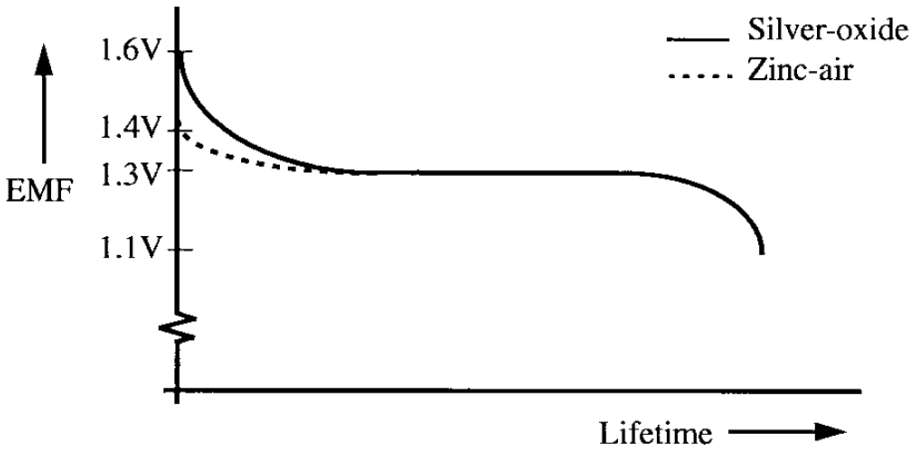


Fig. 3.4 – Open terminal voltage versus lifetime.

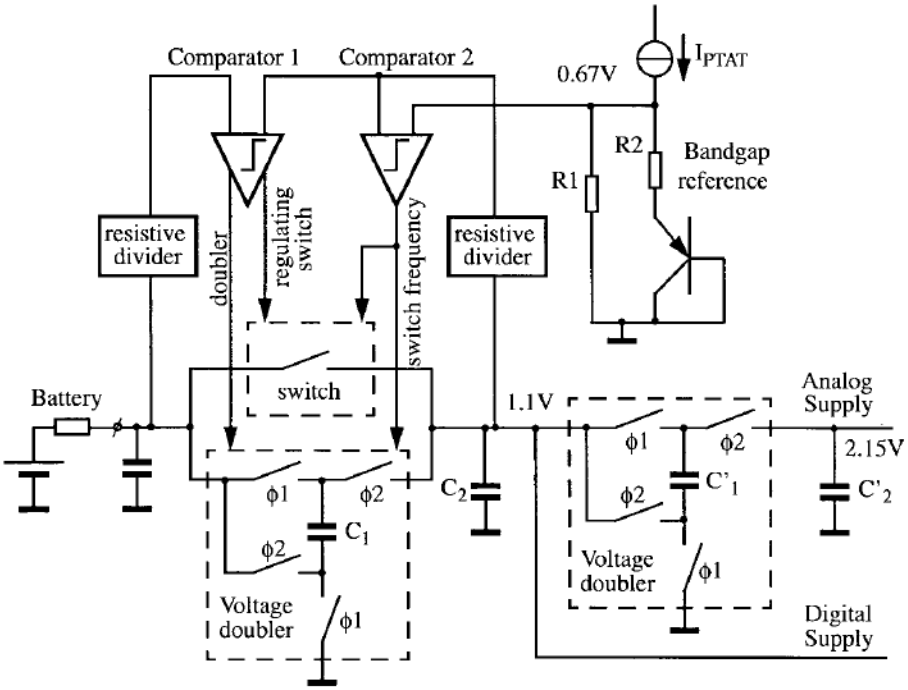


Fig. 3.5 – Power management section of a hearing instrument IC.

resistance can vary in the range 10÷40 Ω. Moreover, due to the peak currents of several mA needed by the earphone at high output, the terminal voltage can temporarily fall to 0.9 V.

A DC/DC conversion scheme is provided to proper supply analog and digital sections, as shown in Fig. 3.5. A switched-mode voltage regulator generates the 1.1 V digital supply voltage. It consists of a regulating switch, a capacitive voltage doubler, two comparators, two resistive dividers and a bandgap reference. For the analog circuitry, an additional capacitive voltage doubler is needed to provide a voltage source of 2.15 V. Comparator 2 is part of a feedback loop which sets the digital supply voltage. Such voltage is compared to a bandgap referred voltage at a 512-kHz repetition frequency. If the digital supply voltage is too low, the comparator generates a 1 μs pulse for either the regulating switch or the voltage doubler. Comparator 1 monitors the battery voltage to determine if the switch or the doubler is to be activated in the feedback loop. The doubler is chosen when the battery voltage drops below the desired digital supply voltage, either because of peak current consumption or when the battery approaches end of lifetime.

A bandgap reference circuit is needed as a voltage control generator for the whole power management section. It is supplied directly from the battery and, therefore, has to guarantee an accurate output voltage from a worst case minimum VDD of 0.9 V. Consequently, a traditional BGR generating an output of 1.25 V is inadequate. The voltage reference proposed in [24] (Fig. 3.6) solves such problem by means of resistive subdivision of the BGR voltage. The PTAT current $I_{PTAT}=V_T/R_0 ln(n)$ is generated by the two left branches of the circuit and injected in the right branches that generate the reference voltage as,

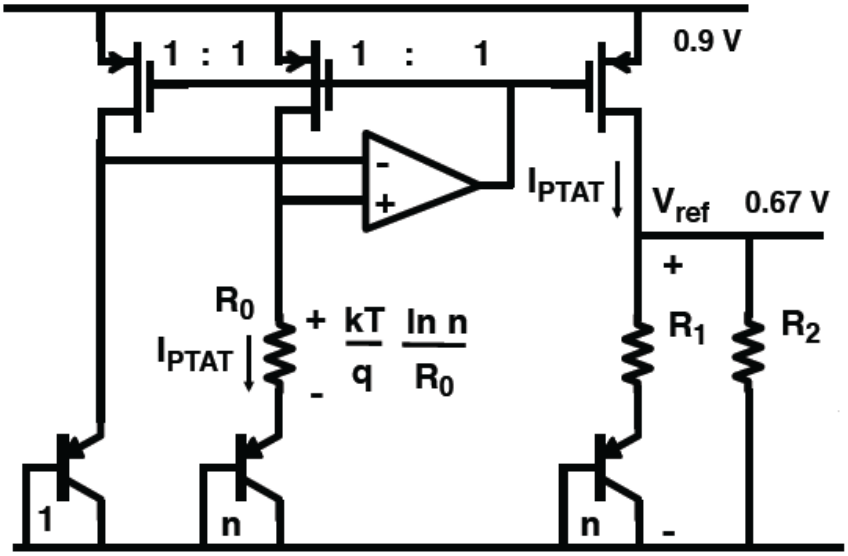


Fig. 3.6 – CMOS sub-Bandgap Reference working with V_{DD} down to 0.9 V.

$$V_{REF} = \frac{R_2}{R_2 + R_1}\left(V_{EB1} + R_1 I_{PTAT}\right) = \frac{R_2}{R_2 + R_1}\left(V_{EB1} + \frac{R_1}{R_0} V_T \ln(n)\right) \quad (3.18)$$

The result is a resistive division of a BGR voltage, i.e. the quantity in round brackets is the same reference voltage found in (3.17). Such circuit configuration achieved an output voltage lower than 1.25 V (0.67 V) and, hence, it permitted to

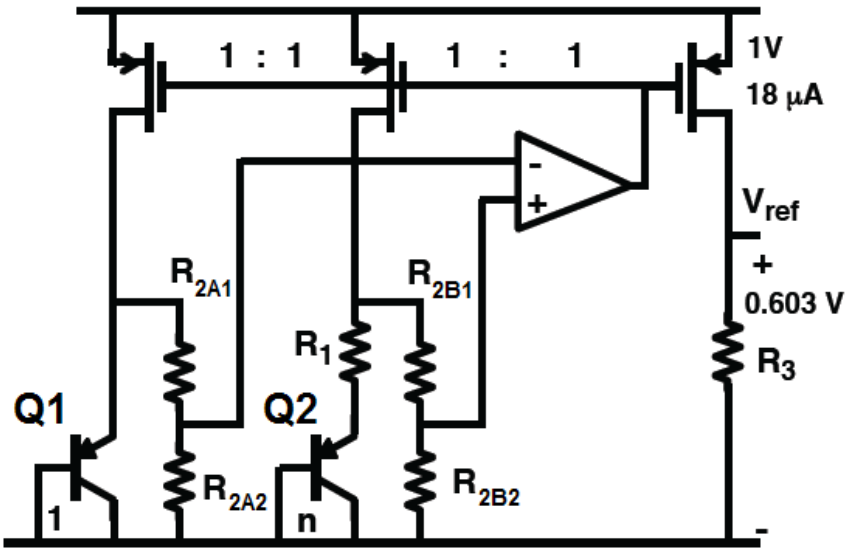


Fig. 3.7 – CMOS sub-Bandgap Reference working with V_{DD} <1 V.

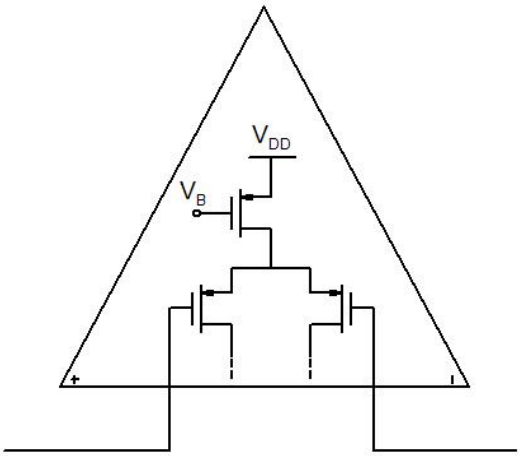


Fig. 3.8 – Op-amp input section: PMOS differential pair.

scale V_{DD} down to 0.9 V.

Another sub-bandgap solution, fabricated in a 0.6-μm n-well CMOS process is the one presented in [48], Fig. 3.7. Again, the op-amp is used to equalize its input voltages. Moreover, if R_{2A1}=R_{2B1}, R_{2A2}=R_{2B2} and the emitter area ratio of Q1 and Q2 is n, the current mirrored by PMOSs is,

$$I = \frac{V_{EB2}}{R_2} + \frac{V_T \ln(n)}{R_1}. \tag{3.19}$$

where R_2=R_{2A1}+ R_{2A2}=R_{2A1}+ R_{2A2}. Therefore the scaled bandgap reference voltage is,

$$V_{REF} = \frac{R_3}{R_2}\left[V_{EB2} + \frac{R_2}{R_1} V_T \ln(n)\right]. \tag{3.20}$$

The minimum V_{DD} (V_{DDmin}) is not limited by the BGR voltage anymore. It can be evaluated by considering the input section of the op-amp in Fig. 3.8. It consists in a PMOS differential pair, therefore V_{DD} is the sum of the V_{SD} of the PMOS mirror device, the PMOS V_{SG} of the differential pair and the input potential of the op-amp. Saturated devices in strong inversion are used in [48], consequently, from Fig. 3.7, V_{DDmin} can be written as,

$$\begin{aligned} V_{DD\min} &= \left(\frac{R_{2B2}}{R_{2B2}+R_{2B1}}\right) V_{EB2} + V_{SG\min} + V_{SDsat} = \\ &= \left(\frac{R_{2B2}}{R_{2B2}+R_{2B1}}\right) V_{EB2} + \left|V_{thp}\right| + 2V_{SDsat} \end{aligned} . \tag{3.21}$$

Forward body biasing of the source-body junction of the op-amp PMOSs has been used to lower their V_{thp} (≈0.9 V without body biasing) and, hence, the minimum supply voltage. Indeed, the V_{DDmin} achieved is 0.98 V. A remarkable insensitivity to

temperature effects were also obtained, in fact, the temperature coefficient in the temperature range 0÷100 °C is just 15 ppm/°C.

3.4. Subthreshold CMOS voltage references

Forward biased pn-junctions are not the only solution for the implementation of voltage references. In particular, several circuit architectures exploited the BGR principle by using MOSFETs operating in weak inversion. Indeed, they exhibit a CTAT behavior, similar to the V_{BE} of BJTs, when biased with a fixed current.
Nevertheless, the BGR principle is not the only starting point in the design of voltage references, in strong and weak inversion.
In both cases, however, subthreshold operation has resulted in sub-microwatt power consumption and lower supply voltage when compared with traditional BGRs.

3.4.1. BGR-principle and weak inversion

One of the first voltage reference topologies based on subthreshold MOSFETs has been presented in [52]. It exploits the fact that the V_{GS} of a subthreshold operated MOSFET, after biasing with a constant drain current, decreases with temperature. Assuming sufficient long channel length and a V_{DS}>3÷4V_T, an expression for I_D in weak inversion regime, equivalent to (2.3), is[36],

$$I_D = \frac{W}{L} I_0 e^{\frac{\varphi_S - 2\varphi_B}{V_T}} , \tag{3.22}$$

Therefore, equating (2.3) and (3.22),

$$\frac{V_{GS} - V_{th} - V_{OFF}}{m} = \varphi_S - 2\varphi_B , \tag{3.23}$$

Thermal behaviors of φ_S and φ_B are, respectively,

$$\begin{aligned} \varphi_S &= \varphi_S(T_0)\frac{T}{T_0} - 3V_T \ln\left(\frac{T}{T_0}\right) + \frac{E_g(T)}{q} - \frac{T}{T_0}\frac{E_g(T_0)}{q} \\ \varphi_B &= \varphi_B(T_0)\frac{T}{T_0} - 3V_T \ln\left(\frac{T}{T_0}\right) + \frac{E_g(T)}{2q} - \frac{T}{T_0}\frac{E_g(T_0)}{2q} \end{aligned} , \tag{3.24}$$

where T_0 is the room temperature. Hence the difference φ_S-2φ_B, as a function of temperature, can be rewritten as,

$$\varphi_S(T) - 2\varphi_B(T) = \left[\varphi_S(T_0) - 2\varphi_B(T_0)\right]\frac{T}{T_0} . \tag{3.25}$$

Consequently, from (3.23) it follows that,

$$V_{GS}(T) = V_{th}(T) + V_{OFF} + \frac{m(T)}{m(T_0)}\frac{T}{T_0}\left[V_{GS}(T_0) - V_{th}(T_0) + V_{OFF}\right] . \tag{3.26}$$

Assuming that the subthreshold slope factor has small variations with temperature (i.e. $n(T) \approx n(T_0)$) and modeling the threshold voltage thermal behavior as,

$$V_{th}(T) = V_{th}(T_0) + K_T\left(\frac{T}{T_0} - 1\right), \tag{3.27}$$

where K_T is a coefficient <0, V_{GS} can be expressed as,

$$V_{GS}(T) = V_{GS}(T_0) + K_G\left(\frac{T}{T_0} - 1\right), \tag{3.28}$$

in which,

$$K_G = K_T + V_{GS}(T_0) - V_{th}(T_0) - V_{OFF}\,. \tag{3.29}$$

For typical values of K_T, V_{OFF} and V_{GS}-V_{th} in weak inversion, K_G is a negative quantity and, therefore, $\partial V_{GS}/\partial T$<0 (typically -1 mV/°C) for any fixed I_D.
The subthreshold voltage reference proposed in [52] is shown in Fig. 3.9. In the right section of the dotted line, the circuit configuration sets the current I_{R1} to be dependent on V_{GS1},

$$I_{R1} = \frac{V_{GS1}}{R_1}\,. \tag{3.30}$$

This CTAT current is then mirrored by M_5 and M_6 in the next circuit subsection. The output voltage is,

$$V_R = R_4 I_{R4} + V_{R3} = R_4\left(\frac{V_{R3}}{R_3} - I_{R1}\frac{(W/L)_6}{(W/L)_4}\right) + V_{R3}\,, \tag{3.31}$$

where,

$$V_{R3} = \frac{(W/L)_5}{(W/L)_4}\frac{R_2}{R_1}V_{GS1} + V_T \ln\left(\frac{(W/L)_8}{(W/L)_7}\frac{(W/L)_5}{(W/L)_6}\right). \tag{3.32}$$

Combining (3.30), (3.31) and (3.32), it is possible to obtain an expression for V_R similar to,

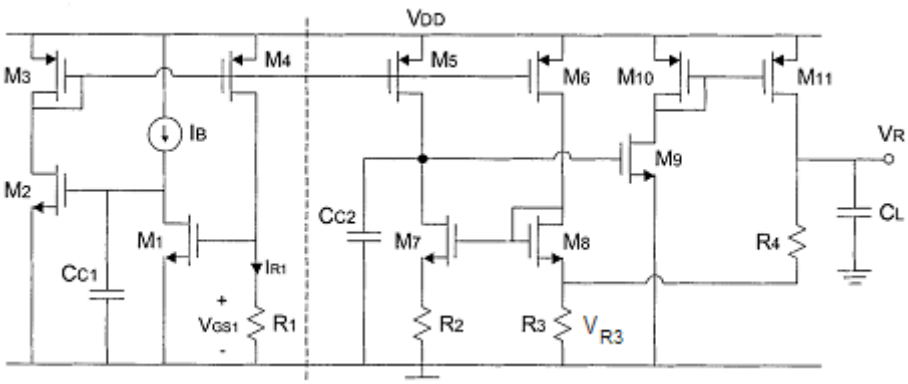


Fig. 3.9 – A voltage reference circuit based on subthreshold MOSFETs.

$$V_R = \alpha V_{GS1} + \beta V_T \,, \tag{3.33}$$

in which α and β only depend on resistor ratios and transistor aspect ratios. Therefore, the condition nullifying the temperature coefficient of V_R (i.e. $\partial V_R/\partial T$=0) is,

$$\frac{\alpha}{\beta} = -\frac{V_T\left(T_0\right)}{K_G} \,. \tag{3.34}$$

The circuit in Fig. 3.9 has been implemented in a 1.2-μm CMOS technology. It generates a reference voltage of 295.3 ±10.8 mV while dissipating 4.3 μW. Furthermore, the temperature coefficient achieved, in the range -25÷125 °C, is 119±35.7 ppm/°C.

A low voltage reference circuit using subthreshold MOSFETs (Fig. 3.10) has been employed in the 100 kHz fifth-order Chebychev low-pass filter of [53], working with supply voltages down to 0.6 V. The architecture is similar to [24] with the main difference that parasitic vertical BJTs have been replaced with subthreshold PMOSs. It is worth noting that the drop voltages is less than 0.4 V in the PMOSs $M_{1\text{-}3}$ of Fig. 3.10, whereas the drop voltage of a forward biased base-emitter junction of a BJT is about 0.7 V, thus not allowing very low voltage operation. The op-amp equalizes its input terminal potentials and the current through R_0 is given by,

$$I_{R0} = \frac{V_{GS1} - V_{GS2}}{R_0} \,. \tag{3.35}$$

Since from (2.3) V_{GS} can be written as,

$$V_{GS} = V_{th} + mV_T \ln\left(\frac{I_D}{I_0}\frac{1}{W/L}\right), \tag{3.36}$$

if the aspect ratio of M_2 is α times that of M_1, I_{R0} becomes,

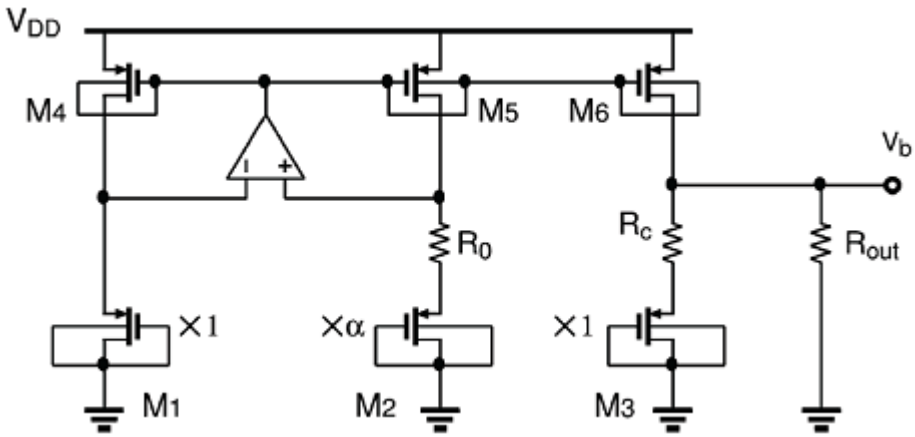


Fig. 3.10 – 0.6-V CMOS voltage reference exploiting the BGR principle on PMOSs in weak inversion.

$$I_{R0} = \frac{mV_T \ln(\alpha)}{R_0}, \tag{3.37}$$

Transistor M_6 injects this PTAT current in the reference voltage branch and the reference voltage is expressed as the sum of a CTAT quantity and a weighted PTAT quantity,

$$V_b = \frac{R_{out}}{R_{out} + R_C}\left[V_{GS3} + \frac{R_C}{R_0} mV_T \ln(\alpha)\right]. \tag{3.38}$$

The temperature coefficient achieved is very low, 25 ppm/°C in the temperature range -40÷80°C and the power consumption is 9 μW.

3.4.2. Not BGR-based architectures

Although BGR based voltage references are the most widely used in electronic ICs, the growing interest towards ultra low power circuits to be integrated in energy constrained systems is focusing the attention on innovative circuit configurations that could achieve sub-microwatt power consumption. Subthreshold operation along with new circuit topologies allowed to develop solutions working with just few nW.

A sub-microwatt CMOS voltage reference with very small temperature sensitivity and line sensitivity was developed in [55]. The schematics of the circuit is shown in Fig. 3.11, it consists of a current source subcircuit and a bias voltage subcircuit. The start-up circuit is also shown, such a circuit is usually needed in all voltage references to set the correct operating point of the circuit, as will be explained in detail in the following paragraphs. All MOSFETs are operated in subthreshold region, except for the MOS resistor M_{R1} which is operated in strong inversion, deep triode region. In the current source subcircuit V_{GS1} can be written as,

$$V_{GS1} = V_{GS2} + V_{DSR1}. \tag{3.39}$$

From (3.36) V_{DSR1} can be expressed as,

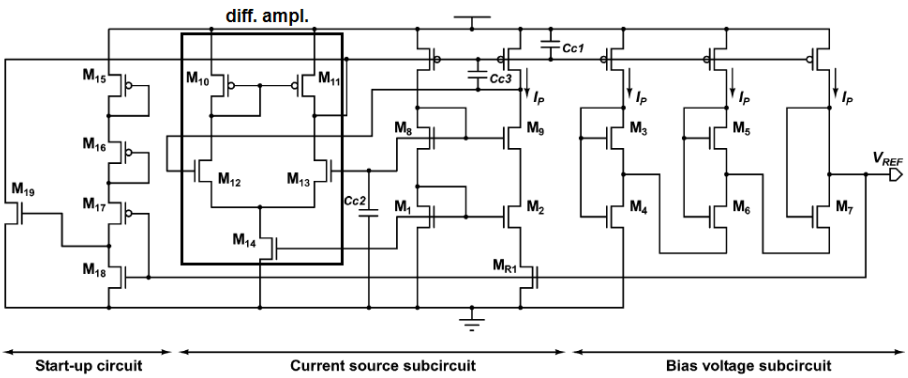


Fig. 3.11 – CMOS voltage reference circuit consisting of subthreshold MOSFETs

$$V_{DSR1} = mV_T \ln\left[\frac{(W/L)_2}{(W/L)_1}\right]. \tag{3.40}$$

The resistance of M_{R1}, operated in strong inversion, deep triode region is,

$$R_{M_{R1}} = \frac{1}{(W/L)_{R1}\,\mu C_{OX}\left(V_{REF} - V_{th}\right)}, \tag{3.41}$$

Therefore, the generated current I_P is equal to,

$$I_P = \frac{V_{DSR1}}{R_{M_{R1}}} = (W/L)_{R1}\,\mu C_{OX}\left(V_{REF} - V_{th}\right) mV_T \ln\left[\frac{(W/L)_2}{(W/L)_1}\right], \tag{3.42}$$

Further analysis can show that the generated current has a positive TC and, hence, increases with temperature.
In the bias voltage subcircuit, $V_{GS3\text{-}7}$ form a closed loop and the currents in M_4 and M_6 are $3I_P$ and $2I_P$, respectively. As a result, V_{REF} is given by,

$$\begin{aligned} V_{REF} &= V_{GS4} - V_{GS3} + V_{GS6} - V_{GS5} + V_{GS7} = \\ &= V_{GS4} + mV_T \ln\left[\frac{2(W/L)_3 (W/L)_5}{(W/L)_6 (W/L)_7}\right] = \\ &= V_{th} + mV_T \ln\left[\frac{3I_P/I_0}{(W/L)_4}\right] + mV_T \ln\left[\frac{2(W/L)_3 (W/L)_5}{(W/L)_6 (W/L)_7}\right] \end{aligned}, \tag{3.43}$$

Equation (3.43) shows that V_{REF} can be expressed as a sum of a V_{th} and V_T scaled by transistor sizes. Since V_{th} has a negative TC and V_T has a positive TC, V_{REF} with a zero TC can be obtained optimizing the transistor aspect ratios for $\partial V_{REF}/\partial T$=0.
The V_{REF} value for zero TC is equal to the threshold voltage (3.27) with T=0 K, so its value depends on process variations. The TC of V_{th} can be evaluated from (2.7) as,

$$\frac{\partial V_{th}}{\partial T} = -(2m-1)\frac{k_B}{q}\left[\ln\left(\frac{\sqrt{N_C N_V}}{N_{ch}}\right) + \frac{3}{2}\right] + \frac{m-1}{q}\frac{\partial E_g}{\partial T}, \tag{3.44}$$

where N_C and N_V are the effective densities of states in the conduction and valence bands, respectively. V_{th} and $\partial V_{th}/\partial T$ both depend on N_{ch}, that is a process dependent quantity. However the change of $\partial V_{th}/\partial T$ due to process variations is very small since it logarithmically depends on N_{ch}. In [55] the effects of process variations on V_{th} and its TC have been studied for a 0.35 µm CMOS process. Fig. 3.12 shows the simulated values of V_{th} and $\partial V_{th}/\partial T$ as a function of N_{ch}. The dashed lines represent the range of N_{ch} for the used technology. In this concentration range, V_{th} changes by ±20% with N_{ch}, while $\partial V_{th}/\partial T$ only changes by ±2%. The measured TC of V_{REF} varies in the range 7÷45 ppm/°C for a set of 17 samples.

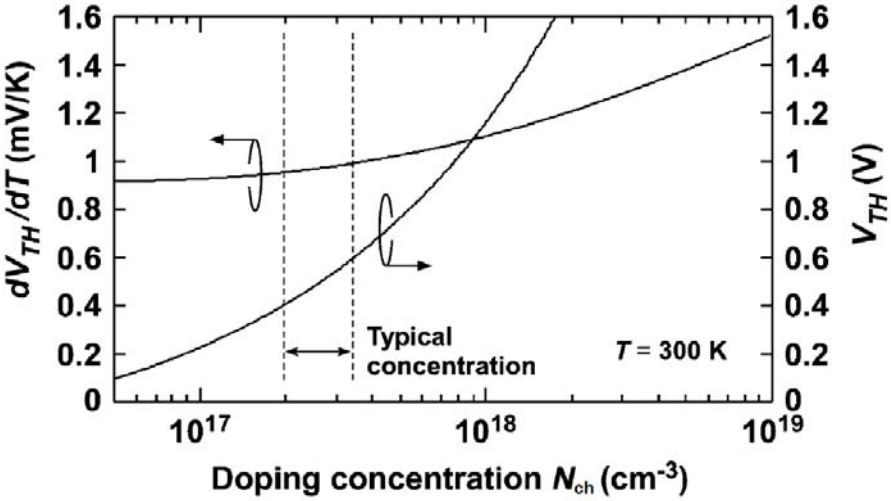


Fig. 3.12 – Threshold voltage and its TC as a function of N_{ch} for a 0.35m CMOS process, at room temperature.

Moreover, the circuit in Fig. 3.11 shows a remarkable insensitivity to V_{DD} variations thanks mainly to the differential amplifier and the cascoded current mirror in the current source subcircuit. In fact the line sensitivity is just 20 ppm/V. Although cascode configurations are useful to reduce the line sensitivity, they consume more supply voltage headroom than those based on simple current mirrors, indeed the minimum is V_{DD} 1.4 V. Current dissipation is only 214 nA at room temperature, this result is principally due to the fact that all MOSFETs are biased with low subthreshold currents.

3.4.2.1. Sub-100 nW voltage references

In low power, low voltage design it is often preferred to exploit the MOS characteristics in both strong and weak inversion regions. Concerning voltage reference design, that is the case of [18], where a power consumption lower than 100 nW and a V_{DD}<1 V have been achieved. Indeed, a novel temperature compensation technique has been developed exploiting both operating regions. The circuit configuration is shown in Fig. 3.13, the current generator provides a current I_0, that is mirrored into the active load M_{10} which generates a temperature compensated V_{GS}=V_{REF}. Assuming that M_{10} works in the saturation region of strong inversion, its V_{GS} is given by,

$$V_{GS10} = V_{REF} = V_{th10} + \sqrt{\frac{2I_0}{(W/L)_{10}\,\mu C_{OX}}}\ . \tag{3.45}$$

Temperature influences on the reference voltage are mainly due to the temperature dependence of V_{th}, μ and I_0. Assuming that V_{th} varies linearly with temperature (3.27), therefore, in order to achieve the temperature compensation, with a perfect cancellation of temperature dependence of mobility, the current generator should provide a current $I_0 \propto \mu T^2$.

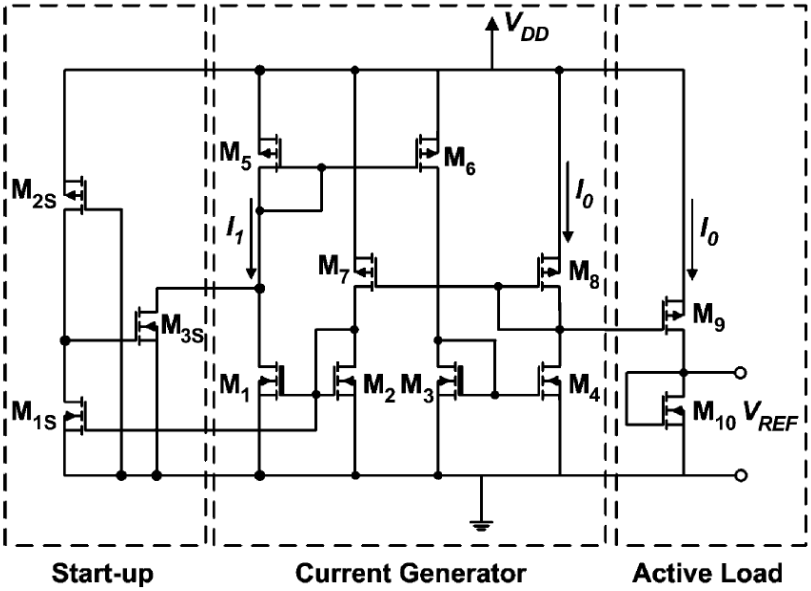


Fig. 3.13 – Sub-100 nW, CMOS voltage reference working with MOSFETs operated in weak and strong inversion regions.

In the current reference circuit $M_{1,2}$ are high V_{th} transistors, while the remaining ones are standard V_{th} transistors. This allow to bias $M_{1,3}$ in subthreshold region and other transistors in strong inversion region. $M_{5,6}$ impose equal current I_1 in $M_{1,3}$, while $M_{7,8}$ impose equal current I_0 in $M_{2,4}$. I_0 can be found considering that $V_{GS1}=V_{GS2}$ and $V_{GS3}=V_{GS4}$, therefore, expressing the characteristic current in (2.3) as $\mu C_{OX}V_T^2$, it follows that,

$$\begin{cases} V_{th1} + mV_T \ln\left[\dfrac{I_1}{\mu C_{OX}V_T^2 (W/L)_1}\right] = V_{th2} + \sqrt{\dfrac{2I_0}{(W/L)_2\,\mu C_{OX}}} \\ V_{th3} + mV_T \ln\left[\dfrac{I_1}{\mu C_{OX}V_T^2 (W/L)_3}\right] = V_{th4} + \sqrt{\dfrac{2I_0}{(W/L)_4\,\mu C_{OX}}} \end{cases}, \tag{3.46}$$

where it has been assumed that subthreshold MOSFETs are saturated (i.e. V_{DS}>3÷4V_T). Solving (3.46) for I_0 gives the result,

$$I_0 = \frac{m^2 C_{OX}(W/L)_4}{2(N-1)^2}\mu V_T^2 \ln^2\left[\frac{(W/L)_3}{(W/L)_1}\right] \quad \text{with} \quad N \equiv \sqrt{\frac{(W/L)_4}{(W/L)_2}}\,. \tag{3.47}$$

As expected $I_0 \propto \mu T^2$, therefore, substituting it in (3.45) V_{REF} becomes,

$$V_{REF} = V_{th10} + V_T \frac{m}{N-1}\sqrt{\left[\frac{(W/L)_4}{(W/L)_{10}}\right]}\ln\left[\frac{(W/L)_3}{(W/L)_1}\right]. \tag{3.48}$$

By proper dimensioning of the transistors it is possible to obtain zero TC for V_{REF}. The developed temperature compensation technique allowed achieving a TC of 10 ppm/°C.
The minimum supply voltage is imposed by the current generator circuit. In particular, to ensure that M_5 operates in saturation region with $V_{GS}<V_{th5}$ (V_{thp}=-0.75 V for the 0.35 μ CMOS process considered) and M_1 has a $V_{DS}>4V_T$, the V_{DDmin} should be, theoretically,

$$V_{DD\min} = \left|V_{GS5}\right| + V_{DS1\min} = \left|V_{thp}\right| + 4V_T \approx 0.85 \text{ V}. \tag{3.49}$$

From experimental results, the V_{DDmin} measured is 0.9 V and the power consumption at room temperature is just 36 nW.
Extreme low voltage operation along with extreme low power dissipation have been attained in the CMOS solution in Fig. 3.14, presented in [57]. All transistors in the circuit configuration operate in subthreshold region.
The whole circuit is biased with an ultra low power self-biased current reference (similar to that shown in [60]) that generates voltage V_b. Drain currents of M_2 and M_4 are given by,

$$\begin{aligned} I_{D2} &= \left(\frac{W}{L}\right)_2 I_0 e^{\left(\frac{V_{GS1}-V_{GS4}-V_{th}}{mV_T}\right)} \\ I_{D4} &= \left(\frac{W}{L}\right)_4 I_0 e^{\left(\frac{V_{GS4}-V_{th}}{mV_T}\right)} \end{aligned} . \tag{3.50}$$

Considering that $I_{D2}\equiv I_{D4}$, V_{GS4} can be extracted from (3.50),

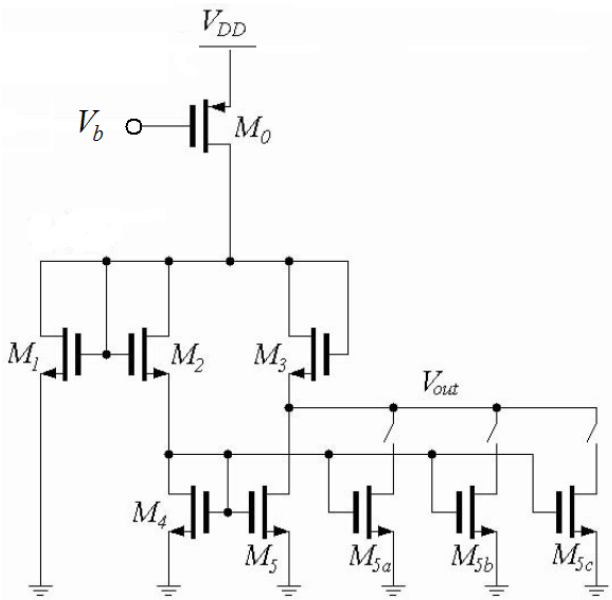


Fig. 3.14 – 0.6 V voltage reference configuration working with all transistors in weak inversion.

$$V_{GS4} = \frac{V_{GS1}}{2} + mV_T \ln\left(\sqrt{\frac{(W/L)_2}{(W/L)_4}}\right). \tag{3.51}$$

From Fig. 3.14 it can be recognized that V_{OUT}=V_{GS4}+(V_{GS2}-V_{GS3}), remembering (3.36), V_{OUT} becomes,

$$V_{OUT} = \frac{V_{th}}{2} + mV_T \ln\left(\frac{(W/L)_3}{(W/L)_5}\sqrt{\frac{(W/L)_2}{(W/L)_4}}\right). \tag{3.52}$$

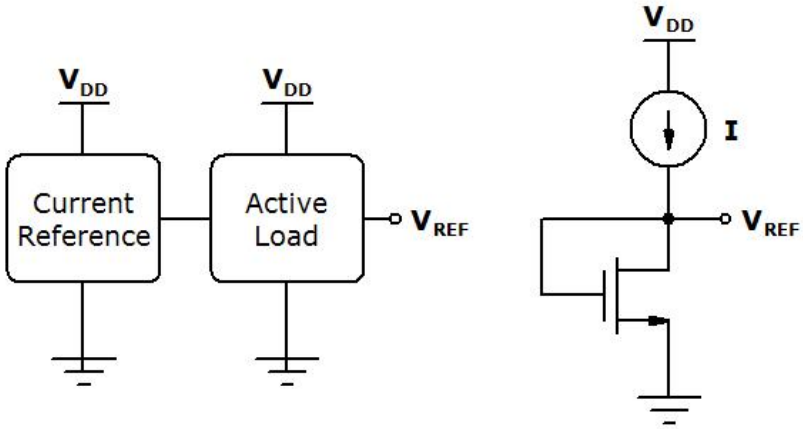
Again a CTAT voltage (V_{th}) is added to the PTAT thermal voltage weighted with transistor aspect ratios. Therefore by choosing the appropriate aspect ratios of $M_{2\text{-}5}$, a reference voltage with $\partial V_{OUT}/\partial T$=0 is obtained. An array of transistors with binary-weighted channel width and equal channel length is placed beside M_5 ($M_{5a,5b,5c}$) to digitally control, by means of switches, the effective aspect ratio of M5 for a post-fabrication tuning of the TC.
The circuit has been fabricated in a 0.18μm CMOS technology and experimental results showed a TC of 127 ppm/°C in the range -20÷120 °C. Subthreshold operation of all MOSFETs allowed correct operation with extremely low V_{DD}s (i.e. down to 0.6 V) and a power consumption of less than 40 nW at room temperature and V_{DD}=0.7 V.

3.5. Proposed voltage references

In this section new CMOS voltage reference configurations, based on a temperature compensation technique derived from the subthreshold regime of operation, are described. Moreover, subthreshold operation is exploited to minimize both the power consumption and the supply voltage. Indeed the power consumption achieved is in the order of few nW, whereas the proposed voltage references are able to work with supply voltages down to 0.45 V.
Within the circuits, n-channel MOSFETs with two different threshold voltages are present and the reference voltage values can be approximated, in last analysis, to the difference of those two threshold voltages at room temperature. Measurements on silicon confirm that the proposed configurations allow minimum power consumption at the least one order of magnitude lower than the best solutions found in the literature. Among the compared circuits, one of the three proposed configurations is characterized by the lowest supply voltage.
This paragraph is organized as follows: in sections 3.5.1. and 3.5.2 the principle of operation of the proposed voltage references and the realized circuit configurations are described in detail; dynamic range and power consumption considerations are discussed in section 3.5.3.; in section 3.5.4. measurement results are shown and compared with low-power low-voltage competitors. A statistical analysis of the figures of merit of each circuit configuration is also carried out in this section.

3.5.1. Operating principle

Basically, the scheme of a voltage reference can be represented as in Fig. 3.15 (a). The current reference provides a current, as independent as possible of supply voltage variations, to the active load that generates the reference voltage. The



(a) (b)

Fig. 3.15 – .(a) Basic diagram of a voltage reference, (b) Scheme of principle of a simple voltage reference.

current injected in the active load compensates temperature effects on the voltage generated. Process variations effects compensation can act in both sections, current generator or active load.

A voltage reference can be simply represented by a current source and a diode connected NMOS, as shown in Fig. 3.15 (b). Such a structure has been adopted in [18], where the load NMOS works in the saturation region providing a temperature compensated reference voltage (V_{REF}) higher than its threshold voltage (V_{th}) at room temperature.

In this work we exploited the subthreshold operation of the load transistor, in order to obtain $V_{REF}<V_{th}$ and, consequently, to further reduce the supply voltage of the overall voltage reference circuit. It is worth noting that biasing the load NMOS in subthreshold region in [18] would not allow temperature compensation, consequently even the current source solution adopted in this work is different from the one used in [18].

Investigating the subthreshold behaviour of the load, a formula for the injected current (i.e. I in Fig. 3.15 (b)), compensating the reference voltage dependence on temperature, can be obtained. To this aim, the I-V characteristics of an NMOS operating in subthreshold region considered is similar to (2.3) and is given by,

$$I_D = S\mu V_T^2 \exp\left(\frac{V_{GS}-V_{th}}{mV_T}\right)\left[1-\exp\left(-\frac{V_{DS}}{V_T}\right)\right], \tag{3.53}$$

where I_0W/L in (2.3) has been replaced by $S\mu V_T^2$, with $S=(W/L)C_{OX}$. From (3.53), for $V_{DS}\geq 3\div 4V_T$, I_D becomes almost independent of V_{DS}, i.e. the transistor works in the so called saturation region of weak inversion. Thus V_{REF} of NMOS in Fig. 3.15 (b) can be approximated by (3.54)

$$V_{REF} = V_{GS} = V_{th} + mV_T \ln\left(\frac{I_D}{S\mu V_T^2}\right). \tag{3.54}$$

Let us note that, at a first approximation, $V_T(T)$, $V_{th}(T)$ and $\mu(T)$ introduce dependence on temperature. The BSIM3v3.2 model provided by the foundry design kit of the CMOS technology used for simulation and fabrication of the circuits, assumes that V_{th} decreases linearly with temperature [46],

$$V_{th} = V_{th}(T_0) + \left(k_{t1} + k_{t2}V_{BS}\right)\left(T/T_0 - 1\right), \tag{3.55}$$

where $V_{th}(T_0)$ is the threshold voltage at the reference temperature ($T_0 \cong 300.15$°K), V_{BS} is the body-to-source voltage of the transistor. Temperature coefficients k_{t1} and k_{t2} have negative values. Moreover, at a first approximation the subthreshold slope parameter m can be considered independent of temperature. However, in the next paragraph, experimental results of $V_{th}(T)$ and $m(T)$ thermal behaviors will be shown.

In order to have a temperature compensated output reference voltage, the following condition must be satisfied,

$$\frac{\partial V_{REF}}{\partial T} = 0. \tag{3.56}$$

Based on equation (3.54), a simple solution of (3.56) can be obtained by generating the current I_D with the following temperature dependence,

$$I_D\left(T\right) = \alpha\mu T^2, \tag{3.57}$$

Therefore, combining (3.57) and (3.54), 3.56 becomes,

$$\frac{\partial V_{REF}}{\partial T} = \frac{\partial V_{th}}{\partial T} + m\frac{k_B}{q}\ln\left(\frac{\alpha}{C_{OX}\left(W/L\right)\left(k_B/q\right)^2}\right) = 0, \tag{3.58}$$

and a value for α (constant with temperature) can be found as,

$$\alpha = C_{OX}\left(\frac{W}{L}\right)\left(\frac{k_B}{q}\right)^2 \exp\left(\frac{q\left|k_{t1}\right|}{mk_BT_0}\right). \tag{3.59}$$

Substituting it in (3.54) and remembering that V_{BS}=0 V in the scheme of Fig. 3.15 (a), the reference voltage becomes,

$$V_{REF} = V_{th}\left(T_0\right) + \left|k_{t1}\right|. \tag{3.60}$$

However, the above solution does not ensure subthreshold operation for the active load.

For this reason, more detailed analysis has to be done, considering for the current source the more general form,

$$I\left(T\right) = \alpha\mu T^2 f\left(T\right), \tag{3.61}$$

where $f(T)$ is a generic function of temperature. Substituting it in (3.54), the condition (3.56) becomes a differential equation in $f(T)$,

$$\frac{\partial V_{REF}}{\partial T} = \frac{k_{t1}}{T_0} + m\frac{k_B}{q}\ln\left(\frac{\alpha}{S\left(k_B/q\right)^2}\right) + \\ +m\frac{k_B}{q}\left\{\ln\left(f\left(T\right)\right) + \frac{T}{f\left(T\right)}\frac{\partial}{\partial T}\left(f\left(T\right)\right)\right\} = 0 \quad , \tag{3.62}$$

In order to achieve temperature compensation the term in curly brackets in (7) should be constant with temperature. A possible solution for $f(T)$ could be simply a constant, but this would lead to the case found in (3.57). Another solution, which can ensure subthreshold operation, can be written as,

$$I_D\left(T\right) = \alpha\mu T^2 \exp\left(\frac{AT+B}{CT}\right), \tag{3.63}$$

where A, B and C are constant with temperature. This solution has been chosen since it was easy to generate with a subthreshold MOSFET based current reference, as will be clear in the next section. It is worth noting that the current reference used in [18] does not allow implementing equation (3.63), therefore, in that case, it is not possible to compensate temperature effects by biasing the load in subthreshold regime. Substituting (3.63) in (3.54), from (3.56) an expression for α can be found which leads to the temperature-compensated reference voltage in (3.65).

$$\frac{\partial V_{REF}}{\partial T} = 0 \Rightarrow \alpha = C_{OX}\left(\frac{W}{L}\right)\left(\frac{k_B}{q}\right)^2 \exp\left(\frac{q\left|k_{t1}\right|}{mk_BT_0} - \frac{A}{C}\right). \tag{3.64}$$

$$V_{REF\_OPT} = V_{th}\left(T_0\right) + \left|k_{t1}\right| + m\frac{k_B}{q}\frac{B}{C}. \tag{3.65}$$

In this case, if the term B/C is negative, differently from (3.60), the load transistor can work in subthreshold region. In the following, how to implement circuit configurations like the one in Fig. 3.15 (a), based on the proposed operating principle, will be shown.

3.5.1.1. Experimental thermal behavior of V_{th} and m

Several works related to low-power voltage references, in which subthreshold operated MOSFETs are involved, consider, at a first approximation, m independent of temperature (e.g. see [18], [55], [56]).

The m dependence on temperature for three NMOSs belonging to the same voltage reference dies and with a W/L of 12 µm / 50 µm, have been measured. The results have been reported in Fig. 3.16. Varying the temperature from 25 °C to 125 C the maximum deviation of m from the mean value evaluated at 25°C is 0.086, that is a variation of only 7% or 715 ppm/°C. Since in (3.54) $V_T(T)$, $V_{th}(T)$ and $\mu(T)$ introduce a much more stronger dependence on temperature, the subthreshold slope parameter can be assumed independent of temperature, in the operating principle section.

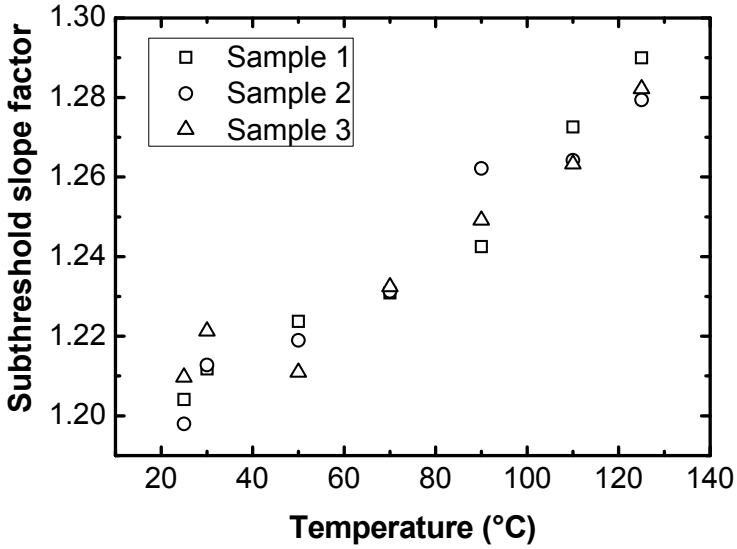


Fig. 3.16 – NMOS m-T characteristic

For what it concerns the V_{th} dependence on temperature, in order to study the accuracy of the approximation reported in (3.55), the V_{th} dependence on temperature for NMOSs with a W/L of 12 µm / 50 µm has been measured. In such analysis it has been set V_{BS}=0 V.

The threshold voltage at room temperature, $V_{th}(T_0)$, predicted by the model is

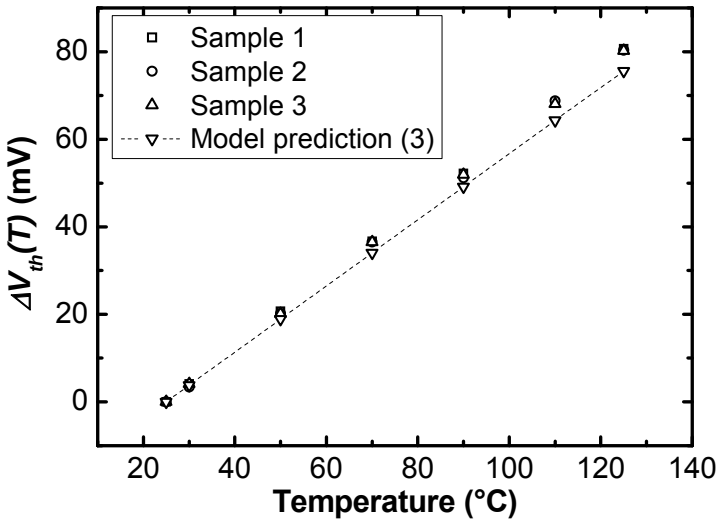


Fig. 3.17 – Measured and predicted V_{th} dependence on temperature.

320 mV, only 1.5 mV higher than the measured mean value over a set of 40 NMOSs.
In Fig. 3.17 the $\Delta V_{th}(T)=V_{th}(T)-V_{th}(T_0)$ temperature dependence of the three NMOSs previously considered and the one predicted by the model, for the CMOS process UMC 0.18 μm, is shown.

3.5.2. Electrical configurations

The developed temperature compensation technique has been employed in the design and implementation of three voltage references in the 0.18 μm UMC CMOS technology.
TABLE 3.1 summarizes information about the technology chosen to implement the proposed circuit. In particular, the table contains the allowed voltage ranges and the threshold voltage values of both Standard V_{th} (SVT) and High V_{th} (HVT) MOSFET models.

Table 3.1
Small-signal MOSFET parameters

Technology	UMC 0.18-μm CMOS process.		
	Voltage range	**V_{th}-NMOS**	**V_{th}-PMOS**
SVT-MOSFET	0 V $\leq V_{GS}$, V_{DS}, $\lvert V_{BS}\rvert \leq$ 1.8 V	0.320 V	-0.456 V
HVT-MOSFET	0 V $\leq V_{GS}$, V_{DS}, $\lvert V_{BS}\rvert \leq$ 3.3 V	0.600 V	-0.720 V

3.5.2.1. Voltage reference 1

The architecture of the proposed voltage reference 1 is illustrated in Fig. 3.18. The self-biased current source has no resistors and works with all MOSFETs in the subthreshold region. It generates the current to be injected in the load transistor (M_{10}) branch. A start-up circuit has been added as a precautionary measure to ensure bias in the desired state.
TABLE 3.2 shows the size of transistors in the current reference, in the active load and in the start-up circuit. The number of parallel MOSFETs used in the circuit layout has also been reported.

3.5.2.1.1. Current reference

The purpose of the current generator in Fig. 3.18 is to provide a current, as independent as possible of supply voltage variations, which compensates temperature effects on V_{REF}. A current reference configuration capable of generating a current almost independently of V_{DD} is the self-biased configuration.
A current in the form (8) can be obtained taking into account a linear combination of NMOS V_{GS} voltages in the subthreshold region. Among the possible solutions for the current reference we chose a self-biased configuration where only three subthreshold operated NMOSs, $M_{1\text{-}3}$, perform such V_{GS} combination, giving

$$V_{GS2} = V_{GS1} + V_{GS3}\,. \tag{3.66}$$

In the current reference, PMOSFETs M_5 and M_7 form a subthreshold current mirror providing ratioed currents $I_2/I_1=(W/L)_7/(W/L)_5=a$. All the transistors in the proposed solution are SVT-MOSFETs, except for M_2 and $M_{1S,}$ which are HVT-MOSFETs.

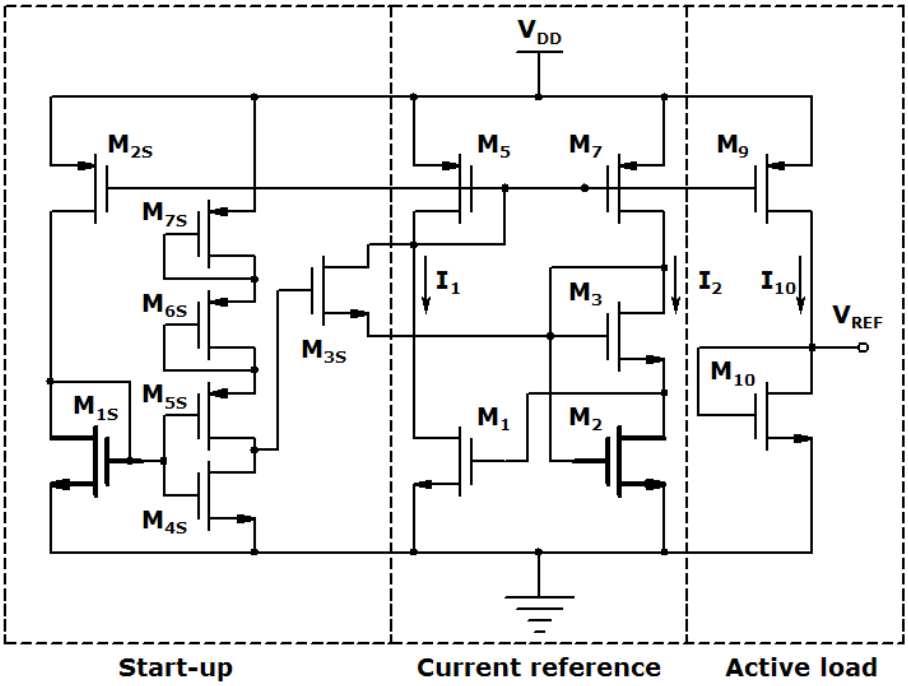


Fig. 3.18 – Voltage reference 1 circuit configuration.

Table 3.2

Transistor sizes of the current reference, the active load and the start-up circuit

Transistor	Value (*W/L*)
M_1	8 μm / 50 μm
M_2	100 μm / 24 μm = (50 μm / 24 μm) × 2
M_3	2 μm / 2 μm
M_5	68 μm / 50 μm = (17 μm / 50 μm) × 4
M_7	52 μm / 50 μm = (26 μm / 50 μm) × 2
M_9	97 μm / 29 μm = (48.5 μm / 29 μm) × 2
M_{10}	1.5 μm / 50 μm
M_{1S}	0.5 μm / 50 μm
M_{2S}	0.5 μm / 50 μm
M_{3S}	0.5 μm / 1 μm

M_{4S}	0.5 μm / 100 μm
M_{5S}	1 μm / 50 μm
M_{6S}	1 μm / 50 μm
M_{7S}	1 μm / 50 μm

In order to ensure subthreshold operation for the whole current source, for M_2 a high V_{th} device was chosen. Using (3.54), (3.66) becomes

$$m_2 V_T \ln\left(\frac{I_2}{S_2\mu_2 V_T^2}\right) = \Delta V_{th} + m_3 V_T \ln\left(\frac{I_2}{S_3\mu_3 V_T^2}\right) + m_1 V_T \ln\left(\frac{I_1}{S_1\mu_1 V_T^2}\right), \quad (3.67)$$

where $\Delta V_{th}=V_{th1}+V_{th3}-V_{th2}$. Assuming that electron mobilities in (3.67) are identical ($\mu \approx \mu_1 \approx \mu_2 \approx \mu_3$) and using $I_2=aI_1$, the current flowing in the left branch of the current generator can be evaluated from (3.67) as

$$I_1 = Q^{1/\Sigma_m} \mu V_T^2 \exp\left(-\frac{\Delta V_{th}}{V_T \Sigma_m}\right), \quad (3.68)$$

wherein,

$$\begin{aligned} Q &= a^{m_2-m_3}\,(S_3^{m_3} S_1^{m_1} / S_2^{m_2}) \\ \Sigma_m &= m_1 + m_3 - m_2 \end{aligned} . \quad (3.69)$$

Considering the V_{th} thermal behaviour shown in (3.55), it can be concluded that the subthreshold current reference architecture in Fig. 3.18 generates a current in the form (3.63).

The stability of V_{REF} with V_{DD} variations mainly depends on the current generator's insensitivity to supply voltage variations. This characteristics is achieved by means of the self-biased current source architecture, where the high impedance elements per branch (M_7 and M_1) are capable of absorbing supply voltage variations leaving their current almost unchanged [44].

Significant improvement of the reference voltage line sensitivity and the power supply rejection ratio (PSRR) can be obtained by using an additional current mirror, in the self biased current reference circuit [55]. The use of additional current mirrors in the current generator, which add transistors in stack increasing the minimum achievable V_{DD}, were avoided.

3.5.2.1.2. Start-up circuit

Self-biased current generators and, generally, all the circuits with two possible DC operating points need a start-up circuit [44,61]. In the case of self-biased references, a degenerate bias point can occur if, for example, transistors carry zero current. They may remain off indefinitely because the loop can support a zero current in both branches.

Basically, a start-up circuit checks a possible zero current condition and by injecting current in a suitable point forces the circuit to move from zero state and brings it to the correct point of operation.

In Fig. 3.18, PMOS M_{S2} checks the current generator. In case it is zero, the current injected in M_{1S} is zero too and V_{GSS1}=0 V. The gate of M_{S1} is connected to the input of an inverter which, for a low input gives a high output. The output of the inverter coincides with the gate of M_{3S}, whose source and drain terminal goes to ground and V_{DD} voltage in case of zero current in the current reference branches. In that case, therefore, M_{3S} starts injecting a current in both current generator branches, thus allowing it to leave the zero current state. In normal conditions, M_{2S} injects a current in M_{1S} that provides a high input to the inverter, so that transistor M_{3S} remains OFF.
The main drawback of start-up circuits is that they dissipate power even when they are not working, i.e. the self-biased works properly. Several solutions have been developed to reduce power consumption in start-up circuits [62, 63]. In the proposed voltage reference the start-up circuit current consumption has been drastically reduced by mirroring a current in M_{2S} more than one hundred times lower (see current mirror sizes in TABLE 3.2) than that injected in the three branches of the current reference and the active load. Moreover, two diode-connected PMOSs ($M_{6S,7S}$) have been added in the inverter branch to reduce the current consumption due to start-up. Therefore, thanks to a proper dimensioning of the start-up circuit used in the proposed voltage references, the power dissipation of the start-up section can be neglected when compared to that of the whole circuit.
Furthermore, the proposed circuit configurations have been fabricated with and without the start-up circuit. From measurements, the same behavior of both topologies has been observed. In particular none of the circuits without start-up has entered in the zero current state, as usually expected for a voltage references based on self-biased current generator. Since the proposed voltage references operate with very low current, in the order of few nA, it is possible that the zero current state in the self-biased architecture becomes not as stable as it is when biased with currents in the order of mA or μA. This could lead to reconsider the voltage reference design techniques, in particular the need of a start-up section, when operating with extremely low current levels.
Finally, bearing in mind that the power consumption of the start-up circuits is negligible when compared with the whole power dissipated in the proposed voltage references, the only drawback owed to the start-up section in such case is related to the additional silicon area occupation.

3.5.2.1.3. Active load

The generated current is then mirrored into a diode connected transistor through M_9. The current mirror gain can be expressed as $I_{10}/I_1=(W/L)_9/(W/L)_5=c$. Therefore, with M_{10} working in the subthreshold region, the reference voltage becomes, from (3.54) and (3.68),

$$V_{REF} = V_{th10} + m_{10} V_T \ln\left\{ Q^{1/\Sigma_m} \frac{c}{S_{10}} \right\} - \frac{m_{10}}{\Sigma_m} \Delta V_{th} . \qquad (3.70)$$

The temperature dependence of the reference voltage is firmly related to the thermal behaviour of V_T and V_{th}. As already obtained in [55, 57, 18], the reference voltage is a linear combination of quantities depending on V_{th}, hence with a CTAT behavior, and on $V_T=k_BT/q$. It is possible, therefore, to dimension the MOSFETs in the circuit configuration in order to achieve the zero TC of V_{REF}.

According to the linear approximation made in (3.55), setting $\partial V_{REF}/\partial T$=0, a value of $(W/L)_{10}$ for temperature compensation can be extracted as,

$$\left(\frac{W}{L}\right)_{10OPT} = \frac{Q^{1/\Sigma_m}\, c/C_{OX10}}{\exp\left\{\frac{q}{k_B T_0}\left[\frac{1}{\Sigma_m}\left(K + k_{t2,3}V_{BS3}\right) - \frac{k_{t1,10}}{m_{10}}\right]\right\}}, \tag{3.71}$$

where $K=k_{t1,1}+k_{t1,3}-k_{t1,2}$ ($k_{t1,i}$ and $k_{t2,i}$ are the threshold voltage temperature coefficients of the i-th transistor).
Hence, satisfying (3.70), V_{REF} becomes,

$$\begin{aligned} V_{REF\_OPT} &= V_{th10}\left(T_0\right) + \left|k_{t1,10}\right| + \frac{m_{10}}{\Sigma_m}\left(V_{th2}\left(T_0\right) + \left|k_{t1,2}\right|\right) - \\ &- \left[\frac{m_{10}}{\Sigma_m}\left(V_{th1}\left(T_0\right) + V_{th3}\left(T_0\right) + \left|k_{t1,1}\right| + \left|k_{t1,3}\right| + \left|k_{t2,3}\right|V_{BS3}\right)\right], \end{aligned} \tag{3.72}$$

the term indicated in square brackets ensures subthreshold operation of M_{10}.
As is clear from (3.72), the reference voltage value only depends on process parameters. A simpler expression for V_{REF_OPT} can be obtained considering that, as usual, the m parameters of different NMOSFETs biased in the subthreshold region have similar values (then $m_{10}/\Sigma_m \approx 1$) and $k_{t1,1} \approx k_{t1,2} \approx k_{t1,3} \approx k_{t1,10}$. Therefore, the reference voltage value can be approximated to the threshold voltage difference between an HVT-NMOS (V_{th2}) and an SVT-NMOS (V_{th1} or V_{th3}). For the chosen technology this difference is enough for the subthreshold operation of M_{10} and, consequently, it is possible to exploit the temperature compensation theory developed in the first section. The mean reference voltage measured is 257.5 mV,

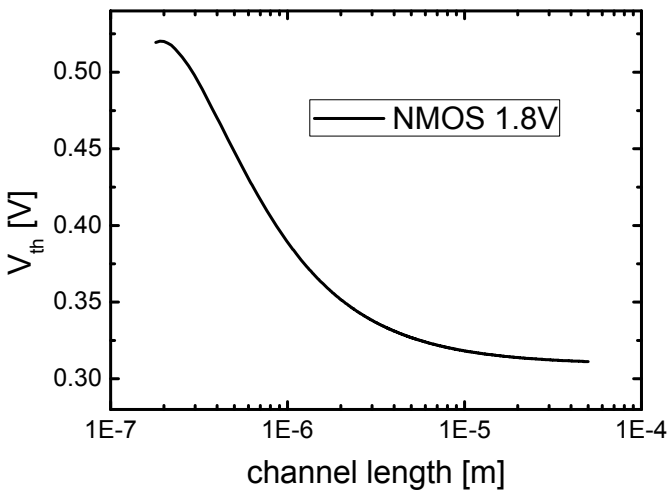


Fig. 3.19 – Simulated V_{th} dependence channel length for a 1.8 V NMOS.

about 23 mV lower than the roughly predicted value. This difference is mainly due to the increase of the threshold voltage of M_3 caused by both body effect and the effect of lateral pocket implants (M_3 has a medium channel length of 2 μm). The simulated V_{th} dependence on channel length for a SVT-NMOS with W=10 μm is shown in Fig. 3.19.
The active load aspect ratio, for temperature compensation, predicted by (14), for the chosen technology, neglecting the body effect on M_3, is 0.0273, while the one chosen after circuit simulation for temperature coefficient optimization, is $(W/L)_{10}$=1.5 μm/50 μm=0.030. The good agreement confirms the validity of the proposed theoretical approach for subthreshold temperature compensation.

3.5.2.1.4. Minimum current consumption

The minimum current consumption is not limited by the operating region of transistors inside the proposed configuration. Thanks to subthreshold operation, all $|V_{GS}|$ values are smaller than the absolute value of MOSFET threshold voltages. Recalling (3.54), it can be stated that, in principle, there is no lower limit for the current supplied in the subthreshold voltage reference of Fig. 3.18. Nevertheless, the proposed temperature compensation technique imposes a lower bound for the generated current, as explained in the following.
The total current drawn from the power supply, neglecting the current flowing in the start-up circuit, can be expressed as the sum of the currents injected into the current source and the active load branches,

$$I_{DD} = I_1\left(1 + a + c\right) = I_1 B \,. \tag{3.73}$$

From (3.68) and (3.71), I_{DD} can be rewritten as,

$$I_{DD} = \frac{A}{E}\frac{B}{c}\left(\frac{W}{L}\right)_{10OPT} , \tag{3.74}$$

wherein A and E are the process and temperature dependent parameters found in (3.68) and (3.71), respectively,

$$\begin{aligned} A &= \mu V_T^2 \exp\left(-\frac{\Delta V_{th}}{V_T \Sigma_m}\right) \\ \frac{1}{E} &= C_{OX10} \exp\left\{\frac{q}{k_B T_0}\left[\frac{1}{\Sigma_m}\left(K + k_{t2,3} V_{BS3}\right) - \frac{k_{t1,10}}{m_{10}}\right]\right\} \end{aligned} , \tag{3.75}$$

The remaining terms in (3.74) only rely on MOSFET aspect ratios. In order to reduce the power consumption, while compensating temperature effects on the reference voltage, the load aspect ratio resulting from (3.71) should not exceed the minimum derivable for the technology chosen: W_{MIN}/L_{MAX}. Moreover, to further reduce I_{DD}, the term B has to be minimized (from (3.73), B_{MIN}~1), whereas the current mirror gain c has to be maximized (i.e. $c_{MAX}=(W_{MAX}/L_{MIN})/(W_{MIN}/L_{MAX})$).

3.5.2.1.5. Minimum supply voltage

Considering the voltage reference configuration in Fig. 3.18, the supply voltage will redistribute as a sum of a V_{DS} and a V_{GS} in the branches of both current reference and active load. Although subthreshold operation does not limit the minimum V_{GS} achievable, in order to generate a current and hence a reference voltage, as

independent as possible of V_{DD} variations, the high impedance transistors (i.e. M_1, M_7 and M_9) should absorb at least a V_{DS}~$4V_T$. The maximum working temperature of the proposed circuit is 125 °C, therefore, the minimum V_{DS} value able to ensure proper operation, even at that temperature, becomes $4V_T$=137 mV. Consequently, by generating a current (3.68) so that, in each branch, V_{GS}<<$4V_T$, a minimum supply voltage of $4V_T$ can be achieved with a similar voltage reference architecture regardless of the process technology chosen.
Nevertheless, in our specific case, the gate-to-source voltage of M_{10}, obtained for temperature compensation, is process dependent and its value cannot be minimized. So, the minimum V_{DD}, ensuring a $|V_{DS}|>4V_T$ for M_9, will be

$$V_{DDMIN} \geq V_{REF\_OPT} + 4V_T \,. \tag{3.76}$$

Since the mean reference voltage measured is 257.5 mV, it follows that V_{DDMIN} , for the subthreshold circuit configuration in Fig. 3.18, will be equal or higher than 395 mV, considering the thermal voltage value at the maximum operating temperature. Obviously, the higher V_{DDMIN}, the more accurate the approximation done in (3.54) for M_9 and, as a consequence, the smaller the circuit line sensitivity. From measurements, proper operation was observed with V_{DD}≥450 mV, which is the smallest obtained so far for voltage reference circuits.

3.5.2.1.6. Sensitivity to supply voltage variations

A common parameter used for the evaluation of the reference voltage sensitivity to V_{DD} variations is the line sensitivity. It is defined as $LS=(\Delta V_{REF}/\Delta V_{DD})/\ V_{REF}$ (%), where ΔV_{DD} is the V_{DD} range of operation (1.8 V-0.45 V=1.35 V for the proposed circuit), ΔV_{REF} is the absolute variation of the reference voltage, in the V_{DD} range considered, and V_{REF} stands for the mean output value. The line sensitivity optimization starts from the minimization of ΔV_{REF}. From Fig. 3.18and 3.18, it is clear that it can be easily translated in the minimization of I_{10} variations.
Neglecting the effect of the term in square brackets in (2.3) (i.e. considering a $V_{DS}\geq 4V_T$), we can rewrite the subthreshold drain current, highlighting the V_{DS} dependence,

$$I_D = I_S \exp\left(\frac{V_{GS}}{mV_T}\right)\exp\left(\frac{\lambda_D V_{DS}}{mV_T}\right), \tag{3.77}$$

in which I_S summarizes the product between the aspect ratio and the process and temperature parameters found in (2.3) and λ_D is the drain-induced barrier lowering (DIBL) factor. Subthreshold operated MOSFETs do not suffer from channel length modulation, therefore, for $V_{DS}\geq 4V_T$, I_D will depend on V_{DS} only because of the DIBL effect. As is well known, the DIBL effect significantly decreases with increasing channel length.
The ΔV_{DD} will cause a $\Delta V_{DS}=V_{DSMAX}-V_{DSMIN}$ on the high impedance transistor in the branch. Assuming constant V_{GS}, the resulting I_D variation will be characterized by a maximum to minimum ratio $I_{DMAX}/I_{DMIN}=exp(\lambda_D \Delta V_{DS}/mV_T)$. Recalling (3.54), the ΔV_{REF} owing to V_{DD} variations can be written as,

$$\Delta V_{REF} = m_{10} V_T \ln\left(\frac{I_{DMAX}}{I_{DMIN}}\right) = \frac{m_{10}}{m^*}\lambda_D \Delta V_{DS} \approx \lambda_D \Delta V_{DS} \,, \tag{3.78}$$

where m^* is the subthreshold slope factor of the high impedance transistor.
Therefore, the line sensitivity can be rewritten as,

$$LS \approx \frac{\lambda_D}{V_{REF}} \frac{\Delta V_{DS}}{\Delta V_{DD}} \% \approx \frac{\lambda_D}{V_{REF}} \% , \tag{3.79}$$

As expected, for line sensitivity minimization, the high impedance transistors in the proposed configuration (i.e. M_1, M_7, M_9) should be as long as possible, in order to improve the stability of the generated current and, therefore, the one of V_{REF}.
The result found in (3.79) not only gives a quick evaluation of the expected line sensitivity of the proposed voltage reference configuration, but can also be extended to every voltage reference architecture operated in subthreshold region, where only one high impedance transistor per branch is present. Such architectures are very common in ultra low voltage solutions, where the reduction of transistor stacks helps in lowering the minimum V_{DD} achievable. In [18], where a two-transistor stack per each branch is present, a minimum V_{DD} of 0.9 V was attained.
The DIBL factor was measured for a PMOS of the same channel length as M_9, the transistor that absorbs V_{DD} variations in the active load branch, obtaining a λ_D of about 0.001. From (3.79), therefore, the expected LS is 0.388 %/V, only 13% different from the measured mean line sensitivity, due mostly to the subthreshold slope factor difference between PMOS and NMOS, which has been assumed equal in (3.78).
3.5.2.1.7. Process variations
The generated voltage (3.78) is a linear combination of threshold voltages parameters. As usually happens in low-power CMOS integrated voltage references [18, 55-57], the reference voltage is process dependent. Thus, neglecting matching errors in (3.78), the accuracy of the output voltage is mainly due to the accuracy of threshold voltages of transistors M_1, M_3, M_2, M_{10}. Process variations are generally distinguished in WIthin Die (WID or intra-die) variations and Die to Die (D2D or inter-die) variations [55]. The first type of variations causes mismatch between transistors of the same chip and influences the relative accuracy of transistor parameters. Careful layout techniques [64] and large transistor W/L [61], can help to reduce those effects. D2D variations, instead, influence the absolute accuracy of transistor parameters and their effects are not compensated in the proposed configuration.
Threshold voltage measurements over a set of 40 SVT-NMOSs, with a W/L of 12 µm / 50 µm, were performed. Each transistor belongs to a different die and is placed within the same voltage reference dies. The Vth distribution is shown im Fig. 3.20.
The average and the standard deviation values obtained are 318.5 mV and 6.7 mV, respectively. Expressing V_{REF_OPT} as the threshold voltage difference between an HVT-NMOS and an SVT-NMOS, the standard deviation of V_{REF_OPT} (σ_{VREF}) can be approximated with,

$$\sigma_{VREF} = \sqrt{\sigma_{HVT}^2 + \sigma_{SVT}^2} , \tag{3.80}$$

where σ_{HVT} and σ_{SVT} are the V_{th} standard deviations of an HVT-NMOS and an SVT-NMOS, respectively. In (3.80) has been implicitly assumed that the variation

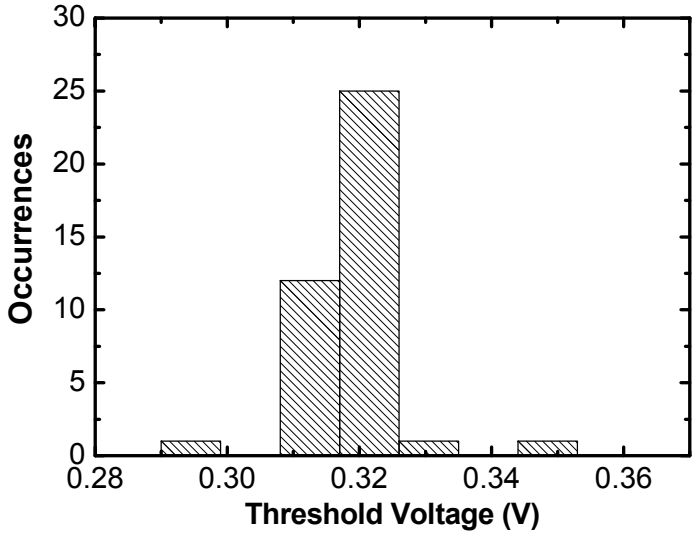


Fig. 3.20 – Measured distribution of V_{th} over a set of 40 SVT-NMOS.

processes of the two V_{th}s are independent, i.e. their covariance is zero. Moreover, assuming that $\sigma_{HVT}=\sigma_{SVT}=\sigma_{Vth}$, it is easy to evaluate the standard deviation of the voltage reference as,

$$\sigma_{VREF} \approx \sqrt{2}\sigma_{V_{th}} , \tag{3.81}$$

The σ_{VREF} obtained from (3.81) is 9.5 mV, only 0.5 mV lower than the measured one. This testifies that the main contributor to the voltage reference inaccuracies are the NMOS threshold voltage variations.

On the other hand, since the threshold voltage dependence of temperature, $\partial V_{th}/\partial T$ has small dependence on process variations, as demonstrated in [55], even the temperature coefficient of the reference voltage will benefit from a similar behaviour with process variations.

3.5.2.1.8. Circuit layout

The circuit has been implemented in UMC 0.18 µm, 1.8 V/3.3 V, CMOS process. The photo of the chip wherein the circuit has been included is shown in Fig. 3.21. In the chip are also present the other voltage reference configurations, with and without start-up. NMOS and PMOS sections indicate the transistors used for the experimental characterization, i.e. threshold voltage and subthreshold slope parameter extraction. The sizes of the chip are 1.5 mm × 1.5 mm, these are the allowed minimum sizes by the Multi Project Wafer of the Europractice service for the UMC 0.18 µm technology.

Due to the passivation layer, is not possible to see the underneath layers. The circuit layout realized with Cadence Virtuoso is reported in Fig. 3.22. Each voltage reference needs three pads: ground, power supply and output reference voltage. As ESD protections a reverse biased MOSFET diode has been placed between V_{DD} and ground pads. Clamp diodes has been used as protection against voltage

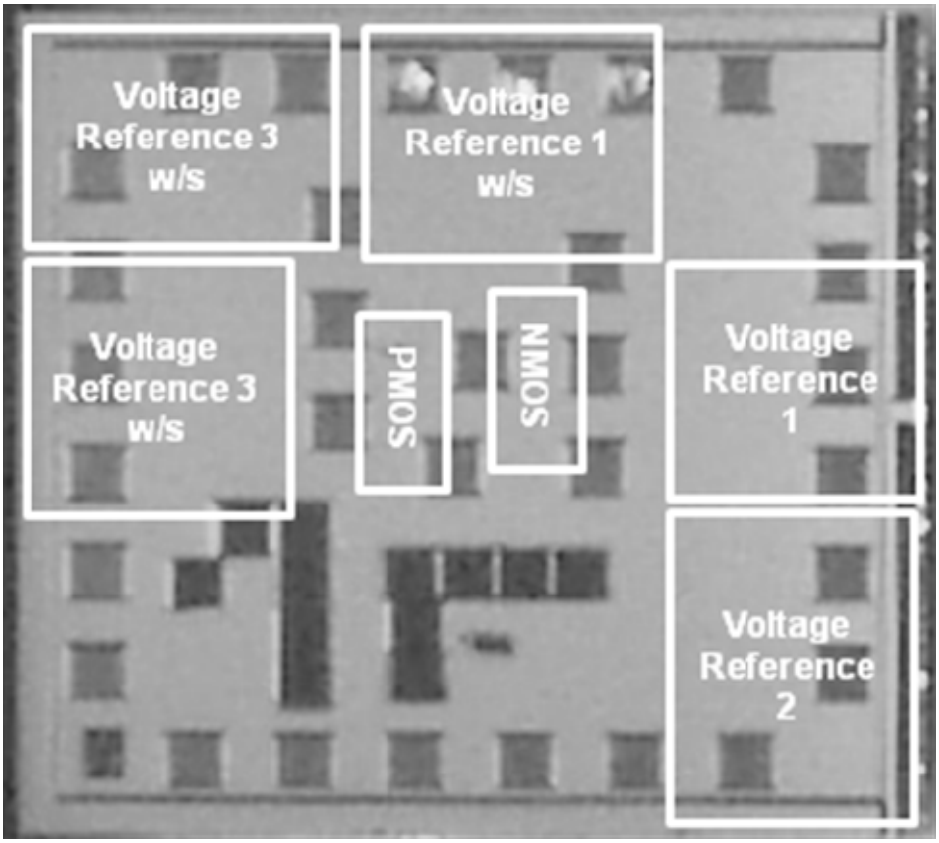


Fig. 3.21 – Chip photo.

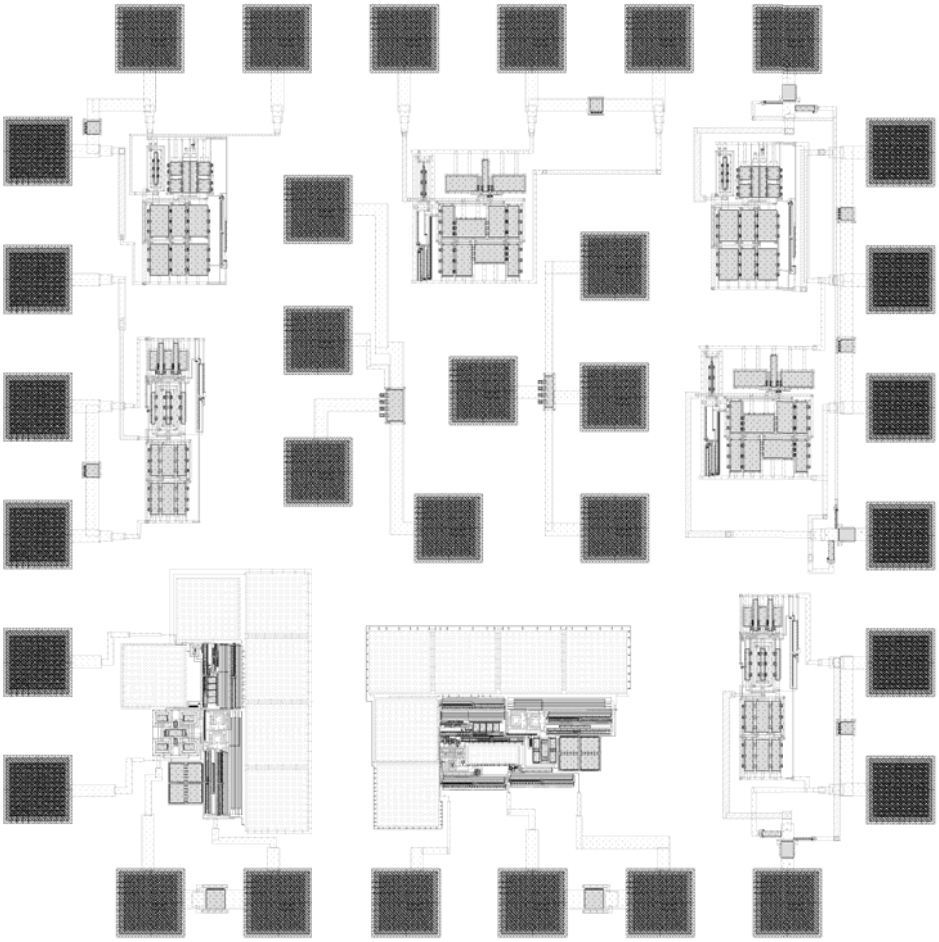
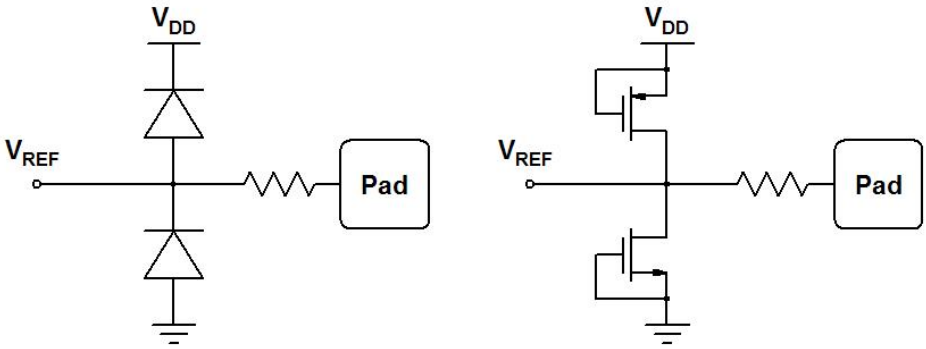


Fig. 3.22 – Layout of the chip.



(a) (b)

Fig. 3.23 – (a) Clamp diodes against voltage overshoots, (b) MOSFET implementation.

peaks on the reference voltage pad, as shown in Fig. 3.22 (a).Drain-body pn-junctions have been employed as clamp diodes (Fig. 3.22 (b)).

A zoom of the circuit layout on the voltage reference 1 with start-up circuit is shown in Fig. 3.24. The circuit area overhead due to the start-up section is only a small fraction of the whole circuit. The occupied chip area is just ≈0.0430 mm^2, thanks mainly to the reduced number of transistors in the core of the circuit.

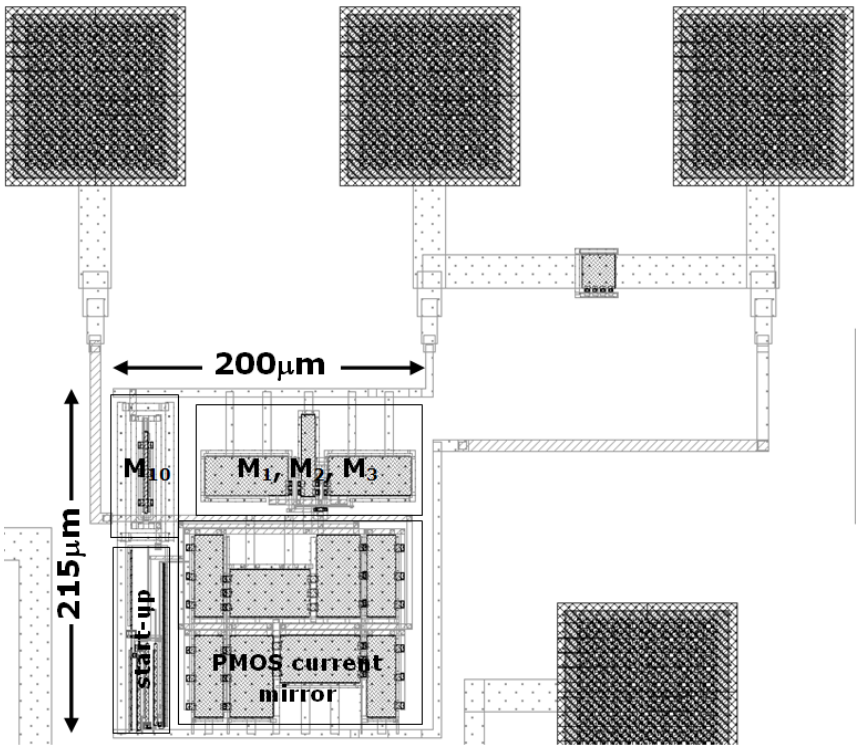


Fig. 3.23 – Layout of the proposed voltage reference 1.

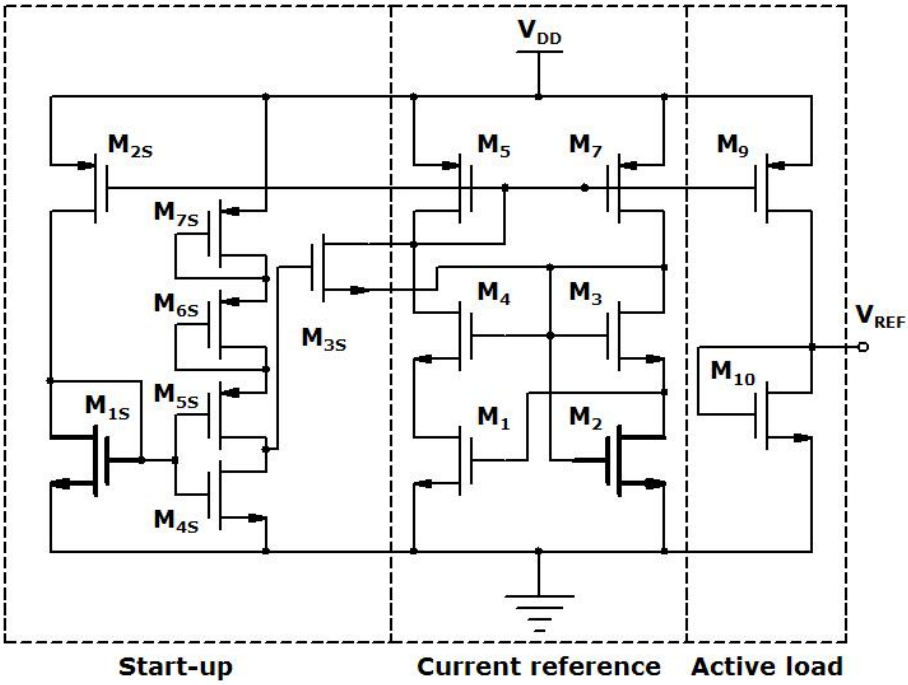


Fig. 3.24 – Voltage reference 2 circuit configuration.

3.5.2.2. Voltage reference 2

The architecture of the proposed voltage reference 2 with start-up [75] is shown in Fig. 3.24. The circuit configuration is very similar to that of voltage reference 1. The same start-up circuit has been employed. The operating principle for temperature compensation is the same adopted previously. Indeed, the current reference generates a current (in the form (3.68)) based on the V_{GS} combination of transistors $M_{1\text{-}3}$ and injects it into the active load. Transistor sizes of the current reference and the active load are reported in table 3.3. Similar design considerations done for the previous configuration are valid for the voltage reference 2.

The main difference is the presence of the high impedance transistor M_4 in the left branch of the current reference.

3.5.2.2.1. Improved self-biased current reference configuration

A useful parameter for describing the variation of the output voltage with the power supply is the sensitivity,

$$S_{V_{DD}}^{V_{REF}} = \frac{\partial V_{REF}/V_{REF}}{\partial V_{DD}/V_{DD}} = S_{V_{DD}}^{I_{REF}} S_{I_{REF}}^{V_{REF}} = \frac{\partial I_{REF}/I_{REF}}{\partial V_{DD}/V_{DD}} \frac{\partial V_{REF}/V_{REF}}{\partial I_{REF}/I_{REF}}, \qquad (3.82)$$

For the circuit proposed is helpful to consider the sensitivity of the reference voltage as the product of two sensitivities (as shown in the relations above): the one of the reference current with the power supply and the sensitivity of the reference voltage with the generated reference current. In order to improve the sensitivity of the circuit these two sensitivities should be reduced.

Table 3.3
Transistor sizes of the current reference and the active load of the voltage reference 2.

Transistor	Value (*W/L*)
M_1	12 μm / 50 μm = (6 μm / 50 μm) × 2
M_2	36 μm / 30.5 μm = (12 μm / 36.5 μm) × 3
M_3	1.5 μm / 1.5 μm
M_4	12 μm / 50 μm = (6 μm / 50 μm) × 2
M_5	26 μm / 50 μm = (13 μm / 50 μm) × 2
M_7	26 μm / 50 μm = (13 μm / 50 μm) × 2
M_9	40 μm / 50 μm = (20 μm / 50 μm) × 2
M_{10}	6 μm / 50 μm

In the voltage reference 1 configuration, only one high impedance transistor per branch is present in the current generator, therefore, the only way to reduce the I_{REF} sensitivity to V_{DD} variations ($S_{V_{DD}}^{I_{REF}}$) is to use channel lengths as long as possible for those transistors.
Nevertheless, changing circuit configuration it is possible to significantly reduce $S_{V_{DD}}^{I_{REF}}$. In fact, using a cascode-like current mirroring scheme in the self-biased current reference allowed to significantly improve the overall line sensitivity. In particular M_4 in Fig. 3.24 acts as a cascode device, it shields M_1 from V_{DD} variations. It is worth pointing out that M_1 is involved in the current generation: in the configuration of Fig. 3.24 the V_{GS}s combination that leads to the generation of current (3.68) remains (3.66). In the voltage reference 1 the drain current of M_1 can vary according to V_{DD} variations, because of DIBL effect, in the new circuit configuration the I_{D1} variation is significantly reduced since the ΔV_{DS} of M_1 due to a ΔV_{DD} is dramatically reduced thanks to the additional cascode device M_4. This at the cost of the additional voltage headroom consumed by M_4.
HSPICE simulation results of the sensitivities in (3.82) are shown in Fig. 3.25 and 3.26 for voltage reference 2 and 1, respectively. With V_{DD}s higher than the V_{DDmin} for correct operation (0.45 V in voltage reference 1 and 0.6 V in voltage reference 2), the sensitivities of V_{REF} to I_{REF} variations ($S_{I_{REF}}^{V_{REF}}$) are quite similar for both circuit configurations, since the active load branches are the same as well as the channel lengths of the high impedance element in that branch. Significant reduction of $S_{V_{DD}}^{I_{REF}}$ with the new current reference configuration is obtained. The overall sensitivity of V_{REF} to V_{DD} variations in voltage reference 2 is roughly reduced by a factor 10x with respect to the other configuration. This result will be verified by line sensitivity measurements in the experimental results paragraph.

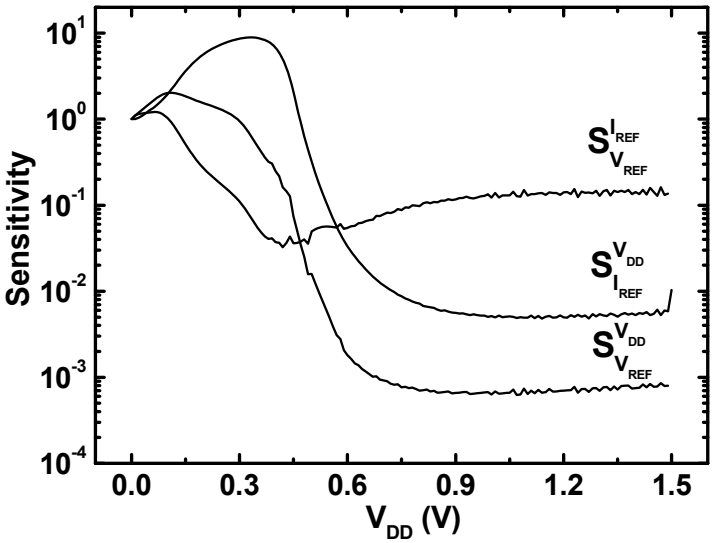


Fig. 3.25 – $S^{V_{REF}}_{V_{DD}}, S^{I_{REF}}_{V_{DD}}, S^{V_{REF}}_{I_{REF}}$ *versus* V_{DD} *for voltage reference 2.*

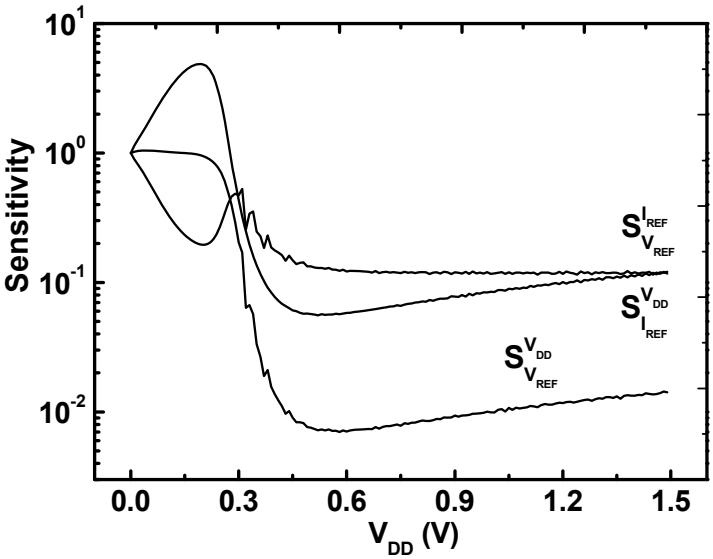


Fig. 3.26 – $S^{V_{REF}}_{V_{DD}}, S^{I_{REF}}_{V_{DD}}, S^{V_{REF}}_{I_{REF}}$ *versus* V_{DD} *for voltage reference 1.*

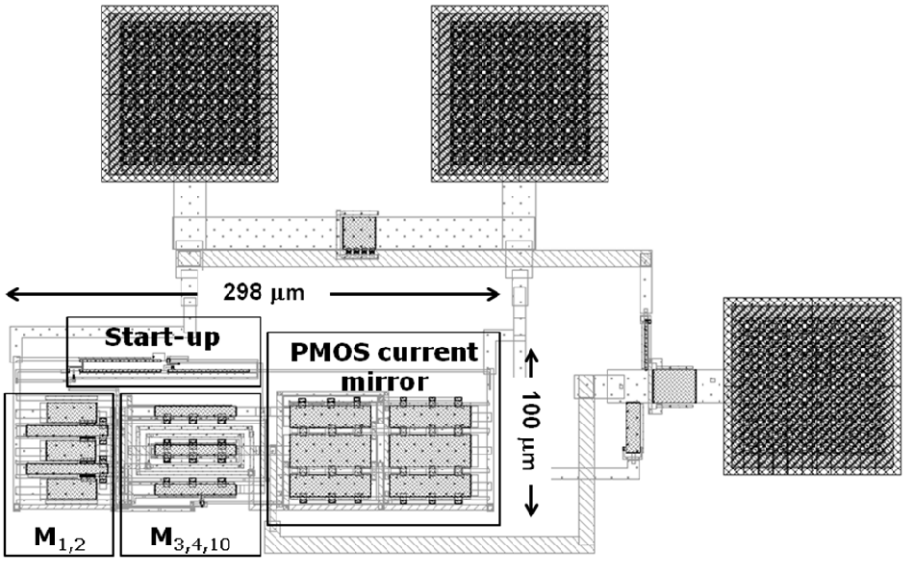


Fig. 3.27 – Layout of the proposed voltage reference 2.

3.5.2.2.2. Circuit layout

Circuit layout of the voltage reference 2 with and without the start-up section have been implemented in the UMC 0.18 µm process. The circuit layout with start-up circuit is shown in Fig. 3.27. The occupied chip area is ≈0.0298 mm^2.

3.5.2.3. Voltage reference 3

The circuit configuration of the voltage reference 3 is directly derived from the subthreshold PTAT current reference proposed in [36]. Furthermore, when operated with saturated MOSFETs in strong inversion, the current reference presented in [36] is the well known β–multiplier. The circuit schematics and the MOSFET sizes are shown in Fig. 3.28 and TABLE 3.4, respectively. Again, the temperature compensation technique is the same developed before as the start-up circuit. Transistors $M_{1\text{-}3}$ perform the V_{GS}s combination for the generation of the current in the form (3.68). Transistors M_1 and M_{1S} are HVT-MOSFETs, while the other transistors within the circuit are SVT-MOSFETs. Only one high impedance transistor per branch absorbs V_{DD} variations, therefore the same line sensitivity of voltage reference 1, if same channel lengths for $M_{1,7,9}$ are used, is expected.

3.5.2.3.1. Circuit layout

The circuit layout of voltage reference 3 has been implemented in the UMC 0.18 µm CMOS process and is shown in Fig. 3.29. Only the configuration with the start-up circuit has been fabricated for that configuration. The occupied chip area is ≈0.0322 mm^2.

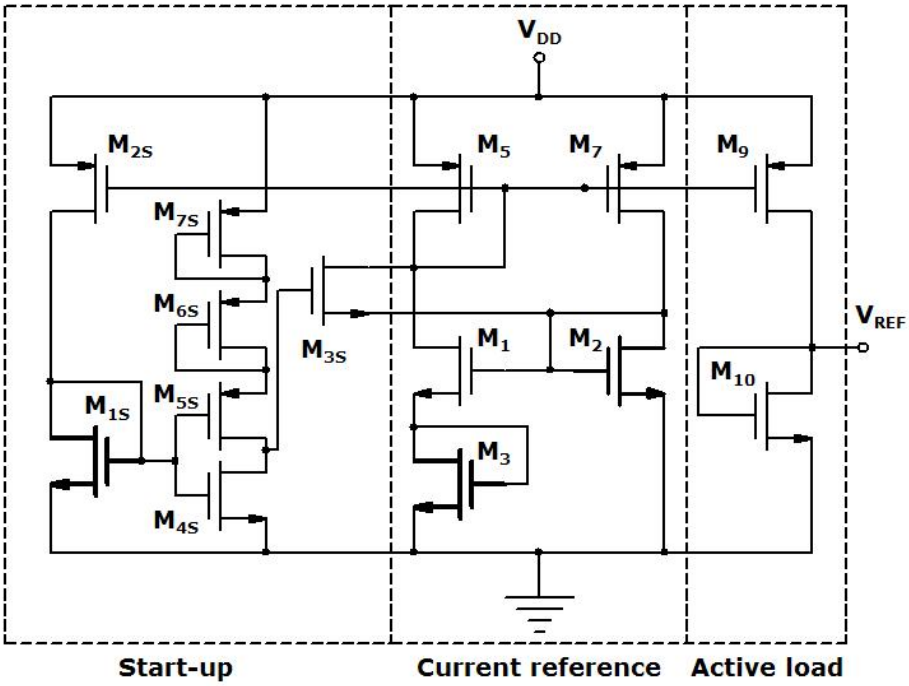


Fig. 3.28 – Voltage reference 3 circuit configuration.

Table 3.4

Transistor sizes of the current reference and the active load of the voltage reference 3.

Transistor	Value (*W/L*)
M_1	12 μm / 50 μm = (6 μm / 50 μm) × 2
M_2	36 μm / 30.5 μm = (12 μm / 36.5 μm) × 3
M_3	1.5 μm / 1.5 μm
M_5	26 μm / 50 μm = (13 μm / 50 μm) × 2
M_7	26 μm / 50 μm = (13 μm / 50 μm) × 2
M_9	40 μm / 50 μm = (20 μm / 50 μm) × 2
M_{10}	6 μm / 50 μm

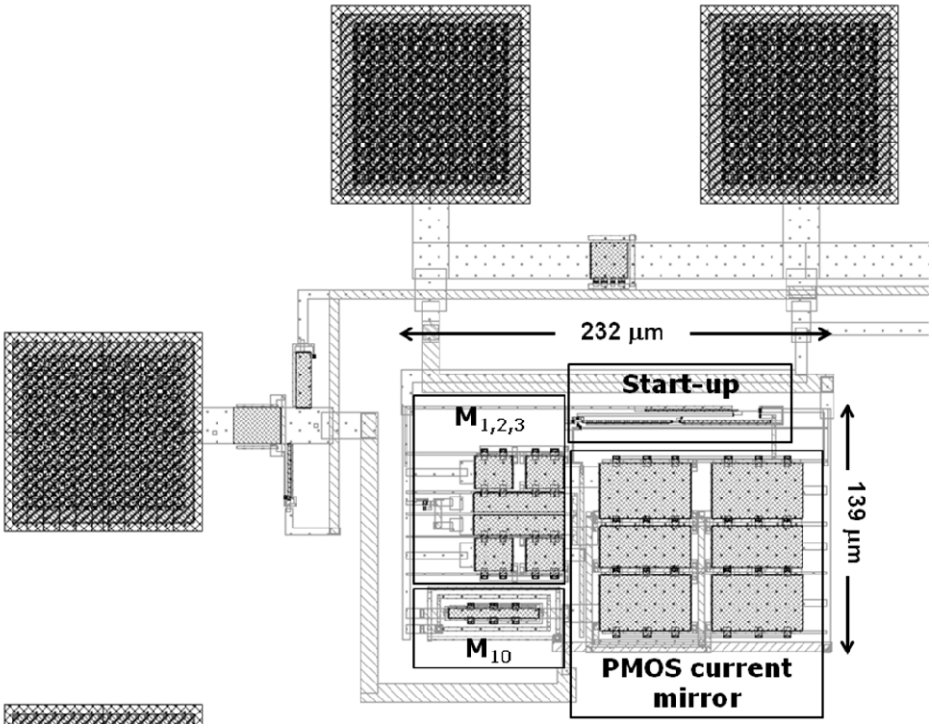


Fig. 3.29 – Layout of the proposed voltage reference 3.

3.5.3. Experimental results

DC measurements were performed at wafer-level DC using a Cascade SUMMIT 11861B prober equipped with a Temptronic chuck temperature controller and a Keithley 4200-SCS semiconductor parameter analyzer. Noise measurements were performed connecting the prober to the National Instruments PXI-4472 dynamic signal acquisition module, equipped with anti-aliasing filters.

In this section, measurement results of the three proposed voltage reference configurations are presented and discussed. Voltage references 1 and 2 have been implemented with and without the start-up circuit. Since no particular differences between the two configurations in terms of performance and power consumption have been observed, only measurements of one typology (with start-up) are discussed in detail, while measurements of the other typology are simply shown.

3.5.3.1. Voltage reference 1

A set of 40 prototypes implemented in UMC 0.18 µm, 1.8V/3.3 V, CMOS process, were measured at wafer-level. In this section, measurement results of a chip with typical behaviour are firstly discussed. Since voltage reference 1 presents the best TC and the lowest power consumption, among the proposed circuits, such results are compared with the best performing low-power low-voltage solutions found in the literature. Subsequently a statistical analysis of the performance, regarding the 40 samples fabricated, is presented. As a typical performance chip, the one with

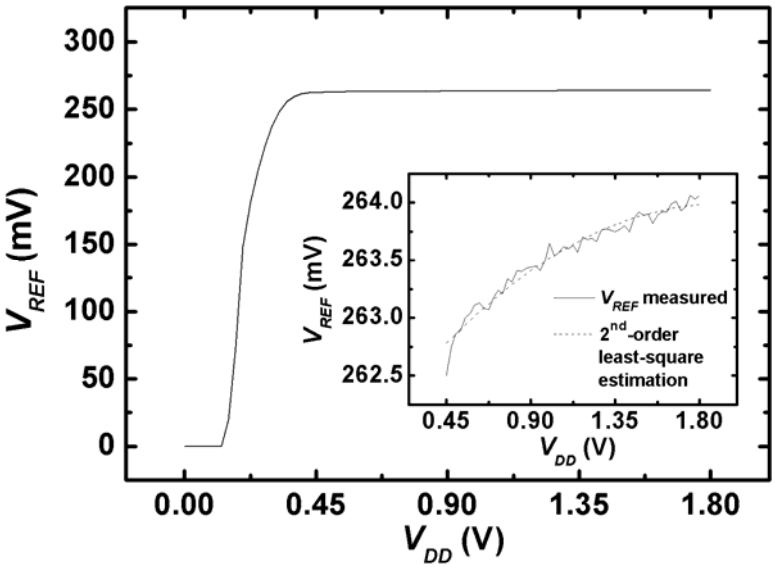


Fig. 3.30 – Measured output voltage as a function of power supply at room temperature and zoom in the V_{DD} operating range.

the median temperature coefficient (TC) was chosen, since it is, commonly, the most critical figure of merit for voltage reference circuits. However, as will be clear from statistical analysis, those typical results are very close to the average ones.

In Fig. 3.30, the V_{REF}-V_{DD} characteristic at room temperature is reported. The circuit starts working properly with V_{DD}=0.45 V. In the supply voltage range from 0.45 V to 1.8 V a mean reference voltage of 263.5 mV is generated. In this supply voltage range, the output voltage changes at most by 1.6 mV, thus leading to a line sensitivity of 0.440 %/V.

The current consumption changes slightly with V_{DD}, at room temperature it varies from 7.0 nA to 7.9 nA in the supply voltage range under consideration. Therefore, under the above-mentioned operating conditions, the minimum power consumption is just 3.2 nW. In Fig. 3.31 (a) the measured values of the current drawn from the power supply, in the temperature range between 0 °C and 125 °C, are shown. The supply current increases with temperature according to (3.68), reaching 36.0 nA at V_{DD}=0.45 V@125 °C, and 40.8 nA at V_{DD}=1.8 V@125°C.

The reference voltage dependence on temperature is shown in Fig. 3.31 (b). Averaging the V_{REF} variation on the V_{DD} range considered, it results ≈4.7 mV, leading to an average TC of 142 ppm/°C. From Fig. 3.31 (b), it is clear that the TC is quite independent of V_{DD}.

The measured and simulated PSRR values at room temperature, at minimum supply voltage, are illustrated in Fig. 3.32. Because of the very high value of the output impedance of the circuit, PSRR measurements have been performed by connecting the voltage reference output to a CMOS operational amplifier (TLC2201). Indeed, from post-layout simulations, the evaluated output impedance

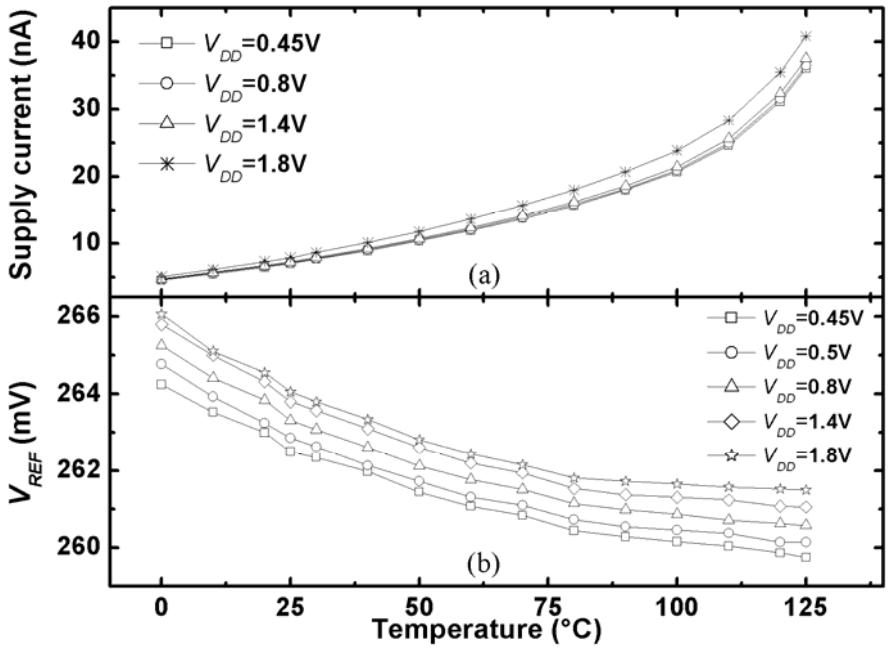


Fig. 3.31 – (a) Supply current versus temperature for different supply voltages. (b) Temperature dependence of the generated reference voltage for different

(Z_{OUT}) is 55 MΩ//0.3 pF, higher than typical input impedance of commercial oscilloscope probes. The opamp input capacitance (C_{IN}) is about 20 pF (considering packaged amplifiers, generally, C_{IN} does not go below 10 pF) and forms a low-pass filter with Z_{OUT} having a -3 dB frequency of 144 Hz. For this reason, the high output impedance of the circuit did not allow an accurate measurement of the PSRR for frequencies above 30 Hz. However, from Fig. 3.32, it is clear that the simulation results are in good agreement with the measurements obtained.

The measured equivalent output noise amplitude, without filtering capacitors, is shown in Fig. 3.33. At 10 Hz its value is 2.3 μV/√Hz and the root mean square voltage noise, measured in a bandwidth from 0.1 Hz to 10 Hz, reported in Fig. 3.34, is 22.0 μV.

It is clear that the designed circuit can only drive high impedance loads such as the capacitive loads offered by MOSFET-based comparators or buffers. Most often voltage references do not drive directly loads, as in the case of voltage supply circuits [58].

TABLE 3.5 summarizes the measurement results of the typical performance chip and compares them with the best low voltage, low power voltage references found in literature, implemented in standard CMOS process. From the comparison it results that the proposed configuration achieves the lowest supply voltage and current dissipation, thus leading to at least a 10× reduction in power consumption with respect to the other proposed solutions. Moreover, those results are obtained

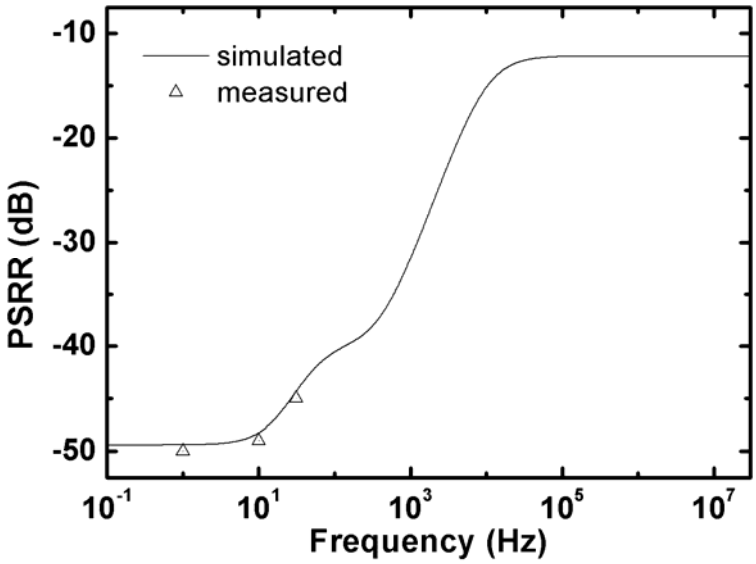


Fig. 3.32 – PSRR at room temperature, for V_{DD}=0.45V

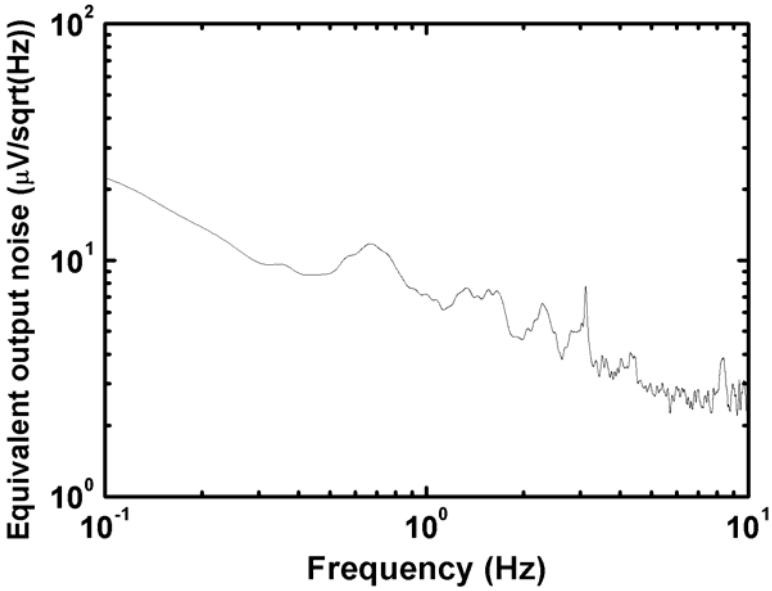


Fig. 3.33 – Measured equivalent output noise (V/√Hz)from 0.1 Hz to 10 Hz.

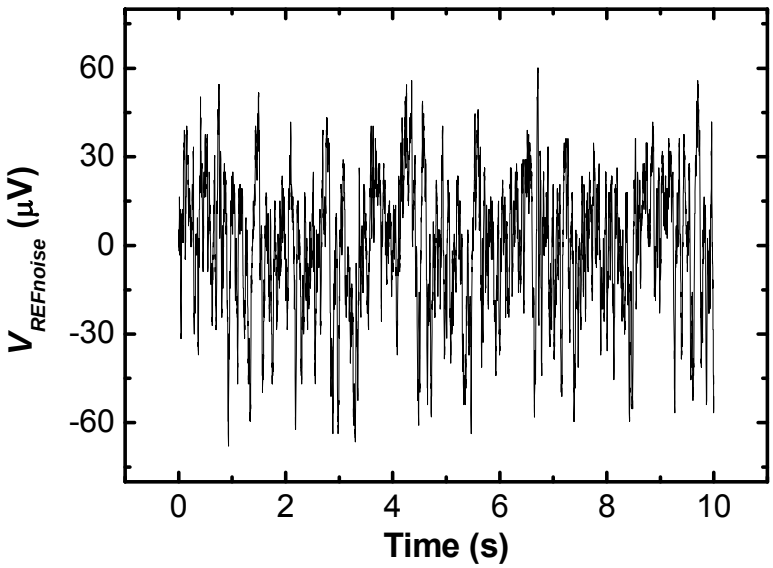


Fig. 3.34 – Output noise fluctuations versus time corresponding to a measurement bandwidth from 0.1 Hz to 10 Hz.

while preserving competitive line sensitivity, temperature coefficient, PSRR and area occupation.

Measurements were carried out over a set of 40 and 38 samples for configuration with and without start-up circuit. In TABLE 3.6, a brief summary of the statistical analysis results, reporting most relevant figures of merit, is given. In particular mean (μ) and standard deviation (σ) values of the temperature coefficient (TC_{AVG}), the line sensitivity (LS), the power consumption at V_{DD}=0.45 V (P_{DISS}) and the reference voltage (V_{REF}) at room temperature were reported. TC_{AVG} and V_{REF} refer to the averaged values on the V_{DD} range considered.

In Fig. 3.35 the distributions of these parameters for the circuits with the start-up section are shown. To our knowledge, detailed statistical information about key performance indicators like TC, LS or power consumption, are not given in works concerning low power voltage references. For this reason it was not possible to compare the statistical results in detail.

The mean power consumption at the minimum V_{DD} and at room temperature is 2.6 nW and it remains lower than 5 nW in the worst case (see Fig. 3.35 (a)). However, even in that case, the proposed circuit continues to achieve the lowest power consumption, considering that the best low power solutions reported in literature do not go below 23 nW [56].

The line sensitivity distribution, in the V_{DD} range from 0.45 V to 1.8 V, is shown in Fig. 3.35 (b) and the mean value is 0.444 %/V. It varies from a minimum value of 0.329 %/V to a maximum value of 0.606 %/V.

Table 3.5

Comparison with low-voltage low-power CMOS Voltage References

	Voltage reference 1 [19]	*Ref. [18]*	*Ref. [57]*	*Ref. [51]*	*Ref. [55]*	*Ref. [56]*	*Ref. [54]*
Technology	0.18 µm CMOS	0.35 µm CMOS	0.18 µm CMOS	0.18 µm CMOS	0.35 µm CMOS	0.35 µm CMOS	0.6 µm CMOS
Supply voltage (V)	**0.45 to 2**	**0.9 to 4**	**0.6 to 2.3**	**0.85 to 2.5**	**1.4 to 3**	**1.1 to 4**	**1.4 to 3**
Supply current (µA) @ room temperature	**0.007 @0.45V** **0.008 @1.8V**	**0.040 @0.9V** **0.055 @4V**	**<0.040 @0.7V** **–**	**3.882 @0.85V** **average**	**0.2143 @1.4 V** **-**	**~0.021 @1.1V** **~0.024 @4V**	**<9.7**
V_{REF} (mV)	263.5	670	~220	221	745±25	96.6±4.0 vers-1 108.9±3.1 vers-2	309.3±19.26
TC (ppm/°C), T range(°C)	142 [0:125]	10 [0:80]	127 [-20:100]	194 [-20:120]	7 [-20:80]	11.4 – [-20:80] , vers-1 9.2 – [-20:80] , vers-2	36.9 [0:100]
Line Sensitivity (%/V)	0.44	0.27	~2.730	0.905	0.002	0.090 , vers-1 0.170 , vers-2	0.083
PSRR (dB)	V_{DD}=0.45V	V_{DD}=0.9V	-	-	V_{DD}=2V	V_{DD}=3V	V_{DD}=1.4V
Low freq [≤100Hz]	-45	-47	-41	-	-45	<-60	-47
High freq [≥10MHz]	(-12.2 sim.)	-40	-	-	~ -22@10kHz	<-40	-20
Die area (mm^2)	0.043	0.045	0.004	0.0238	0.055	0.0189 , vers-1 0.0193 , vers-2	0.055

Table 3.6
Statistical analysis of performance of voltage reference 1

Configuration and number of samples	With start-up 40		Without start-up 38	
	μ	σ	μ	σ
TC_{AVG} (ppm/°C)	165	100	183	135
LS (%/V) @ 25°C	0.444	0.058	0.434	0.066
P_{diss} (nW) @ 0.45V & 25°C	2.6	0.7	3.3	0.8
V_{REF} (mV) @ 25°C	257.5	10.0	265.6	9

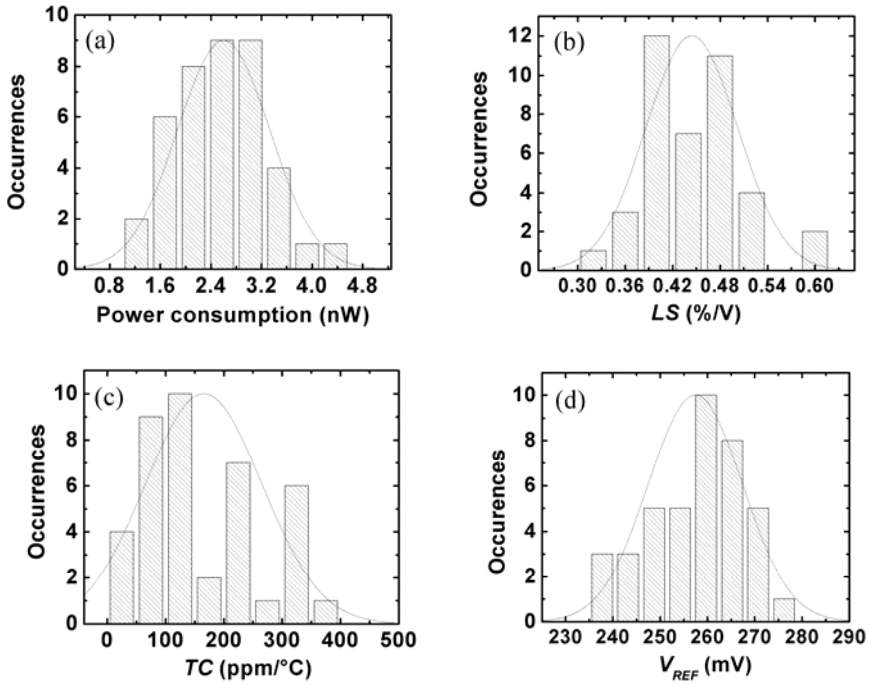


Fig. 3.35 – Distributions of the most relevant figures of merit for the 40 measured samples: (a) Power consumption @ 25 °C and V_{DD}=0.45 V. (b) Line sensitivity @ 25°C. (c) Temperature coefficient. (d) Generated voltage reference @ 25°C.

The best TC_{AVG} is 39 ppm/°C, which means an average variation of the output voltage of 1.2 mV in the temperature range from 0 °C to 125 °C. The worst TC_{AVG} (357 ppm/°C), instead, gives an average variation of about 12 mV.
From Fig. 3.35 (c) it is evident that the TC is the quantity with the largest dispersion, indeed, the coefficient of variation σ/μ is ≈61%. The temperature

coefficient is a very sensitive parameter, in [48] a 4-bit trimming network is used to optimize it, by changing resistor values after fabrication.

Fig. 3.35 (d) shows the distribution of V_{REF} at room temperature. The σ/μ is 3.9%, such spread in reference voltage is higher in the proposed circuit than when using BGRs, e.g. in [65] the σ/μ is around 1.5%. However, the proposed solution exhibits a power consumption and a minimum supply voltage significantly lower compared to bandgap-like circuits.

3.5.3.2. Voltage reference 2

Measurements over a set of 37 and 35 samples of voltage reference 2 with and without start-up circuit, respectively, were carried out. A significant improvement in line sensitivity has been observed. As expected from sensitivity simulations (see paragraph 3.5.2.2) the line sensitivity in such case improved roughly by a factor 10 with respect to voltage reference 1. In Fig. 3.36 the V_{REF}-V_{DD} characteristics is shown and a zoom in the operating V_{DD} range is also provided. In the V_{DD} range [0.6;1.8] V, V_{REF} varies just a few tenths of millivolts, thus leading to a line sensitivity of 0.065 %/V. Notwithstanding, adding a transistor in the stack of the left branch of the current reference (see Fig. 3.24), increased the V_{DDmin} from 0.45 V to 0.6 V.

TABLE 3.7 summarizes the measurement results, providing mean and standard deviation values of the most relevant figures of merit. Again, no particular differences have been observed between the behaviors of circuits with and without start-up. Although the circuit shows a remarkable insensitivity to V_{DD} variations, unfortunately the temperature coefficient does not follow the same trend. The TC of

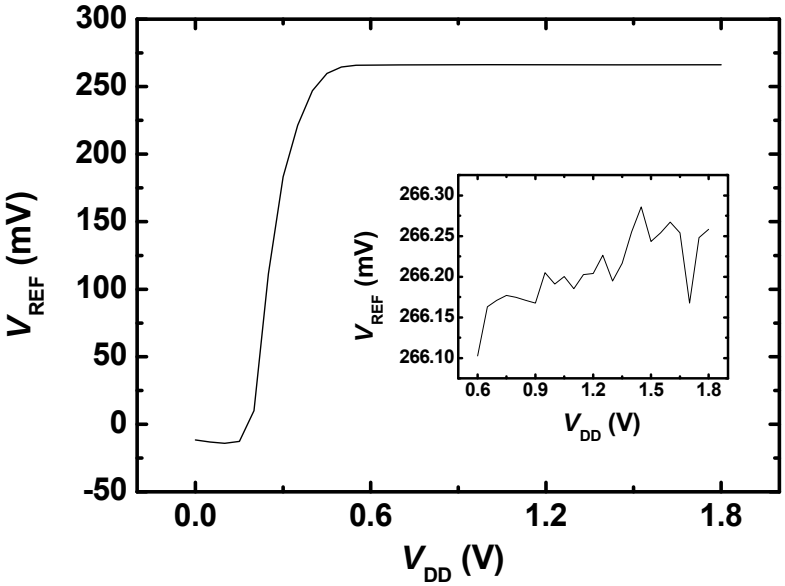


Fig. 3.36 – Measured V_{REF}-V_{DD} and zoom in the V_{DD} operating range [0.6:1.8] V.

voltage reference 2 is around three times higher than that measured on voltage reference 1, even though the temperature compensation technique employed is the same. In order to discover the worst case behavior of the circuit, process corner simulations of the post-layout circuit configuration were performed before fabrication. Nevertheless they showed a worst case TC of only 32 ppm/°C.

Table 3.7

Statistical analysis of performance of voltage reference 2

Configuration and number of samples	With start-up 37		Without start-up 35	
	μ	σ	μ	σ
TC_{AVG} (ppm/°C)	462	134	463	141
LS (%/V) @ 25°C	0.065	0.041	0.070	0.047
P_{diss} (nW) @ 0.45V & 25°C	22.3	5.4	21.9	5.3
V_{REF} (mV) @ 25°C	259	9.8	258.3	9.8

Probable mispredictions of the subthreshold behavior with the employed BSIM3v3.2 models might have occurred. However, as usually happens in the design of integrated circuits, a satisfactory result in post-layout simulations is still no guarantee for a completely successful product (there is no substitute for “real silicon”).

Power consumption is higher than the previous solution, while V_{REF} spread remains quite similar.

3.5.3.2. Voltage reference 3

A set of 39 samples of voltage reference 3 have been fabricated and tested. Measurement results are shown in TABLE 3.8.

Even in this case, the measured TC is significantly different from those predicted by post-layout simulations. The worst case TC from process corner simulations is about 20 ppm/°C. Line sensitivity remains similar to that measured on voltage

Table 3.8

Statistical analysis of performance of voltage reference 3

Configuration and number of samples	With start-up 37	
	μ	σ
TC_{AVG} (ppm/°C)	246	134
LS (%/V) @ 25°C	0.357	0.041
P_{diss} (nW) @ 0.45V & 25°C	15.4	5.4
V_{REF} (mV) @ 25°C	242.1	9.8

reference 1, indeed, in the circuit configuration of the current reference only one transistor per branch absorbs V_{DD} variation as for voltage reference 1.

4. SUBTHRESHOLD OPERATIONAL AMPLIFIER DESIGN

4.1. Introduction

Lowering the supply voltage is the main approach to reduce power consumption in digital electronic circuits. Moreover, considering relentless device dimension scaling, reliability issues make supply voltage reduction necessary. Nevertheless, reducing the supply voltage poses significant challenges in designing operational amplifiers (op-amps), which are almost always present in analog sections of modern mixed-signal integrated circuits. For this reason, op-amp design has become a crucial aspect in order to put to use potential benefits guaranteed by technology scaling.

Different techniques were developed to implement op-amps working at very low supply voltage, such as the use of charge pumps, the use (where possible) of the MOSFET body terminal as signal input or as biasing terminal to lower the device threshold voltage, and the use of transistors operating in subthreshold (or weak inversion) region [66]-[70]. Subthreshold design is particularly gaining attention [67] [69], [70] to implement ultra low-voltage low-power amplifiers characterized by larger output swing and lower bias current with respect to traditional amplifiers working in saturation region.

In this chapter, well-defined design guidelines to build efficient subthreshold op-amps are provided. Subsequently, measurement results of an op-amp designed exploiting such guidelines are shown. The rest of the paper is organized as follows. In the next paragraph ultra low voltage op-amp implementations are described. Subsequently a design procedure for subthreshold op-amps, based on the equation set presented in chapter 2, is reported. A design example of a subthreshold op-amp with a 0.5 V supply voltage and measurement results of prototype chips fabricated in a 0.18 μm CMOS process are finally presented and discussed.

4.2. Low-voltage, low-power CMOS op-amps

The trend towards low voltage systems is pushing to use solutions such as charge-pumping techniques to boost on-chip voltage beyond V_{DD} in order to bias analog sections. However, such techniques consume a significant amount of power and may lead to reliability problems in some cases. Several techniques allowing true low voltage operation without voltage boosting have been developed: the use of subthreshold operated devices, the use of the body terminal as a fourth control or signaling node and the use of circuit topologies avoiding stacked transistors.

4.2.1. Bulk-driven MOSFETs

In order to bias a MOSFET in the saturation region of strong inversion, in which $V_{DSsat}=V_{GS}-V_{th}$, the supply voltage must satisfy the following requirement [66],

$$V_{DD} \geq V_{GS} = V_{DSsat} + V_{th} \,. \tag{4.1}$$

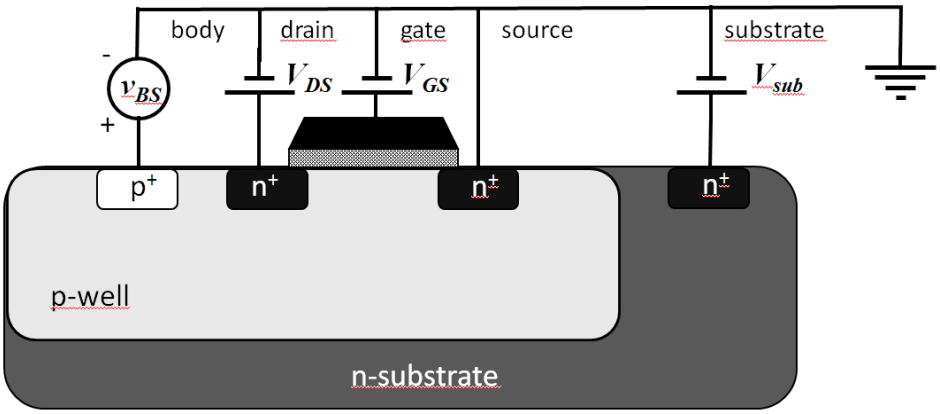


Fig. 4.1 – Cross section of a bulk-driven NMOS (p-well CMOS technology).

Furthermore, if the MOSFET is gate-driven, the supply voltage requirement becomes,

$$V_{DD} \geq V_{GS} = V_{DSsat} + V_{th} + V_{signal} \,. \tag{4.2}$$

Consequently, the threshold voltage requirement constraints signal swing, with regard to the Input Common Mode Range (ICMR), and, hence, the dynamic range.
A widespread solution for low voltage analog design is the bulk-driven MOSFET, Fig. 1. The V_{GS} is taken to a DC voltage >V_{th} to bias the device in strong inversion, the drain is connected normally and the signal is applied between the body and the source (v_{BS}). As a result, the main advantage of bulk-driven MOSFET with respect to traditional gate-driven ones, is the possibility to achieve larger ICMRs, since strong inversion is ensured by V_{GS}, whereas the signal is applied on v_{BS}.
Moreover, considering (2.10f) in TABLE 2.1, the g_{mb} can be theoretically higher than g_m if,

$$V_{BS} \geq 2\varphi_F - 0.25\gamma^2 \approx 0.5 \text{ V} \,, \tag{4.3}$$

although this will exceedingly forward bias the body-source junction thus dissipating unacceptable current levels. One disadvantage of bulk-driven MOSFETs is their input capacitance. While for gate-driven transistors it is the C_{gs} or the series of C_{gs} and C_D in strong and weak inversion, respectively, the input capacitance of bulk-driven ones is the parallel of the body-to-source and the body-to-substrate capacitances (i.e. C_{bs}+C_{bsub}). In [66] it has been demonstrated that the transition frequency (f_{Ti}) of gate-driven MOSFETs in strong inversion is twelve times larger than that of bulk-driven MOSFETs.
In Fig. 4.2 it is shown the schematic of the bulk-driven op-amp proposed in [66]. The input stage consists of a bulk-driven differential pair, $M_{1,2}$, that provides rail-to-rail ICMR. The emitter follower Q_6 serves as a level-shifter between the input stage and the output stage. The output stage uses a NMOS driver, M_7, loaded by the active current source, M_{12}, to obtain a high PSRR. R_z and C_C are the compensation elements for the Miller pole-splitting technique. The op-amp gain-bandwidth product (GBW) and the open-loop gain (A_{v0}) are described by

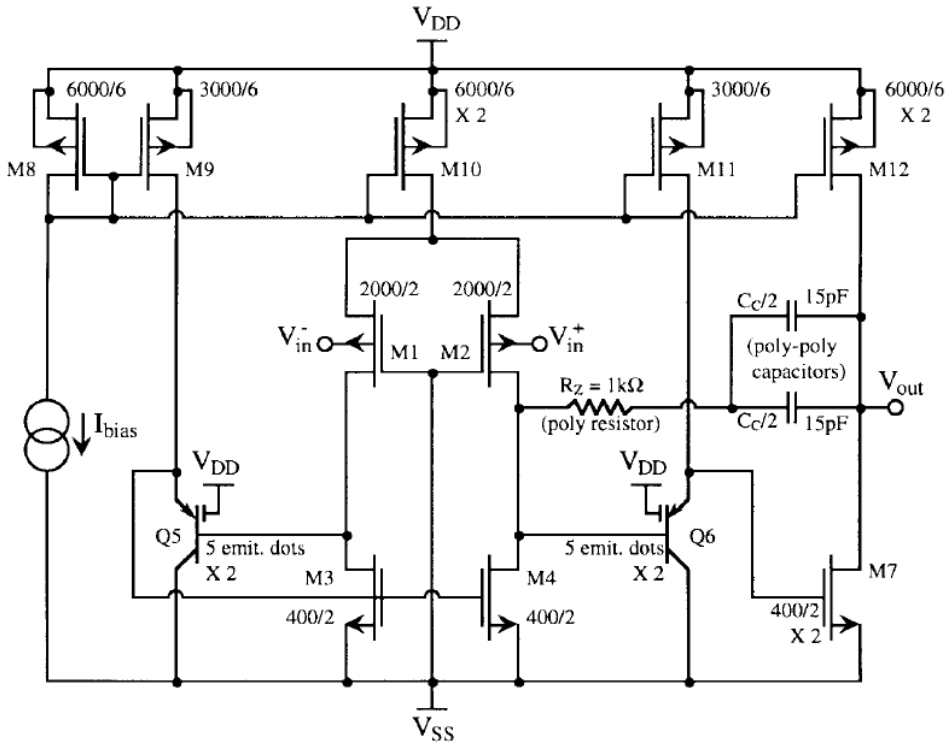


Fig. 4.2 – 1-V CMOS op-amp with bulk-driven differential pair.

$$GBW \approx \frac{g_{mb1}}{2\pi C_C}, \tag{4.4}$$

$$A_{v0} \approx \frac{g_{mb1} g_{m7}}{\left(g_{ds2} + g_{ds4}\right)\left(g_{ds7} + g_{ds12}\right)}. \tag{4.5}$$

The circuit has been implemented in a 2 µm CMOS technology. Bulk-driving allowed to bias the op-amp with a supply voltage as low as the MOSFET V_{th} (0.8 V) plus a V_{DSsat} of 200 mV and a power dissipation of 300 µW. The op-amp achieved rail-to-rail ICMR and output swing. Measurements showed a phase margin of 57° and a unity-gain frequency of 1.3 MHz for a 22 pF load capacitance.

4.2.2. Forward body bias

For high-performance applications that require high bandwidth or sampling rate, MOSFETs are biased in the saturation region of strong inversion, i.e. V_{GS}-V_{th}>0.2 V. Forward biasing of the body-source junction, usually applied in low voltage digital circuits, has been also applied in [68] to lower the MOSFET's V_{th}s within analog modules. In particular, a forward bias of 250 mV has been applied, which results in a reduction of the V_{th} by about 50 mV. The process used for fabrication is a triple well 0.18 m CMOS with standard 0.5 V- V_{th} devices.

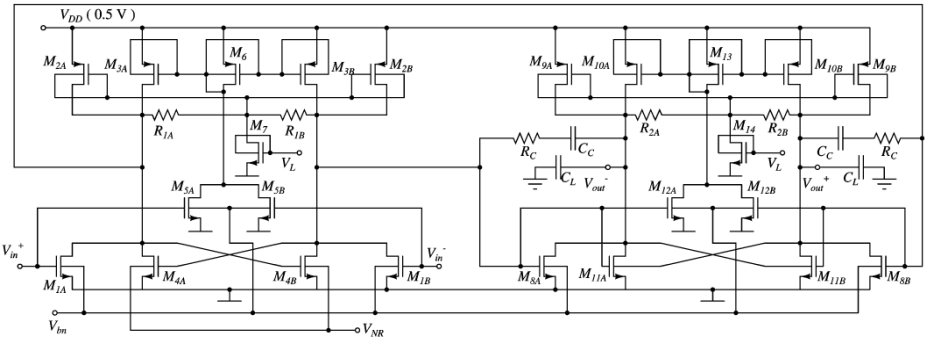


Fig. 4.3 – Two-stage fully differential op-amp with forward body bias.

An op-amp working with a 0.5 V V_{DD}, exploiting forward body bias, has been proposed in [68]. Its circuit configuration is shown in Fig. 4.3. The two stages consist of two identical differential amplifiers cascaded. Considering the input stage, the differential pair M_{1A} and M_{1B} and the active loads M_{2A} and M_{2B} amplify the differential input voltage. Resistors R_{1A} and R_{1B} provide common-mode feedback through the active load. The bodies of M_{2A} and M_{2B} are connected to the gates to further reduce their V_{th}. Moreover, to lower the V_{th} of the input differential pair, the body of M_{1A} and M_{1B} is forward biased.
The overall DC gain is,

$$A_{v0} \approx \frac{g_{m1}}{g_{ds1} + g_{ds2} + g_{ds3} + g_{ds4} + 1/R_1 - g_{m4}} . \tag{4.6}$$

The op-amp is stabilized through the Miller capacitors C_C and resistors R_C across the second stage. The gain-bandwidth product is,

$$GBW \approx \frac{g_{m1}}{2\pi C_C} , \tag{4.7}$$

The circuit dissipates 75 μW and operates with V_{DD}=0.5 V. The DC gain is 62 dB, the unity gain bandwidth and the phase margin are 10 MHz and 60° for a 20 pF load capacitance, respectively.

4.2.3. Subthreshold operation

The main drawback on implementing low voltage CMOS circuit is the threshold voltage. Under weak inversion, the signal swing is larger than in strong inversion due to the low V_{DS}, but the frequency response is reduced due to the extremely low currents. Therefore, a tradeoff between speed and signal swing should be taken into account.
In [67] an op-amp working with transistors operating in weak inversion is presented. The proposed configuration is an improved Miller OTA circuit using composite transistors. The structure of a NMOS composite transistor is shown in Fig. 4.4. From such scheme it follows that,

$$I_{Da} = I_{Db} \ , \ V_{DSa} = V_{GSa} - V_{GSb} . \tag{4.8}$$

Considering saturated MOSFETs in weak inversion, then V_{DSa} is given by (4.9).

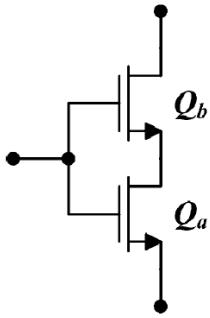


Fig. 4.4 – Composite transistor structure.

This expression is independent of the V_{GS} of the transistor, it only depends on transistor aspect ratios and CMOS parameters.

$$V_{DSa} = V_T \ln\left[1 + \frac{(W/L)_b}{(W/L)_a}\right]. \tag{4.9}$$

The schematic of the modified Miller OTA working with subthreshold operating MOSFETs is shown in Fig. 4.5. A bulk-driven differential pair allows further reduction of the supply voltage. Transistors Q_{3a}-Q_{3b} and Q_{4a}-Q_{4b} form composite transistors. They allow the differential active load and the common gate amplifier to be biased by the same potential. V_{DS3a} and V_{DS4a} are equal and constant. Therefore V_{DS1} and V_{DS2} will be equal and constant too, thus reducing significantly mismatch of the differential pair and, hence, the output offset voltage. The minimum V_{DD} can be defined as,

$$V_{DD\min} = V_{GS1} + V_{DSsat5}\,. \tag{4.10}$$

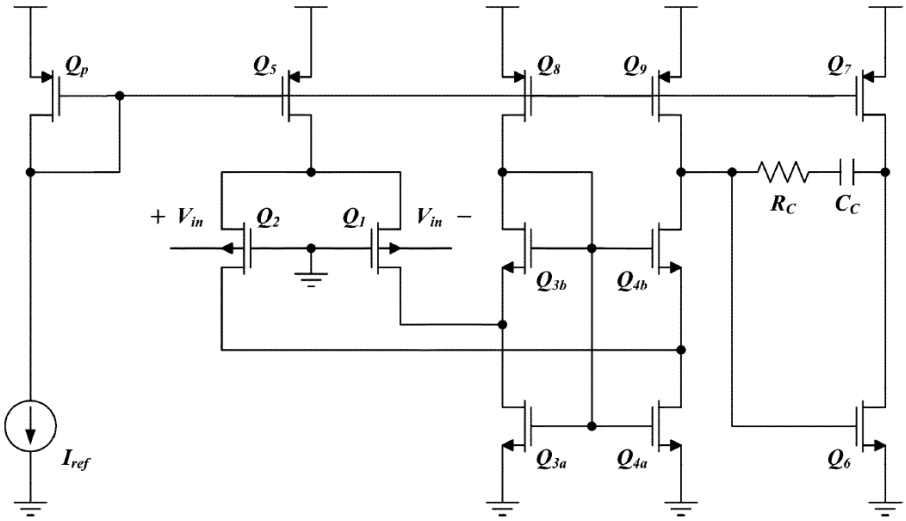


Fig. 4.5 – Modified Miller OTA operating with transistors in weak inversion.

Thanks to the subthreshold operation, V_{GS1} is lower than V_{th} (for NMOS of the considered CMOS process, V_{th} =490 mV)
Since the differential input pair is bulk-driven, the GBW can be approximated with (4.7), while the open loop gain is given by,

$$A_{v0} \approx \frac{g_{mb1} g_{m6}}{\left(g_{ds6} + g_{ds7}\right)\left[\frac{g_{ds4} + g_{ds9}}{g_{mb4} + g_{ds4}}\left(g_{ds2} + g_{ds4}\right) + g_{ds9}\right]}. \tag{4.11}$$

Tests on silicon chips fabricated in a 0.35 μm CMOS process confirmed a V_{DDmin} of 0.6 V and a power consumption of just 550 nW. The measured open loop gain is 69.4 dB, while the unit gain frequency and the phase margin, for a load capacitance of 15 pF, are 11.35 kHz and 65.1°, respectively.

4.3. Design of ultra low-power, low-voltage subthreshold opamps

Subthreshold operation promises several advantages in the field of ultra low power, low voltage op-amp design. Since it is not an issue that has been addressed in detail, in the following well-defined design guidelines to build energy efficient subthreshold op-amps are provided.
As discussed in chapter 2, although MOSFETs stay in subthreshold region provided that V_{GS}<V_{TH}, a minimum V_{DS}=3V_T should be imposed for amplifier transistors in order to reduce the I_D dependence on V_{DS}. Also for subthreshold

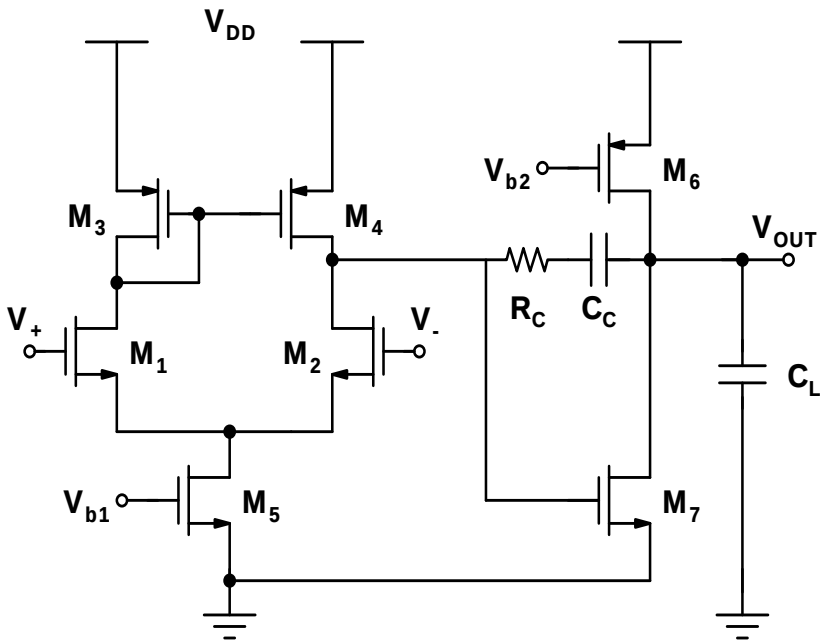


Fig. 4.6 – Two-stage Miller compensated op-amp.

op-amps, two-stage topologies are therefore more suitable than single-stage cascode configurations aiming to reach sufficiently large DC gain and output swing values under very low-voltage supply conditions.

4.3.1. Design guidelines

The transistor-level configuration of the considered op-amp is shown in Fig. 4.6. It represents the well-established two-stage Miller-compensated op-amp in which no particular techniques exploiting the MOSFET body terminal [66-71] (such as bulk-driven or body-bias) are employed. In the circuit, MOSFETs $M_{1\text{-}5}$ and $M_{6,7}$ are assumed to operate in subthreshold region and implement the first and the second amplifier stages, respectively. The capacitance C_C and the resistance R_C constitute the frequency compensation network required to achieve closed-loop stability of the amplifier. The design of the op-amp in the figure is a complex multidimensional problem in which several different constraints have to be taken into account. Aiming to help the designer in addressing this problem, a design procedure was developed to fix the sizes and bias currents of transistors, and the compensation network elements in order to reach the desired op-amp performance. Obviously the procedures assumes that fundamental design specifications regarding the open loop DC gain, the noise performance, the bandwidth, the slew rate and the open loop phase margin are given. In TABLE 4.1 the subthreshold parameters m and I_0 in (2.3) are reported for both PMOS and NMOS of the CMOS process UMC 0.18 μm.

TABLE 4.1

NMOS AND PMOS SUBTHRESHOLD PARAMETERS (W/L = 10 μM/2 μM)

	I_0	m
NMOS	288 nA	1.20
PMOS	74 nA	1.35

4.3.1.1. Open loop DC gain

The procedure starts from the design constraint on the DC gain. Taking into account the expressions of g_m and r_d reported in TABLE 2.1, the DC gain of the two-stage op-amp under examination can be approximately written as follows:

$$A_{V0} = g_{m1} g_{m7} \left(r_{d1} // r_{d3}\right)\left(r_{d6} // r_{d7}\right) = \\ = \frac{1}{m_1 m_7}\left[\left(\frac{\lambda_{D1}}{m_1} + \frac{\lambda_{D3}}{m_3}\right)\left(\frac{\lambda_{D7}}{m_7} + \frac{\lambda_{D6}}{m_6}\right)\right]^{-1} \approx \\ \approx \frac{1}{\left(\lambda_{D1} + \lambda_{D3}\right)\left(\lambda_{D7} + \lambda_{D6}\right)} \tag{4.12}$$

where g_{mi}, r_{di}, m_i and λ_{Di} are the transconductance, drain-source resistance, subthreshold slope parameter and DIBL effect coefficient of the *i*-th transistor, respectively. Above equation relates the DC gain to the λ_D values of transistors $M_{1\text{-}4}$, $M_{6,7}$. Since λ_D values are related to transistor lengths (see Fig. 2.5), eq. (4.12)

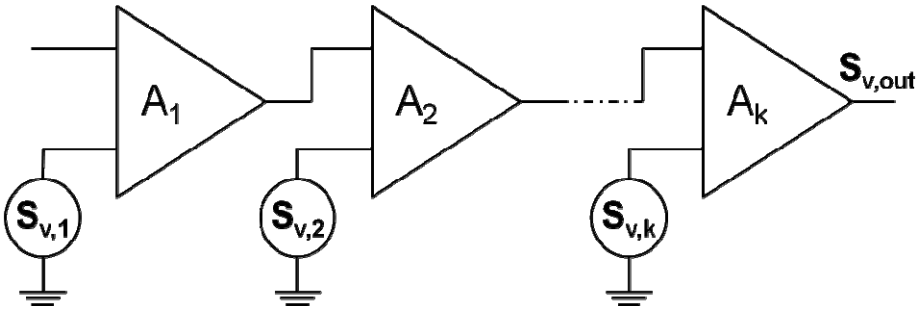


Fig. 4.8 – Noise in a cascade of gain stages.

can be used to estimate the minimum lengths of MOSFETs $M_{1\text{-}4}$, $M_{6,7}$ required to reach the desired DC gain.

4.3.1.2. Noise performance

Noise performance of the two-stage amplifier are related to the noise of single transistors within the circuit. For the cascade of gain stages, the most important contribution comes from the first stage. Indeed, assuming to represent the noise of each stage with an input referred generator, as shown in Fig. 4.8, the output noise power spectral density is,

$$S_{v,out} = \left(A_1 A_2 ... A_k\right)^2 S_{v,1} + \left(A_2 ... A_k\right)^2 S_{v,2} + ... + A_k^2 S_{v,k} \,, \tag{4.13}$$

referred to the input chain, it results,

$$S_{v,in} = S_{v,1} + \frac{S_{v,2}}{A_1^2} + ... + \frac{S_{v,k}}{\left(A_1 A_2 ... A_{k-1}\right)^2} \,. \tag{4.14}$$

Therefore, if the gain of the first stage is large enough the input referred noise power spectral density is mainly controlled by the first stage.
The input referred noise of the differential stage of Fig. 4.3 can be derived as [44],

$$S_{v,in} = 2 S_{v,in_{1,2}} \left[1 + \left(\frac{g_{m3}}{g_{m1}} \right)^2 \right], \tag{4.15}$$

In which, $S_{v,in1,2}$ is the input referred noise power spectral density of the differential pair transistors. Neglecting flicker noise contribution, according to (2.13) in weak inversion and (2.10a), the bias current of the transistors $M_{1,2}$ of the differential pair (namely $I_{D1,2}$) can be calculated as follows, to achieve the required input-referred noise power spectral density $S_{vw,in}$:

$$I_{D1,2} = \frac{4\left(mV_T\right)^2}{qS_{vw,in}} \left[1 + \left(\frac{m_{NMOS}}{m_{PMOS}} \right)^2 \right]. \tag{4.16}$$

From (2.10a), above equation automatically also fixes the transconductance of the first amplifier stage (g_{m1}).

Furthermore, by inverting (2.3), for $V_{DS} \geq 3V_T$,, the aspect ratio of $M_{1,2}$ is then chosen in order to fix a suitable V_{GS} which allows the transistor to operate in the subthreshold region:

$$\left(\frac{W}{L}\right)_{M_{1,2}} = \frac{I_{D1,2}}{I_0} e^{\left(\frac{V_{TH}-V_{GS1,2}}{m_{1,2}V_T}\right)} . \tag{4.17}$$

4.3.1.3. Speed and stability requirements

Once g_{m1} is fixed, the compensation capacitance C_C is set according to the following well-known equation in order to achieve the desired gain-bandwidth product (GBW):

$$C_C = \frac{g_{m1}}{2\pi \cdot GBW} . \tag{4.18}$$

The bias current of $M_{6,7}$ ($I_{D6,7}$) is then set to impose the op-amp slew rate (SR). The SR depends on the slews of both the output node of the differential stage (SR_{INT}) and the output node of the second stage (SR_{EXT}). Therefore the overall op-amp SR is given by the minimum between the internal slew rate and the external slew rate of the second amplifier stage [72]:

$$SR = \min\left(SR_{INT}, SR_{EXT}\right) = \min\left(\frac{2I_{D1,2}}{C_C}, \frac{I_{D6,7}}{C_L + C_C}\right). \tag{4.19}$$

Clearly, if the value of $I_{D1,2}$ chosen from (4.16) to meet the noise requirements is insufficient to achieve also the desired slew rate, according to (4.19), it has to be opportunely increased and the previous design steps have to be retraced. Alternatively, in a first design step, SR_{EXT}=SR_{INT}=SR can be set, obtaining,

$$I_{D6,7} = 2I_{D1,2}\left(1 + \frac{C_L}{C_C}\right). \tag{4.20}$$

From (2.10a), the transconductance of the second amplifier stage (g_{m7}) is automatically fixed once $I_{D6,7}$ is fixed. Moreover, since $I_{D6,7}$ is fixed, it is possible to find the $M_{6,7}$ aspect ratio, similarly to relation (4.17).

In Fig. 4.9 the small signal equivalent circuit of the two-stage amplifier considered is shown. The first stage is represented by the transconductance g_{m1} of the differential pair, the output resistance R_1 is equal to the parallel of the drain

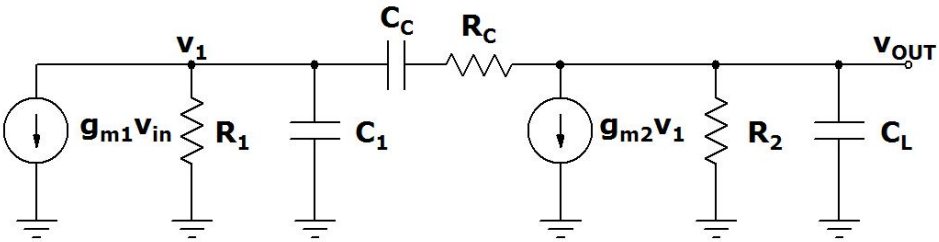


Fig. 4.9 – Small signal equivalent circuit of a two-stage Miller amplifier.

resistances of NMOSs and PMOSs in the first stage and C_1 is the output capacitance of such stage, i.e. the total parasitic capacitance at the output node of the first amplifier stage. g_{m2} and R_2 are related to the common source stage, while C_L is the load capacitance of the op-amp.
Referring to the compensation network, it is known that the compensation capacitor (C_C) alone provides pole-splitting, between the dominant pole and the non-dominant pole, and includes a zero in the right s-plane of the transfer function of the two-stage op-amp. The frequencies of the dominant pole, the second pole and the zero are [44], respectively,

$$f_{fp} \approx \frac{g_{m1}}{2\pi C_C A_{v0}} . \tag{4.21}$$

$$f_{sp} \approx \frac{g_{m2}}{2\pi\left(C_1 C_L / C_C + C_1 + C_L\right)} . \tag{4.22}$$

$$f_z \approx \frac{g_{m2}}{2\pi C_C} . \tag{4.23}$$

A zero in the right s-plane introduces a phase shift like a negative pole in the left s-plane, therefore it should be placed far away from the unity gain frequency of the amplifier. Considering as unity gain frequency the $GBW=g_{m1}/2\pi C_C$, the ratio between GBW and f_z is equal to g_{m1}/g_{m2}. Therefore, the greater g_{m2} to g_{m1}, farther will be the zero from the unity gain frequency. With saturated MOSFETs in strong inversion is difficult to achieve large differences between the two transconductances, gm changes with $\sqrt{I_D}$ and the aspect ratio. Differently, with subthreshold operating MOSFETs, as BJTs, g_m is proportional to I_D and, therefore, g_{m2} can be designed suitably higher than g_{m1}. So, concerning stability issues, in weak inversion regime the frequency compensation by means only of the Miller compensation capacitor C_C requires less effort compared with strong inversion operation. Notwithstanding, in the proposed guidelines the nulling resistor (R_C) approach has been adopted in order to improve the amplifier GBW.
The presence of R_C modifies the zero in,

$$f_z = \frac{1}{2\pi C_C\left(1/g_{m2} - R_C\right)} . \tag{4.24}$$

Standard nulling resistor approach sets $R_C=1/g_{m2}$ in (4.24) sending the zero to frequency ∞, therefore without changing the amplifier bandwidth. In [72] a different technique is proposed, the zero is exploited in the left s-plane to compensate the second pole (4.22). Therefore the compensation sets the following condition,

$$\frac{g_{m2}}{C_L} = \frac{g_{m2}}{C_C\left(g_{m2} R_C - 1\right)} . \tag{4.25}$$

Accomplished (4.25), the frequency of the second pole becomes,

$$f_{sp} = \frac{1}{2\pi R_C C_1}, \tag{4.26}$$

that is higher than (4.22) and, hence, a greater unity gain frequency can be reached. Thus, according to the design strategy developed in [72] to maximize the bandwidth of two-stage op-amps, based on the relation (4.25), the optimal value of R_C can be fixed. The nulling resistor is sized to stabilize the op-amp, by imposing the desired open-loop phase margin (ϕ_M), as follows,

$$R_C = \frac{1}{2g_{m7}}\left(1+\sqrt{1+\frac{4g_{m7}C_L}{g_{m1}C_1\tan(\phi_M)}}\right) \tag{4.27}$$

As shown in the next section, the systematic design procedure introduced above is very useful aiming to achieve very low-voltage low-power op-amps working in subthreshold region.

4.3.2. Two-stage op-amp design

To demonstrate both the actual potential and the validity of the design procedure exposed in the previous paragraph, the two-stage op-amp of Fig. 4.6 was designed and fabricated in a commercial 0.18 µm CMOS technology by using a 0.5 V supply voltage. The transistor dimensions and the values of compensation network elements are shown in TABLE III. According to (4.16) the differential pair was biased with the $I_{D1,2}$ reported in TABLE III, in order to have white noise levels comparable with the best low-power low-voltage solutions (see Table 4.3). Consequently, transistor aspect ratios, C_C and R_C were chosen according to the above mentioned design guidelines.

TABLE 4.2

MOSFET DIMENSIONS AND COMPENSATION NETWORK ELEMENTS

M_5	50 µm / 1 µm = (5 µm / 1 µm) × 10
M_6	100 µm / 10.3 µm = (2.5 µm / 10.3 µm) × 40
$M_{1,2}$	160 µm / 2 µm = (20 µm / 2 µm) × 8
$M_{3,4}$	100 µm / 0.4 µm = (12.5 µm / 0.4 µm) × 8
M_7	5.6 µm / 1 µm
R_C	1.1 MΩ
C_C	2.1 pF
C_L	30 pF
$I_{D1,2}$	17.5 nA
$I_{D6,7}$	110 nA

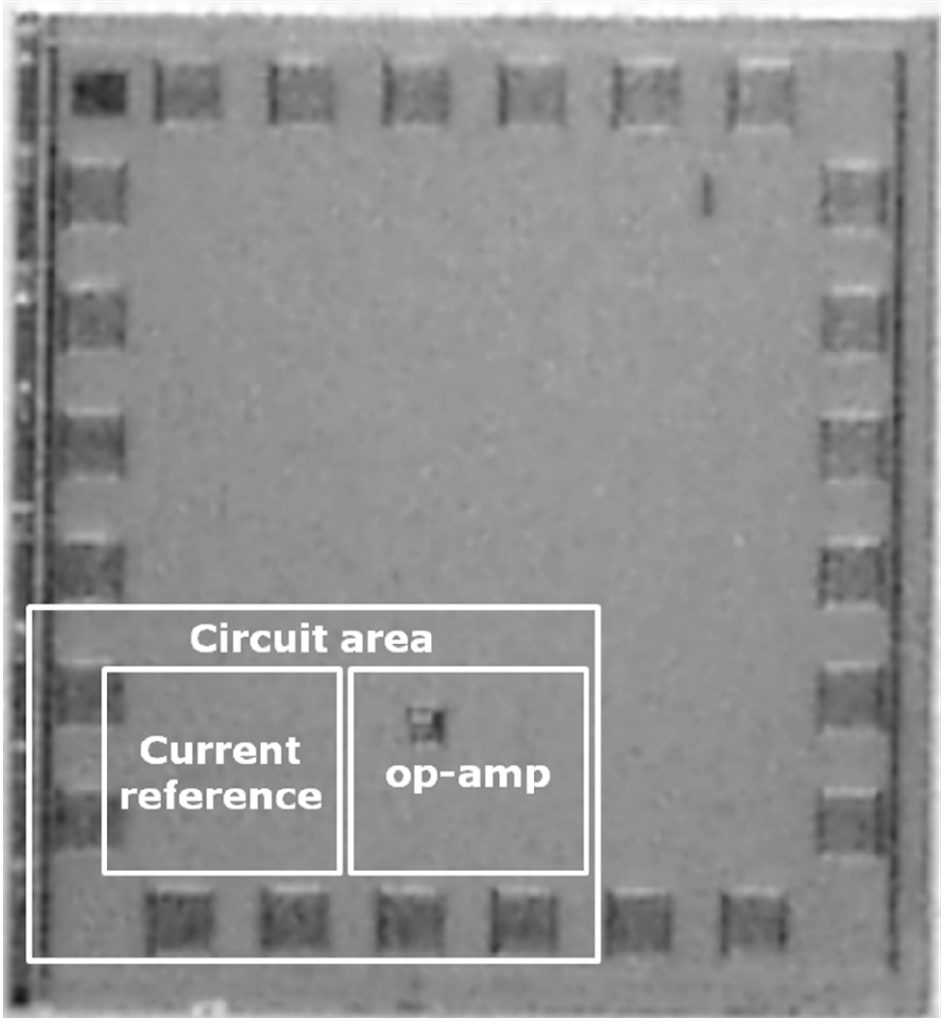


Fig. 4.10 – Chip photo indicating current reference and op-amp sections.

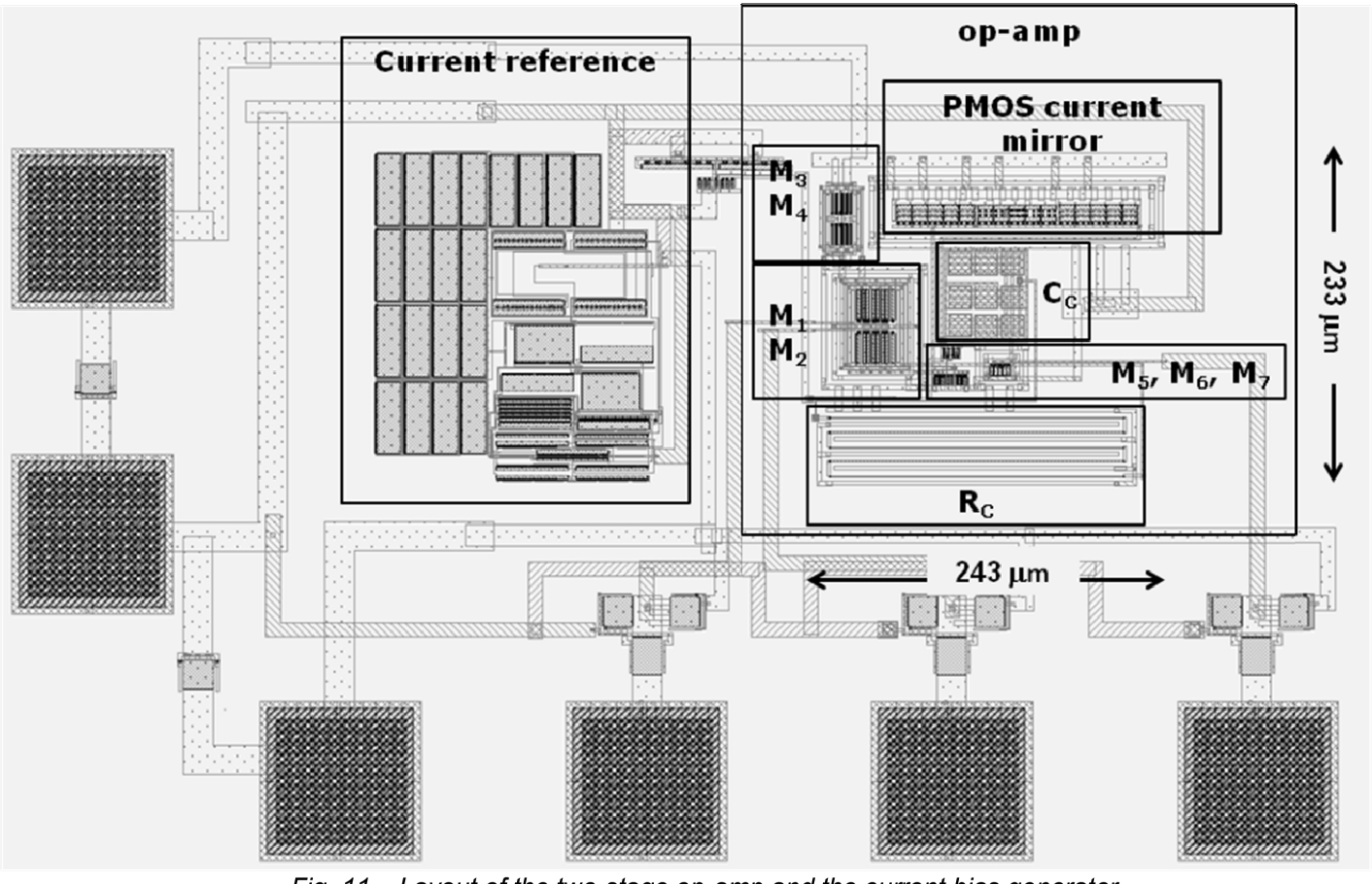


Fig. 11 – Layout of the two-stage op-amp and the current bias generator

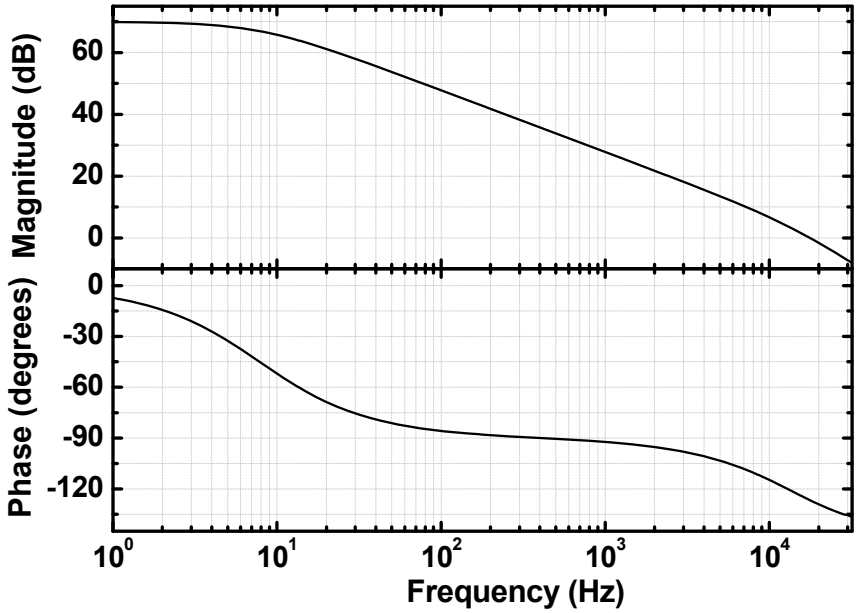


Fig. 4.12 – Measured amplifier frequency response.

The chip photo and the corresponding layout, indicating the circuit subsections, are shown in Fig. 4.10 and 4.11, respectively. Area occupation is just 0.057 mm^2. In the layout is also showed the current reference (similar to that presented in [74]) used to generate the bias currents $I_{D1,2}$ and $I_{D6,7}$ through current mirrors.

4.4. Experimental results

Current consumption and DC gain measurements were performed connecting the packaged samples to the parameter analyzer 4200-SCS Keithley by means of a test fixture. Noise measurements were performed by means of the National Instruments PXI-4472 dynamic signal acquisition module, equipped with anti-aliasing filters. For the remaining measurements the high impedance probe Tektronix P6139A was used.

Fig. 4.12 illustrates the measurement of the open-loop amplifier frequency response which shows DC gain, GBW and phase margin of 70 dB, 18 kHz and 55°, respectively. The measured transient response in unity-gain closed-loop configuration is reported in Fig. 4.13, and it is characterized by a slew rate (SR) of 3 V/ms.

The input-referred voltage noise frequency behavior has been measured by connecting the proposed operational amplifier as a unity gain buffer and amplifying its output through a low noise amplifier (TLC2201) in non inverting configuration, as shown in Fig. 4.14. Moreover, to reduce external noise contributions, the measurement setup has been shielded by a metal box and supplied with ±6 V sealed lead acid batteries. To filter the noise fluctuations coming from the batteries $V_{DDIBIAS}$, $V_{DDOPAMP}$ and V_+ have been filtered by means of a low-pass filter (LPF),

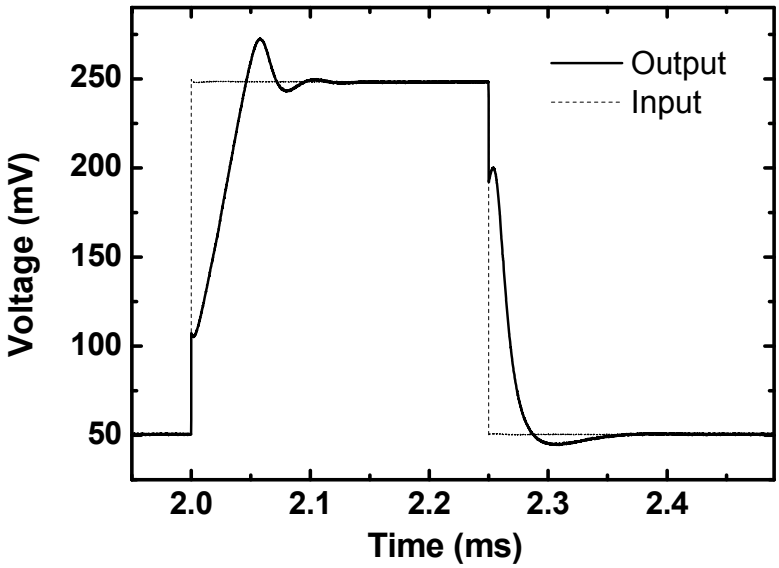


Fig. 4.13 – Measured step response of the amplifier in unity-gain closed-loop configuration.

whose scheme is shown in Fig. 4.15. The potentiometer R^* of each LPF has been adjusted to provide $V_{DDIBIAS}$ =1 V, $V_{DDOPAMP}$ =0.5 V and V_+=250 mV.
The measured input-referred voltage noise is shown in Fig. 4.16. As clear, the flicker noise contribution becomes significant below roughly 100 Hz. This low frequency corner is an additional advantage of subthreshold operation, since 1/*f* noise power is proportional to the square of the DC channel current, while the shot noise is proportional to the DC channel current. Indeed, in [66], where saturated

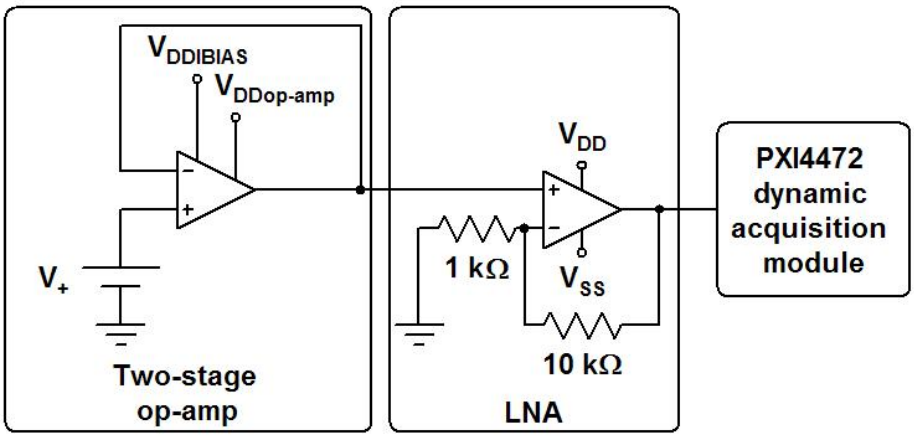


Fig. 4.14 – Noise measurement setup.

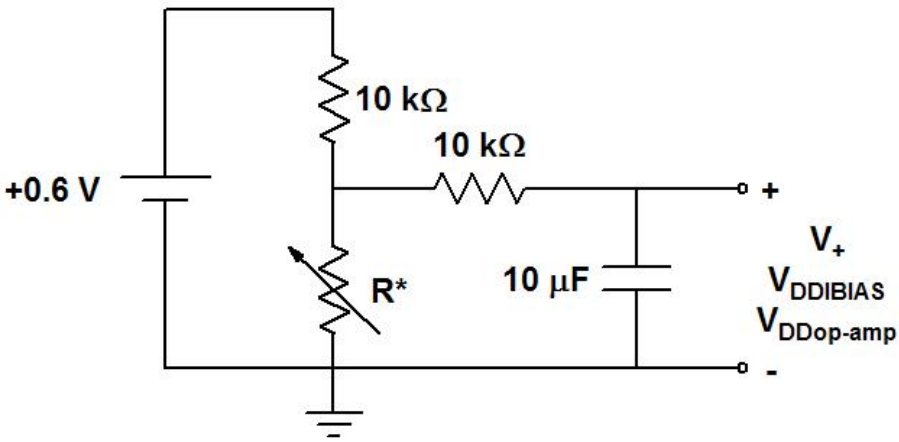


Fig. 4.15 –Low-pass filter for the generation of $V_{DDIBIAS}$, $V_{DDOPAMP}$ and V_{+}.

transistors are involved, the flicker noise affects the noise frequency response up to 1 MHz.

TABLE 4.3 compares the main amplifier performance indicators with results from published low-voltage low-power op-amp designs. The "operation mode" in TABLE 4.3 also highlights the operating region of the MOSFETs and the techniques used to implement the amplifiers (i.e. body-biasing and bulk-driving).

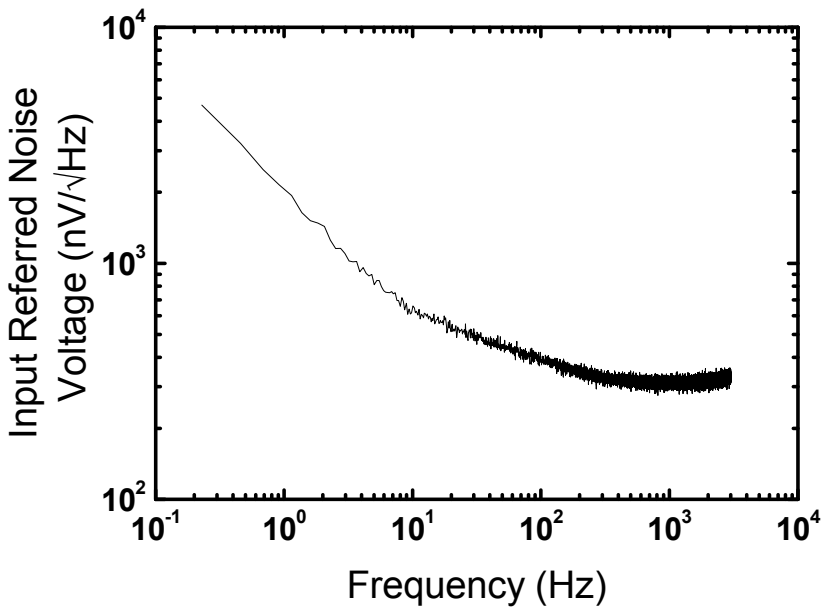


Fig. 4.16 – Measured input referred noise voltage ($nV/\sqrt{Hz}$).

TABLE 4.3

MEASURED OP-AMP PERFORMANCE AND COMPARISON WITH LOW-VOLTAGE LOW-POWER OP-AMPS PRESENTED IN THE LITERATURE

	This work	[68] (gate-input)	[66]	[67]	[70]	[69]
A_{V0} (dB)	70	62	49	69.4	69	79
GBW (MHz)	0.018	10	1.3	0.011	2	0.006
ϕ (°)	55	60	57	65	57	62
SR (V/μs)	0.003	2	1.6	0.015	0.5	____
FOM_{GBW} (MHz•pF/μA)	3.6	1.33	0.09	0.18	1	0.14
FOM_{SR} (V•pF/(μs•μA))	0.68	0.27	0.12	0.25	0.25	____
Input referred noise voltage @ 1 kHz ($nV/\sqrt{Hz}$)	310	____	376	290	____	____
I_{TOT} (μA)	0.15	150	300	0.9	40	0.5
Supply voltage (V)	0.5	0.5	1	0.6	1	0.9
Output swing (V)	0.30	____	0.95	0.50	0.80	0.88
Power (μW)	0.075	75	300	0.54	40	0.45
C_L (pF)	30	20	22	15	20	12
Area (mm^2)	0.057	0.017	1.515	0.060	____	0.500
Operation mode	Subth.	Sat., body-bias	Sat., bulk-driven	Subth., bulk-driven	Subth., bulk-driven	Subth.
Technology (μm)	0.18	0.18	2	0.35	0.5	0.5

The op-amp exhibits the lowest supply voltage (along with [67]) and the lowest power consumption, which is almost one order of magnitude lower than the minimum values found in literature. I_{TOT} shown in TABLE 4.3 refers to both the current injected in the op-amp stages and that flowing in the current mirrors for the $V_{b1,2}$ generation (see Fig. 4.6). Silicon area occupation is comparable to the other low-power solutions. It appears that the proposed solution is characterized by smaller GBW and SR values than some other amplifiers which, however, dissipate a significantly larger amount of power and/or drive smaller load capacitances. Therefore, to compare coherently the dynamic performances of the op-amps under examination, two well known figures of merit (FOM_{GBW} and FOM_{SR}) are used, which also take into account the load capacitance and total current consumption (I_{TOT}) [68], [73]:

$$FOM_{GBW} = \frac{GBW \cdot C_L}{I_{TOT}} \qquad FOM_{SR} = \frac{SR \cdot C_L}{I_{TOT}} \quad . \tag{4.28}$$

Interestingly, the proposed op-amp shows very significant improvements in both FOM_{GBW} and FOM_{SR} (TABLE 4.3) compared with other low-voltage, low-power amplifiers, allowing a more convenient trade-off between speed, power and load capacitance to be reached. Such remarkable efficiency improvement with respect to op-amps working in saturation is mainly owed to the different g_m/I_D behavior in the two operating regions. Finally, the input-referred voltage-noise measured at 1 kHz is only 20 nV/√Hz higher than the best result found at 1 kHz for ultra low-power opamps [67].

Thus subthreshold op-amp design allows a better trade-off among speed, low frequency noise levels and DC bias currents to be achieved compared with strong inversion. Moreover, the comparison with previously reported subthreshold op-amps evidences the effectiveness of the design procedure to fully exploit subthreshold conduction for op-amp design.

CONCLUSION

This work covered several aspects of CMOS circuit design with MOSFETs biased in the subthreshold region. Attention was focused on the design of basic blocks such as voltage references and operational amplifiers that are to be found in nearly all analog sections of integrated circuits. In particular, new circuit configurations and new design methodologies were proposed in order to extremely reduce power consumption and to enable such analog subsystems to be operated at ultra low supply voltages.

At first, the motivations for innovative design approaches in emerging energy-constrained applications were presented. Therefore, subthreshold regime was considered in the design of the analog subsections. The I-V characteristic was deeply examined and small signal equivalent circuit parameters were extracted.

Subsequently, exploiting the relations and considerations discussed above, ultra low voltage, extremely low power voltage reference circuits fabricated in the UMC 0.18 μm CMOS process, were presented. The new circuit configurations work with all transistors in subthreshold region, thus allowing a remarkable reduction of minimum supply voltage and power consumption. Indeed, the proposed solutions represent a significant advance in low-power low-voltage reference circuit design: in one of the three configurations the power dissipation is just 2.6 nW, which is one order of magnitude lower than that of the best results found in literature, and its minimum supply voltage for correct operation falls to 0.45 V. However, power consumption and minimum supply voltage of the other two solutions remain well below 30 nW and 0.7 V, respectively. In addition, a temperature compensation technique in subthreshold-operated voltage reference circuits were presented and employed in the three circuit architectures. Moreover, the accuracy of the circuits were improved against supply voltage variations too, thus obtaining a line sensitivity as low as 0.065 %/V.

Subsequently, a design methodology for ultra low voltage, low power subthreshold op-amp was presented. To demonstrate the potential of subthreshold op-amps, a two-stage Miller-compensated amplifier was fabricated in a commercial 0.18 μm CMOS technology. The designed op-amp is supplied with 0.5 V and dissipates just 75 nW. It achieves a DC open loop gain of 70 dB and a GBW of 18 kHz with a 30 pF load capacitance while the phase margin is 55°. From a comparative analysis with the best low-power, low-voltage solutions found in literature, the proposed one shows the highest values of figures of merit which take into account speed, power consumption and load capacitance (FOM_{GBW} and FOM_{SR}).

The whole study demonstrates that well-designed analog modules operating in subthreshold regime represent very efficient solutions to be included in modern energy-constrained electronics systems, such as in portable biomedical applications or wireless sensor nodes.

REFERENCES

1. N. Mohamed and I. Jawhar, "A fault tolerant wired/wireless sensor network architecture for monitoring pipeline infrastructures," in *Proc. 2nd Int. Conf. Sensor Technol. Appl.*, Aug. 2008, pp. 179–184.
2. Y. Tachwali, H. Refai, and J. E. Fagan, "Minimizing HVAC energy consumption using a wireless sensor network," in *Proc. 33rd Annu. Conf. IEEE Ind. Electron. Soc.*, Nov. 2007, pp. 439–444.
3. N. G. Elvin, N. Lajnef, and A. A. Elvin, "Feasibility of structural monitoring with vibration powered sensors," *Smart Mater. Structures*, vol. 15, pp. 977–986, Jun. 2006.
4. D. Malan, T. Fulford-Jones, M. Welsh, and S. Moulton, "CodeBlue: An ad hoc sensor network infrastructure for emergency medical care," in *Proc. Int. Workshop Wearable Implantable Body Sensor Netw.*, Jun. 2004.
5. G. Chen, S. Hanson, D. Blaauw, and D. Sylvester, "Circuit Design Advances for Wireless Sensing Applications," *Proc. of the IEEE*, Vol. 98, No. 11, Nov. 2010.
6. A. Wang, B. H. Calhoun, and A. P. Chandrakasan, **Sub-threshold Design for Ultra Low-Power Systems**. Springer Science+Business Media, LLC, 2006.
7. B. H. Calhoun, and A. P. Chandrakasan, "Ultra-Dynamic Voltage Scaling (UDVS) Using Sub-Threshold Operation and Local Voltage Dithering," *IEEE J. of Solid-State Circuits*, Vol. 41, No. 1, Jan. 2006.
8. "2010 Overall Roadmap Technology Characteristics (ORTC) Tables" *ITRS*, 2009, (Online), http://public.itrs.net.
9. L. Magnelli, F. A. Amoroso, F. Crupi, G. Cappuccino and, G. Iannaccone, "Design of a 75 nW, 0.5 V Subthreshold CMOS Operational Amplifier," *Int. Jour. Of Circ. Theory and Appl.*, submitted on Jul. 01st, 2011.
10. U. Karthaus and M Fischer, "Fully Integrated Passive UHF RFID Transponder IC With 16.7-μW Minimum RF Input Power," *IEEE J. of Solid-State Circuits*, Vol. 38, No. 10, Oct. 2003.
11. G. De Vita and G. Iannaccone, "Ultra low power RF section of a passive microwave RFID transponder in 0.35 μm BiCMOS," *IEEE International Symposium on Circuits and Systems*, 2005 – *ISCAS* 2005, vol. 5, pp. 5075-5078, Kyoto 2005.
12. "Superstructures Engineering: Adding sensors and other devices to bridges, tunnels and buildings can turn them into "smart structures" capable of sensing and, in some cases, even responding to problems," *The Economist online*, Dec. 9th, 2010.
13. G. Chen, S. Hanson, D. Blaauw, and D. Sylvester, "Circuit Design Advances for Wireless Sensing Applications," *Proc. of the IEEE*, Vol. 98, No. 11, pp. 1808-1827, Nov. 2010.

14. N. Verma, A. Shoeb, J. V. Guttag, and A. P. Chandrakasan, “A micro-power EEG acquisition SoC with integrated seizure detection processor for continuous patient monitoring,” in *Proc. Symp. Very Large Scale Integr. (VLSI) Circuits*, Jun. 2009, pp. 62–63.
15. M. van Elzakker, E. van Tuijl, P. Geraedts, D. Schinkel, E. A. M. Klumperink, and B. Nauta, “A 10-bit charge-redistribution ADC consuming 1.9 _w at 1 MS/s,” *IEEE J. Solid-State Circuits*, vol. 45, no. 5, pp. 1007–1015, May 2010.
16. A. Agnes, E. Bonizzoni, P. Malcovati, and F. Maloberti, “A 9.4-ENOB 1 V 3.8 _W 100 kS/s SAR ADC with time-domain comparator,” in *Proc. IEEE Int. Solid-State Circuits Conf.*, 2008, pp. 246–610.
17. F. Lee and A. Chandrakasan, “A 2.5 nJ/bit 0.65 V pulsed UWB receiver in 90 nm CMOS,” *IEEE J. Solid-State Circuits*, vol. 42, no. 12, pp. 2851–2859, Dec. 2007.
18. G. De Vita, G. Iannaccone, “A Sub-1-V, 10 ppm/ oC, Nanopower Voltage Reference *Generator”, IEEE J. Solid-State Circuits*, Vol.42, n. 7, pp. 1536-1542, 2007.
19. 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.
20. J. Ko, C. Lu, M. B. Srivastava, J. A. Stankovic, Andreas Terzis, and Matt Welsh, “Wireless Sensor Networks for Healthcare,” *Proc. of the IEEE*, Vol. 98, No. 11, pp. 19471960, Nov. 2010.
21. R. P. Dresher and P. P. Irazoqui, “A compact nanopower low output impedance CMOS operational amplifier for wireless intraocular pressure recordings,” in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 6055–6058.
22. S. Raghunathan, M. P. Ward, K. Roy, and P. P. Irazoqui, “A low-power implantable event-based seizure detection algorithm,” in Proc*. 4th Int. IEEE/EMBS Conf. Neural Eng.*, May 2, 2009, pp. 151–154.
23. M. D. Linderman, G. Santhanam, C. T. Kemere, V. Gilja, S. O’Driscoll, B. M. Yu, A. Afshar, S. I. Ryu, K. V. Shenoy, and T. H. Meng, “Signal processing challenges for neural prostheses,” *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 18–28, Dec. 2008.
24. H. Neuteboom, Ben M. J. Kup, and M. Janssens, “A DSP-based hearing instrument IC,” *IEEE J. Solid-State Circuits*, Vol. 32, No. 11, pp. 1790 – 1806, Nov. 1997.
25. Intel Corporation. (2006). “The SHIMMER sensor node platform,” [Online]. Available: http://fiji.eecs.harvard.edu/CodeBlue
26. Jan M. Rabaey, A. P. Chandrakasan, and B. Nicolić , **Digital Integrated Circuits – A Design Perspective (2nd Ed.)**. Prentice Hall, 2003.
27. K.-S. Yeo and K. Roy, **Low-Voltage, Low-Power VLSI Subsystems**. Mc Graw-Hill, 2005
28. J.C. Shah, S.S. Sapatnekar, “Wiresizing with buffer placement and sizing for power-delay tradeoffs,” *Proc. of VLSI Design*, pp. 346-351, 1996.
29. B. H. Calhoun, F. A. Honoré, 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.

30. J. W. Tschanz, J. T. Kao, S. G. Narendra, R. Nair, D. A. Antoniadis, A. P. Chandrakasan, and V. De, “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.
31. F. Crupi, L. Magnelli, P. Falbo, M. Lanuzza, M. Nafría, R. Rodríguez, “Performance and reliability of ultra-thin oxide nMOSFETs under variable body bias,” *Microelectronic Engineering,* vol. 84, pp. 1947-1950, 2007.
32. Calhoun, B. H., A. Wang, A. Chandrakasan, "Modeling and Sizing for Minimum Energy Operation in Subthreshold Circuits," *IEEE J. Solid-State Circuits*, vol. 40, no. 9, pp. 1778-1786, September 2005.
33. B. C. Paul, A. Raychowdhury, and K. Roy, “Device optimization for digital subthreshold logic operation,” *IEEE Trans. Electron Devices*, vol. 52, no. 2, pp. 237–247, Feb. 2005.
34. F. M. Wanlass and C. T. Sah, "Nanowatt Logic Using Field-Effect Metal-Oxide-Semiconductor TYiodes," in *IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers,* Philadelphia, vol. 6, pp. 32-33, Feb. 1963.
35. R. M. Swanson and J. D. Meindl, "Ion-Implanted Complementary MOS Transistors in Low-Voltage Circuits," *IEEE J. Solid-State Circuits*, vol. 7, no. 2, pp. 146-153, April. 1972.
36. E. Vittoz and J. Fellrath, "CMOS Analog Integrated Circuits Based on Weak Inversion Operation ," *IEEE J. Solid-State Circuits*, vol. 12, no. 3, pp. 224-231, June 1977.
37. H.J. Oguey and D. Aebischer, “CMOS current reference without resistance,” *IEEE J. Solid-State Circuits*, vol.32, pp.1132–1135, July 1997.
38. R. Lyon and C Mead, "An analog electronic cochlea," in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 7, pp. 1119-1134, July 1988.
39. H. Soeleman, K. Roy, and B. C. Paul, “Robust subthreshold logic for ultra-low power operation,” *IEEE Trans. VLSI Syst*, Vol. 9, Issue 1, Feb. 2001.
40. E. A. Vittoz, "Weak Inversion for Ultimate Low-Power Logic," in *Low-Power Electronics Design*, C. Piguet, Ed. CRC Press, 2005.
41. C. C . Enz and E. A. Vittoz, “CMOS Low-Power Analog Circuits Design,” in *Proc. Int. Symp. Circ. and Syst. (ISCAS)*, pp. 79-133, 1993.
42. Yuan Taur, Tak H. Ning, **Fundamentals of Modern VLSI Devices**, Cambridge University Press, 2002.
43. Y. Tsividis, **Operation and Modeling of the Transistor MOS (2nd ed.)**, Oxford University Press, 2003.
44. F. Maloberti, **Analog Design for CMOS VLSI Systems**. Boston, MA: Kluwer Academic, 2001.
45. M. Alioto, “Understanding DC Behavior of Subthreshold CMOS Logic Through Closed-Form Analysis”, *IEEE Trans. on Circ. and Syst. I*, vol. 57, no. 7, pp. 1597-1607, July 2010.
46. W. Liu, X. Jin, J. Chen, M. Jeng, Z. Liu, Y. Cheng, K. Chen, M. Chan, K. Hui, J. Huang, R. Tu, P. K. Ko and C. Hu, *BSIM3v3.2.2 MOSFET Model, Users’Manual*. Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, 1999.

47. 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.
48. K. N. Leung and P. K. T. Mok, “A sub-1-V 15 ppm/°C CMOS bandgap voltage reference without requiring low threshold voltage device,” *IEEE J. Solid-State Circuits*, vol. 37, no. 4, pp. 526–530, Apr. 2002.
49. J. Doyle, Young Jun Lee, Yong-Bin Kim, H. Wilsch, and F. Lombardi, “A CMOS Subbandgap Reference Circuit With 1-V Power Supply Voltage,” I*EEE J. Solid-State Circuits*, vol. 39, no. 1, pp. 252–255, Jan. 2004.
50. B. K. Ahuja, Hoa Vu, C. A. Laber, W. H. Owen, "A very high precision 500-nA CMOS floating-gate analog voltage reference," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2364- 2372, Dec. 2005.
51. P. -H. Huang, H. Lin, and Y.-T. Lin, “A simple subthreshold CMOS voltage reference circuit with channel-length modulation compensation,” *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, no. 9, pp. 882-885, Sept. 2006.
52. G. Giustolisi, G. Palumbo, M. Criscione, and F. Cutri, “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.
53. I. Akita, K. Wada, and Y. Tadokoro, “A 0.6-V Dynamic Biasing Filter With 89-dB Dynamic Range in 0.18-μm CMOS,” *IEEE J. Solid-State Circuits*, vol. 44, no. 10, pp. 2790–2799, Oct. 2009.
54. K. N. Leung and P. K. T. Mok, “A CMOS Voltage Reference Based on Weighted ΔV_{GS} for CMOS Low-Dropout Linear Regulators,” *IEEE J. Solid-State Circuits*, vol. 37, no. 1, pp. 146–150, Jan. 2003.
55. 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.
56. 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.
57. H. Wang and Q. Ye, “A CMOS Voltage Reference Without Resistors for Ultra-Low Power Applications,” *ASIC, 2007. ASICON '07. 7th International Conference on*, pp. 526-529, 22-25 Oct. 2007.
58. M. Oljaca and B. Baker, “How the voltage reference affects ADC performances,” *Analog Application Journal (http://www.ti.com/aaj)*, Texas Instruments Incorporated, 2009.
59. Willy M. C. Sansen, **Analog Design Essential**. Springer-Verlag, May 2006.
60. E. M. Camacho-Galeano, C. Galup-Montoro, and M. C. Schneider, “A 2-nW 1.1-V self-Biased Current Reference in CMOS Technology,” *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 52, no. 2, pp. 61-65, Feb. 2005.
61. B. Razavi, **Design of Analog CMOS Integrated Circuits**. New York: McGraw Hill, 2001.
62. A. Boni, “,” *IEEE J. Solid-State Circuits*, vol. 37, no. 10, pp. 1339-1343, Oct. 2002.
63. Q. A. Khan, S. K. Wadhwa, and K. Misri, “Low Power Startup Circuits for Voltage and Current Reference with Zero Steady State Current*,” in Proc. of Int. Symp. Low Pow. Elect. and Design (ISLPED)*, 2003.

64. A. Hastings, **The Art of Analog Layout. Englewood Cliffs.** NJ: Prentice-Hall, 2001.
65. V. Gromov, A. J. Annema, R. Kluit, J. L. Visschers, P. Timmer, "A Radiation Hard Bandgap Reference Circuit in a Standard 0.13 µm CMOS Technology," *Nuclear Science, IEEE Transactions on* , vol.54, no.6, pp.2727-2733, Dec. 2007.
66. B. J. Blalock, P. E. Allen, and G. A. Rincon-Mora, "Designing 1-V Op Amps Using Standard Digital CMOS Technology", *IEEE Trans. On Circ. And Syst. II*, vol. 45, no. 7, pp. 769-779, Jul. 1998.
67. L. H. C. Ferreira, T. C. Pimenta, R. L. Moreno, "An Ultra-Low-Voltage Ultra-Low-Power CMOS Miller OTA With Rail-to-Rail Input/Output Swing", *IEEE Trans. On Circ. And Syst. II*, vol. 45, no. 10, pp. 843-847, Oct. 2007.
68. S. Chatterjee, Y. Tsvidis, and P. Kinget, "0.5-V Analog Circuit Techniques and Their Application in OTA and Filter Design", *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2373-2387, Dec. 2005.
69. T. Stockstad and H. Yoshizawa, "A 0.9-V 0.5-µA Rail-to-Rail CMOS Operational Amplifier", *IEEE J. Solid-State Circuits*, vol. 37, no. 3, pp. 286-292, Mar. 2002.
70. T. Lehmann and M. Cassia, "1-V Power Supply CMOS Cascode Amplifier", *IEEE J. Solid-State Circuits*, vol. 36, no. 7, pp. 1082-1086, Jul. 2001.
71. C. Sawigun and W.A. Serdijn, "Ultra-low-power, class-AB, CMOS four-quadrant current multiplier", *IET Electron. Lett.*, vol. 45, no. 10, May 2009.
72. G. Palmisano, G. Palumbo and S. Pennisi, "Design Procedure for Two-Stage CMOS Transconductance Operational Amplifiers: A Tutorial", *Analog Int. Cir. Sig. Proc.*, vol. 27, pp. 179–189, 2001.
73. Hoi Lee, Mok, P.K.T, "A CMOS current-mirror amplifier with compact slew rate enhancement circuit for large capacitive load applications," *Int. Symp. Circ. and Syst., ISCAS 2001* , vol. 1, pp.220-223, 6-9 May 2001.
74. G. De Vita, and G. Iannaccone, "A 109 nW, 44 ppm/°C CMOS Current Reference with Low Sensitivity to Process Variations," *IEEE International Symposium on Circuits and Systems, 2007 (ISCAS 2007)*, pp. 3804-3807, New Orleans, USA, 27-30 May 2007.
75. L. Magnelli, F. Crupi, P. Corsonello, and G. Iannaccone, "A sub-1V nanopower temperature compensated subthreshold CMOS voltage reference with 0.065 %/V line sensitivity", *IET Electron. Lett.*, submitted on Jan. 28th, 2012.