# UNIVERSITA' DI PISA <br> Dipartimento di Ingegneria dell'Informazione: <br> Elettronica, Informatica, Telecomunicazioni 

Dottorato di Ricerca in Ingegneria dell'Informazione

# Nanopower CMOS transponders for UHF and microwave RFID systems 

## Giuseppe De Vita

Tutore: Prof. Giuseppe Iannaccone

Tutore: Prof. Bruno Pellegrini

1. Introduction ..... 1
1.1. Low-Power electronics ..... 1
1.1.1. Low-power integrated circuit design techniques ..... 2
Low-power in digital circuits ..... 2
Low-power in analog circuits ..... 4
1.2. Low-power electronics applications ..... 7
1.2.1. Wireless sensor networks ..... 7
1.2.2. Ambient Intelligence ..... 9
1.2.3. Implantable medical devices ..... 10
1.2.4. Radio Frequency IDentification (RFID) Systems ..... 10
1.3. RFID technology ..... 11
1.3.1. Brief history of RFID technology ..... 11
1.3.2. Classification of RFID systems ..... 12
1.3.3. International standards ..... 14
1.3.4. Passive UHF/Microwave RFID systems ..... 15
1.4. References ..... 17
2. Design Criteria and Architecture ..... 21
2.1. Introduction ..... 21
2.2. Voltage Multiplier and Power Matching Network ..... 23
2.2.1. $\quad \mathrm{N}$-stage Voltage Multiplier ..... 23
Power Consumption ..... 26
Power consumption in the presence of substrate losses ..... 28
2.2.2. Input Equivalent Impedance ..... 30
2.2.3. Power Matching Network ..... 30
2.2.4. Non-Linear Effects ..... 34
2.2.5. Matching when Conditions Vary ..... 37
2.3. BACKSCATTER MODULATOR ..... 40
2.3.1. ASK and PSK Backscatter Modulation ..... 40
2.3.2. PSK Backscatter Modulator ..... 44
Circuit description ..... 45
Comparison with other topologies ..... 47
2.4. Modulation Depth and Maximum Operating Range ..... 49
2.4.1. Transponder input power ..... 49
2.4.2. Probability of Error at the Reader ..... 50
Received Signal at the Reader's Antenna ..... 50
Receiver Architecture ..... 51
Noise Spectral Density ..... 53
2.5. References ..... 58
3. Voltage Reference ..... 60
3.1. Introduction ..... 60
3.2. Overview of CMOS-Based Voltage Reference ..... 61
3.3. Proposed Voltage Reference 1 ..... 63
3.3.1. $\quad$ Operating Principle of the Proposed Reference Voltage Generator63
3.3.2. Circuit Description ..... 65
Current Generator Circuit ..... 65
Active Load ..... 67
3.3.3. Supply Voltage Dynamic Range ..... 68
3.3.4. Temperature Compensation ..... 69
3.3.5. Line Sensitivity ..... 71
3.3.6. Experimental Results ..... 71
3.4. Proposed Voltage Reference 2 ..... 73
3.4.1. Circuit Description ..... 73
3.4.2. Temperature Compensation ..... 75
3.4.3. Second order effects on the temperature coefficient ..... 76
Channel length modulation effect ..... 76
3.4.4. Body effect ..... 77
3.4.5. Power Supply Rejection Ratio ..... 78
3.4.6. Experimental Results ..... 81
3.5. Proposed Voltage Reference 3 ..... 83
3.5.1. Circuit Description ..... 83
Current generator circuit ..... 84
Active load ..... 85
3.5.2. Design Consideration ..... 86
Channel length modulation effect ..... 86
Minimum power consumption dimensioning ..... 86
Sensitivity to process variations ..... 87
3.5.3. Dynamic Range ..... 87
3.5.4. Temperature Compensation ..... 88
3.5.5. Second order effects on the temperature coefficient ..... 88
Channel length modulation effect ..... 88
Body effect ..... 90
3.5.6. Experimental Results ..... 90
3.6. Conclusion ..... 94
3.7. References ..... 95
4. Power Supply ..... 97
4.1. Introduction ..... 97
4.2. Architecture of a Passive RFID Transponder ..... 98
4.3. Temperature compensated voltage regulator ..... 100
4.3.1. Circuit description ..... 100
4.3.2. Temperature Coefficient ..... 102
4.3.3. Experimental Results ..... 104
4.4. Negative-Temperature coefficient voltage regulator ..... 104
4.4.1. $\quad$ Series Voltage Regulator ..... 107
4.4.2. Supply Voltage Range ..... 109
4.4.3. Temperature Coefficient ..... 110
4.4.4. Experimental Results ..... 112
4.5. Constant-delay voltage regulator ..... 116
4.5.1. $\quad$ Current Reference Circuit ..... 116
4.5.2. Voltage Reference Generator ..... 118
4.5.3. Sensitivity to Temperature Variations ..... 119
4.5.4. Sensitivity to Process Variations ..... 120
4.5.5. Experimental Results ..... 121
4.6. Conclusion ..... 125
4.7. References ..... 125
5. Non-Volatile Memory ..... 127
5.1. Introduction ..... 127
5.2. Memory Costraints ..... 129
5.3. Memory Cell ..... 129
5.4. Memory Architecture ..... 132
5.5. Memory Circuits ..... 132
5.5.1. Sense Amplifier. ..... 132
5.5.2. Word Line Driver. ..... 134
5.5.3. Input and Output buffer and Y decoder ..... 137
5.6. Experimental Results ..... 137
5.7. Conclusion ..... 143
5.8. References ..... 143
6. Complete implementation of a passive transponder ..... 145
6.1. Transponder architecture ..... 145
6.2. Analog section ..... 146
6.2.1. Voltage multiplier ..... 147
6.2.2. Demodulator ..... 149
6.2.3. ASK backscatter modulator ..... 150
6.3. Digital section ..... 152
6.3.1. Random delay ..... 153
6.3.2. Acknowledgement detector ..... 153
6.3.3 Clock generator ..... 154
Circuit Description ..... 155
Current Reference Generator ..... 156
Voltage Reference ..... 157
Temperature Coefficient ..... 158
Process Variations Sensitivity ..... 158
Simulation Results ..... 159
6.4. Conclusion ..... 162
6.5. System performance ..... 163
6.6. References ..... 163
7. Conclusion ..... 165

## 1. INTRODUCTION

### 1.1. Low-Power electronics

Low-voltage and low-power integrated circuits design techniques were originally developed, more than 30 years ago, for wrist watches. The power consumption of quartz-crystal wrist watches must be smaller than few $\mu \mathrm{W}$ to ensure a sufficient duration of the available energy source. In quartz-crystal wristwatches it is preferable to work at high input frequency so that quartz crystals are cheaper and have a better temperature coefficient. In such condition, we thus need a frequency divider to generate the 1 Hz -clock signal required for the correct operation of the watch. As a consequence, it is very important to reduce as much as possible the dynamic power of the first dividing stages, which work at higher frequency. In the literature it is possible to find different works where some solutions for low-power frequency dividers are proposed [1], [2], [3]. In the design proposed in [3], at a supply voltage of 1.35 V the maximum frequency is 2 MHz and the dynamic power consumption per stage is $1.6 \mathrm{nW} / \mathrm{kHz}$, which is one order of magnitude smaller than that of the designs proposed in [1], [2]. These are the first examples of lowpower and low-voltage integrated circuits. Today, wrist watches have a complexity larger than some thousands of transistors with an on-board microcontroller and they have a power consumption smaller than $0.5 \mu \mathrm{~W}$ at supply voltages smaller than 1.5 V .

In recent years, in virtue of the widespread diffusion of battery-operated devices, especially in the field of medical applications and short-range low frequency communication, there is a growing demand for low-power circuits. In such systems the low-power consumption is the first requirement, whereas other requirements, such as speed or dynamic range, are sacrificed. In battery-operated systems, the demand for low-power circuits is driven by the need to extend the battery life time, in order to reduce the replacing or recharging procedures, which often are very costly, as in the case of sensors or identification devices distributed in a wide area, or difficult to perform, as in the case of medical implantable devices. In such applications, there is also a stringent requirement for small size and weight and this imposes to use small-size batteries, which thus are able to provide quite small energy capacity (few Wh). In many applications scenario, the targeted node lifetime ranges is typically between 2 to 10 years, which sets a drastic requirement on the power consumption. Indeed, in the case of an on-board 1.5 V -AA-battery of 2.6 Ah with a leakage current of $30 \mu \mathrm{~A}$, in order to achieve a lifetime lying between 2 to 7 years, an average power consumption comprised between 10 to $100 \mu \mathrm{~W}$ is required. Since the power consumption of radio transceivers commercially available today
ranges typically in the several tens of mW , for example a Bluetooth transceiver consumes some tens of mW [4], [5], it is clear that there is the need to develop adhoc circuits with very low power consumption.

Moreover, in recent years, in most of VLSI-based systems, including computers and telecommunication products, the urgent need of portability and the growing relative cost of power supplies and heat-removal systems are leading to a strong demand for low-power circuits [6], [7], whereas, until few years ago, the power consumption was the least important requirement and the largest interest was in achieving higher and higher speed and precision. High power consumption in modern VLSI systems causes self-heating, which is a big problem because, when the chip temperature increases, there is a strong reduction of the device reliability and a degradation of the system performance. As a consequence, in order to ensure a safe and high-performance operation, cheap plastic packages can not be used anymore but it is necessary to use air-cooled packages or other kinds of packages, which are much more expensive. Moreover, in order to satisfy the increasing need of power in high performance portable systems, such as laptop and mobile phones, more and more expensive batteries are developed, to provide the power required with a life time long enough. Their cost is, by now, a large fraction of the total cost of the product in which they are installed.. As a consequence, today, the low-power trend in circuit design is mainly driven by two forces: on the one hand, the demand for long-life battery-operated systems and, on the other hand, the technological limitation of high performance VLSI systems. In the last years, the strong demand for low-power consumption has led to the development and use of several techniques, both for digital and analog circuits.

### 1.1.1. Low-power integrated circuit design techniques

As far as low-power design techniques are concerned, it is usually convenient to distinguish between digital and analog circuits.

## Low-power in digital circuits

The power consumption in VLSI digital circuit can be written as,

$$
\begin{equation*}
P=1 / 2 C V_{D D} \Delta V f+I_{s t} V_{D D}+I_{\text {static }} V_{D D}, \tag{1-1}
\end{equation*}
$$

where $V_{D D}$ is the supply voltage, $C$ is the load capacitance, $\Delta V$ is the load voltage swing, $f$ is the load switching frequency, $I_{s t}$ is the shortcircuit current and $I_{\text {static }}$ is the static current due to junction currents, subthreshold currents and gate tunneling currents. The first term is the load switching power, the second term is the shoot through power and the third term is the static power. Several techniques have been developed to reduce each of the three terms.

The load switching power can be drastically reduced through minimization of capacitance, voltage and frequency. The capacitance minimization can be achieved by a power/performance sizing, clock-gating strategies and glitch suppression
procedures. It is clear that the largest capacitances in the chip are the output capacitances due to pads and package and then most of the switching power is consumed to charge/discharge such capacitances. Usually, in order to drive such capacitances, a sizing, based on the gate size multiplier in an exponential horn of inverters, is used. Unfortunately, such strategy is optimized to maximize the performance in term of speed. Device sizing for power efficiency is significantly different than sizing for performance and then a proper strategy, optimized for power efficiency, must be used achieving a good trade-off between power and performance [8]. A final solution to minimize the output capacitances, which are power hungry, is to use System on Chip (SoC): the integration in a unique chip totally eliminates such capacitances. In VLSI systems, from $25 \%$ to $50 \%$ of power consumption is usually due to driving latches, most of which have a low utilization $(10-35 \%)$. In order to reduce the power consumption, a strategy is to gate-off unused latches and associated logic, turning off clocks to unused units or to individual latch banks [9]. Glitches can represent a significant portion of the dynamic power and they can be avoided by using non-glitching logic, such as domino, and by a careful timing dimensioning in order to adjust the delays in a proper way [10].

The voltage minimization can be achieved by lowering the voltage swing and the supply voltage. The most efficient strategy, in this sense, is to lower the supply to lower both $V_{D D}$ and $\Delta V$. The supply voltage reduction is the most promising strategy to drastically reduce the power consumption in modern VLSI systems. Indeed, the maximum frequency is proportional to the supply voltage ( $f \propto V_{D D}$ ) but the power is proportional to $V_{D D}^{3}$ and then, when reducing the supply voltage, the power consumption is drastically reduced while the performance is still acceptable.

Lowering the frequency allows us to reduce the power linearly but does not improve the energy efficiency. Anyway, such solution is important to avoid heating problems.

The shoot through power can represent a not negligible portion of the dynamic power consumption ( $8-15 \%$ ). It can be minimized by lowering the supply voltage or by avoiding slow input signal, in order to minimize the time interval during which both the pull-up and pull-down network of the driven gate are enabled simultaneously.

The static power consumption can be minimized by lowering the supply voltage or by lowering the static currents. In this sense, NMOS, pseudo-NMOS or CMOS CML logics must be avoided, except for very specialized applications. In modern sub-micron IC technologies the gate tunneling current, because of the low oxide thickness, is the most important component of the static current in VLSI systems and is the major power issue especially for the standby power. Since tunnel current is exponentially dependent on the electric field in the oxide [11], a reduction technique consists of reducing the supply voltage, in order to reduce the voltage
across the oxide, or of using new gate materials with higher dielectric constant, in order to increase the equivalent oxide thickness [12]. Both solutions lead to a strong reduction of the tunneling current through the oxide. As a consequence, lowering the supply voltage has a double effect in reducing the stand-by power because it reduces both $V_{D D}$ and $I_{\text {static. }}$. The subthreshold current, instead, is quite significant in fast low-threshold voltage devices. An approach is to use low-threshold voltage devices only in the critical paths, to guarantee high performance, and to use highthreshold voltage devices in the rest of the system, to drastically reduce the subthreshold current [13], [14]. Another possibility to reduce subthreshold current is to use stacked devices [15]. Especially for energy constrained systems, such as battery operated systems, the stand-by power can be drastically reduced by two levels of supply gating: to lower or to turn off power supply to the whole system when inactive and to turn off inactive units while system is active.

In conclusion, some trade-offs must be faced. In order to achieve high performance, we need high supply voltage and low threshold voltage to minimize the propagation delay; in order to minimize dynamic power we need low supply voltage and low threshold voltage; in order to minimize stand-by power we need low supply voltage and high threshold voltage. A trade-off must be found according to the specific application we are interested to.

Power reduction must be also performed at the architectural level through the development of ad-hoc domain-specific computing architectures because general purpose microprocessors can dissipate 500 times more power than an ad-hoc hardware realization [16]. Since most of the power is dissipated for data transfers to the memory, cache and local registers, different methods have been proposed for a better utilization of memory hierarchy and processor cycles [17] and to design optimal memory architectures for low power [18]. Such techniques consist of proper transformations of the code to reduce memory accesses to external memory, to improve data locality in the on-chip cache and to optimize the storage order to reduce address computation. Such transformations drastically reduce the total number of accesses to memory leading to a strong power reduction. Another way to reduce the power consumption is to lower as much as possible the clock frequency by exploiting parallelism. Indeed, in an Ambient Intelligence platform there are many concurrent tasks with different sampling rate and then the best choice is to use a multi-processor architecture with its own localized memory where each processor is optimized for the addressed application with a specialized instruction set and with the minimum allowable clock speed.

## Low-power in analog circuits

While in digital circuits the most efficient way to drastically reduce the power consumption is to reduce the supply voltage, in analog circuits the low-power techniques are quite different. In analog circuits, lowering the supply voltage does not automatically reduce power consumption. Indeed, the power consumption of
analog circuits is basically set by the Signal to Noise Ratio (SNR) and the operation frequency, and its minimum value is independent of the supply voltage. In analog circuits the power is dissipated to keep the energy of the signal larger than the thermal energy. Let us consider a $100 \%$ current efficient single pole transconductor used to charge and discharge an output capacitance. The power $P$, drawn from the supply voltage $V_{D D}$, to have a sinusoidal voltage across the capacitance $C$ with a peak-to-peak value $V_{P P}$ is $C f V_{P P}^{2}\left(V_{D D} / V_{P P}\right)$. The output noise power is $k T / C$, where $k$ is the Boltzmann's constant and $T$ is the absolute temperature. As a consequence the signal-to-noise ratio is,

$$
\begin{equation*}
S N R=\frac{V_{P P}^{2} / 8}{k T / C} . \tag{1-2}
\end{equation*}
$$

Then we can write,

$$
\begin{equation*}
P=8 k T f(S N R) V_{D D} / V_{P P} \tag{1-3}
\end{equation*}
$$

For a given $S N R$ and frequency, the minimum power consumption to realize a single pole is achieved for $V_{P P}=V_{D D}$, and then the minimum power consumption is independent of the supply voltage [19], [20]. From such consideration, it is clear that power efficiency circuits must be designed to be rail-to-rail. Such power limit can be applied at each pole of any linear analog circuit and it is very stringent because it requires a factor 10 of power increase every 10 dB of $S N R$ increase. As a consequence, analog circuits become very power inefficient, in the case of systems that require very large $S N R$. Such considerations do not depend on the technology or on the supply voltage.

Such theoretical lower limit to the power consumption is increased by other technological limitations in practical circuits’ implementations. For example the power dissipated in bias circuitry is wasted and, in addition, if the bias schemes are inadequate, they can increase the noise and then a larger power is required. Since the minimum power consumption is achieved for a voltage swing equal to the supply voltage, it is important that the signal is amplified as early as possible to its maximum value and maintained along the signal path.

Another aspect to be considered is the presence of additional noise sources, such as those internal to active and passive components or the noise coming from the supply voltage, which forces to increase the power consumption to maintain a given $S N R$. Another aspect that leads to an increase in the power is the need for precision, which imposes to use larger dimensions for active and passive components causing an increase of parasitic capacitances and then of the power. In the case of switched capacitors, the clock frequency must be at least twice as large as the maximum frequency of the signal and then the power consumed by the clock could be dominant. Different techniques exist to reduce the effect of such limitations by acting at circuit or device level.

Although in analog circuits the minimum power for a given $S N R$ and frequency does not depend on the supply voltage, because of the technological limits described above, such as the power dissipated in the bias circuitry or the fact that the voltage swing is not equal to the supply voltage, the power dissipated in an analog circuit slightly depends on the supply voltage. The trend in modern CMOS technologies is to reduce the supply voltage, especially to accomplish the power requirement of digital circuits. In micropower analog circuits the use of MOS transistors in weak inversion provides several advantages for low-power and lowvoltage purposes, as will be explained in detail in the following. An important aspect for low-voltage applications is that the drain-source saturation voltage in the weak inversion operation is much smaller than that in strong inversion, since it is sufficient that the drain-source voltage is larger than the thermal voltage; this helps to reduce the supply voltage. Moreover, in the weak inversion the transconductance-to-current ratio of a MOS transistor reaches its maximum value, approaching that of a bipolar transistor. Indeed, the I-V characteristics of a MOS in the saturation and in the weak inversion regions can be well approximated by,

$$
\begin{gather*}
I_{D}=\frac{\mu C_{o x}}{2} \frac{W}{L}\left(V_{G S}-V_{t h}\right)^{2}=\frac{k}{2}\left(V_{G S}-V_{t h}\right)^{2},  \tag{1-4}\\
I_{D}=\mu C_{o x} V_{T}^{2} \frac{W}{L} \exp \left(\frac{V_{G S}-V_{t h}}{m V_{T}}\right)\left[1-\exp \left(-\frac{V_{D S}}{V_{T}}\right)\right], \tag{1-5}
\end{gather*}
$$

where $\mu$ is the carrier mobility in the channel, $V_{T}$ is the thermal voltage, $V_{t h}$ is the threshold voltage, $m$ is the subthreshold slope parameter, $W$ and $L$ are the channel width and length, respectively. As a consequence the transconductance-to-current ratio of a MOS transistor in the saturation and in the weak inversion are given by $2 /\left(V_{G S}-V_{t h}\right) \quad$ and $1 /\left(m V_{T}\right)$, respectively. Since $V_{T}<\left(V_{G S}-V_{t h}\right)$, the transconductance-to-current ratio of a MOS transistor in the weak inversion is larger than that of a MOS in strong inversion. As a consequence, when the current is limited, weak inversion provides maximum gain-bandwidth product for a given load capacitance or minimum input equivalent noise for a given output noise. Moreover, weak inversion also provides maximum gain per device and minimum input referred offset in a differential pair.

On the other hand, the maximum value of the transconductance-to-current ratio of a MOS transistor leads to a maximum mismatch of current mirrors. Indeed, the current mismatch $\Delta I_{D}$ due to the threshold voltage mismatch $\Delta V_{t h}$ is given by $\Delta I_{D}=g_{m} \Delta V_{t h}$ and then, for a given current, a maximum mismatch is achieved in the weak inversion. Moreover, in the weak inversion the device has a higher noise. Indeed, the drain current noise has a spectral density given by $4 k T \gamma g_{0}$ (where $g_{0}$ is the drain-source conductance and is proportional to $g_{m}$ ) [21] and then, for a given
current, maximum noise is achieved. As a consequence, in low-voltage current mirrors a trade-off must be found between, on the one hand, low voltage operation and, on the other hand, worse precision and higher noise. The other drawback of MOS transistors in weak inversion is that the transition frequency does not exceed a few hundreds of MHz and that they can not be used in high-frequency analog applications. In the strong inversion region, the transition frequency of MOS transistors increases to some GHz but the power consumption becomes higher, as previously explained. For such reason, BiCMOS is the best technology for lowpower and high-frequency analog circuits.

In order to reduce power consumption in analog circuits, some techniques at the system level can be adopted. In analog circuits, the limit of the minimum power calculated in (1-3) can not be overcome and, usually, at least ten times more power will be needed for practical reason. This means that it is not possible to implement a 16-bit audio A/D converter $(S N R=98 \mathrm{~dB})$ with a power consumption smaller than $50-100 \mu \mathrm{~W}$ and other power will be needed to amplify the signal after the conversion. A few microwatts per pole will be sufficient for the subsequent signal processing. As a consequence, most of the power in signal processing chain is consumed in the analog interfaces when the dynamic range is smaller than $S N R$. Some power reduction technique can be used if the $S N R$ is much smaller than the dynamic range, as in the case of hearing aids where the speech transmission requires a $S N R$ of only 40 dB but a dynamic range larger than 100 dB . In such cases, power can be reduced by maintaining the $S N R$ at its minimum value, independently of the signal, and the dynamic range at the required value. This means that, if the signal is weak, the noise floor must be low and, if the signal is large, noise floor must be raised to achieve the required $S N R$. This can be achieved by using analog floating point technique, which basically consists in dividing the input signal by a proper factor so that the signal fits within a given range [22]. In such a way, the signal that enters the processor is always within a min-max range and then the SNR is always kept constant. Such technique well fits with switchedcapacitor circuits in which the updating can take place between two sampling instants.

### 1.2. Low-power electronics applications

### 1.2.1. Wireless sensor networks

A network of wireless sensors consists of a large number of energy-autonomous microsensors distributed in an area of interest. Each node monitors its local environment, locally processes and stores the collected data so it can be used by other nodes. It shares this information with the other neighboring nodes by using a wireless link. Specific features of interest to the end-user can be extracted from the
different information collected by several nodes. The wireless sensor networks can be used for several applications in the field of logistics, identification, medical applications, industrial control, etc..

Many applications can be thought in the field of logistics, asset tracking and supply chain management [23]. For example, the wireless sensor networks can be used to solve the problem of tracking container in a large port [24]. In a port there are thousands of containers stacked one on the other and some of them are empty, others are bound for many destinations. To improve the efficiency is necessary that the location of each container is known exactly that it is chosen so that the containers next needed are close to the ship, where they have to be loaded, and are on the top of the stack. The use of a sensor networks with a sensor on each container allows us to determine the position of each container. In the same way, the wireless sensor network can be used in the supply chain management to know the precise location of an item in a large warehouse. This means that it is possible to know the location of an item to be sold or to perform an automatic inventory, drastically reducing the costs.

Another application of wireless sensor network is for health monitoring, such as athletic performance monitoring to store pulse and respiration rate information and to send such information to a personal computer for later analysis, or daily blood sugar monitoring to record blood sugar values. Wireless sensor networks in health monitoring are expected to extend their field of applications to monitor some enzymes or other biological materials.

Several features distinguish the wireless sensor networks from other standard wireless network. The first one is the size. The nodes must be smaller than one cubic centimeter and less than 100 grams so that they can be embedded into the environment. Second, each node must be low-cost to enable the realization of sensor networks with a large number of nodes. This means that the single node, the communication protocol and the network design must have low complexity to satisfy the low-cost requirement. The most critical issue is the stringent power requirement, which requires a node's power consumption smaller than $100 \mu \mathrm{~W}$ [25]. Indeed, nodes are typically battery-operated and, since the nodes are many and they might be deployed in hardly accessible regions, they should not require any maintenance. The nodes have therefore to be energetically autonomous and hence the batteries can not be replaced or recharged. The last few years have seen the emergence of numerous new radio technologies. The trend in these technologies is to offer higher and higher data rates to enable the consumers to transfer larger quantity of data in smaller time. Anyway, such high-data-rate technologies do not address the low-end classes of application, which do not require such high speed and complexity. Among wireless commercial devices available today, the closest match is the Bluetooth transceiver, which consumes more than some tens of mWs and costs more than 10 dollars [4], [5]. Such performance, in terms of power and cost, is orders of magnitude above that required for wireless sensor networks. To reach this stringent power requirement the operating range of each node is limited
to a few meters and the data-rate is limited to a few kbps. Anyway, energy optimization must be performed at each level of the system design process, from the physical layer to the communication protocol.

### 1.2.2. Ambient Intelligence

In the near future, cars, offices and houses will have a distributed network of intelligent devices that provide information, communication and entertainment. These systems will adapt to the user in a context-aware fashion and will differ substantially from contemporary equipment in their appearance in people's environments and in the way users interact with them. Ambient intelligence is the term used to denote such paradigm. Ambient Intelligence refers to the presence of an environment that is sensitive, adaptive and responsive to the presence of people or objects [26], [27]. In a car environment, Ambient Intelligence can serve, for example, the purpose of safety improvements, intelligent navigation and comfort enhancement. In an office environment, Ambient Intelligence will serve productivity enhancement by supporting the office workers and by improving their living conditions. As an example of application, we can consider the management of environmental control systems in large office buildings. Distributed sensors and actuators allow us to monitor and control some environmental parameters, such as temperature, airflow, light, etc., improving the living conditions of the occupants by, for example, giving the possibility to create micro-climates according to occupants' preferences. Moreover, the wireless solution eliminates the costs of wires and of installing wiring for a single sensor. In a home environment, Ambient Intelligence will improve the quality of life by creating the desired atmosphere and functionality via intelligent, personalized inter-connected systems and services. It is possible to think at a remote control that can control all home electronic equipments, such as the television, DVD player, stereo, washing machine, but also the lights, the curtains, the locks, etc.; then, an intelligent system can offer some services, such as closing the curtain when the television is turned on, or automatically muting the television when a call is received. To make Ambient Intelligence a reality, many innovations have to be realized, both in hardware and software. An Ambient Intelligence system exhibits a multitude of environmentsystem interfaces (sensors, actuators and transducers), handling the complete conversion between external information sources/sinks and the digital signal processing world, consisting of a sensor/actuator/transducer combined with RF and mixed signal circuits. Low-data rate sensors and actuator control signals form the interface between environment and system with data rates as low as $1 \mathrm{~b} / \mathrm{s}$ or less. One of the challenges is to find power- and cost-optimized solutions at very low-data-rate, both wireless and wired. For a radio technology to succeed in the Ambient Intelligence applications, it must take into account the driving factors of all applications areas, such as extremely low-cost, ease of installation, short-range operation and reasonable battery life. Also for Ambient Intelligence applications,
there is the stringent need to develop wireless devices with a low complexity, to ensure low-cost, and with a very low-power consumption.

### 1.2.3. Implantable medical devices

Implantable medical devices (IMDs) are used in the treatment of many diseases, including heart diseases, neurological disorders and deafness [28]. IMDs are widely used in the treatment of arrhythmias, which is a condition of heart rhythm problems that occurs when the electrical impulses that coordinate heartbeats do not function properly, causing the heart to beat too fast, too slow or irregularly; such heart disease is treated by the use of pacemakers and Implantable Cardioverter Defibrillators (ICDs) [29], [30], [31], which guarantee the correct heartbeats. Another application field of IMDs is for the treatment of hearing loss [32]. A hearing aid is an electronic, battery-operated device that amplifies and modifies sound to allow for improved communication. Hearing aids receive sound through a microphone, which traduces the sound waves to electrical signals. The audio signal is amplified and sent to a loudspeaker. New ultra-low-power radio frequency technologies are spurring the development of innovative medical tools, from endoscopic camera capsules to implanted devices that wirelessly transmit patient health data. The most important requirement of IMDs is the very low power consumption required to extend the battery life time to several years since such devices are implanted and battery replacement is very difficult. Indeed, implanted battery power is limited and the impedance of the battery is relatively high, limiting peak currents that may be drawn from the supply ( $<6 \mathrm{~mA}$ ). The transceiver must operate in a low-power sleep mode, with an extremely low current $(<1 \mu \mathrm{~A})$, and with the capability to look periodically for a wake up signal..

### 1.2.4. Radio Frequency IDentification (RFID) Systems

In recent years, automatic identification procedures have become very popular in many service industries, purchasing and distribution logistic operations, manufacturing companies and material flow systems. Automatic identification procedures are used to provide information about people, animals, goods and products in transit. In RFID systems, the transfer of power and data from the reader to the transponder and viceversa is performed using radio communication. Also in such kind of applications, the power requirements in the design of the transponder is very critical, in order to extend the battery life time, in the case of an active transponder, or to extend the operating range, in the case of a passive transponder.

To meet the requirements of the applications within IMDs, wireless sensor networks, ambient intelligence, RFID systems, a successful design must have several specific features: extreme low-power, low-cost, low data-rate. The need for these features leads to a combination of interesting technical issues not found in other widespread wireless network technologies, such as Bluetooth, IEEE 802.11.

### 1.3. RFID technology

Radio Frequency Identification (RFID) technology proposes new solutions to replace the traditional automatic identification systems, such as those based on barcodes and smart cards. An RFID system consists of an ensemble of transponders, each applied to the objects to be identified, and a reader that interrogates the transponders via radio waves, [33]. In a barcode system, in order to read the information the barcode must be brought rather close to the reader (few tens of cm ) and in visual line of sight. Moreover, although the bar code is very cheap, it has a very low storage capacity and it can not be reprogrammed.

A more flexible solution is to store the information in a silicon chip. The most common way of electronic data-carrying device in use in everyday life is the smart card based on contact operation. However, the mechanical contact used in smart card is often impractical. In RFID systems, instead, a contactless transfer of data between the data-carrying device and the reader is performed via radiowaves. A transponder can be identified in a unique way by an identification code stored in the transponder. In principle, the silicon chip can store a large amount of information that can be read and written at a distance of several meters.

### 1.3.1. Brief history of RFID technology

The origin of RFID systems can be traced back to the World War II. In that period, exactly in 1935, Watson-Watt had been discovered the radar but there was the problem that the radar was able to warn of approaching planes but not to distinguish if the planes belonged to enemies or not. In order to solve such problem, the British introduced the first example of tag, which was installed on each plane; such tag, when the plane was approaching at the airport, received a signal from the radar stations and answered by transmitting a signal to identify the plane as friendly. This was the first example of transponder and in the following years a larger and larger number of scientists were involved in the research dealing with the identification by exploiting the RF energy [34], [35], [36]. The first example of commercial use of RFID system was the Electronic Article Surveillance systems, which employs 1-bit transponders, that can be set on or off according to if the item was paid or not.

The first patent for an RFID system was received by Mario W. Cardullo on January 23th, 1973 [37]. In the mid-1980s, an RFID system was commercialized for automatic toll payment: a transponder was installed on a car or truck and a reader at the gates. Such system was widely used for the automatic toll payment of roads, tunnels and bridge. In the same years, under the request of the US Agricultural Department, a passive RFID system at 125 kHz was developed for the identification of cows [38] and they are still currently used for the same purposes. Later, passive low-frequency transponders were also used for access control to buildings.

In the following years, RFID systems operating at higher frequencies were introduced to achieve larger operating ranges and larger bandwidth. At first, RFID systems at 13.56 MHz were used for access control, automatic toll payment and in contactless smart card. In 1990s, IBM introduced the first UHF RFID system, which was able to provide an operating range larger than 6 meters, and it was used in several applications, especially in the supply chain management [39]. But the technology was expensive at the time due to the low volume of sales and the lack of open, international standards. UHF RFID had an important boost in 1999, when the Uniform Code Council, EAN International, Procter \& Gamble and Gillette created the Auto-ID Center at the Massachusetts Institute of Technology. The objective of the Auto-ID Center was to develop low-cost RFID tags by putting only a serial number on the tag to keep the price down so that they could be applied on all products to track them through the supply chain. Then the serial number on the tag was read and stored in a database that would be accessible over the Internet. Previously, tags were a mobile database that carried information about the product or container they were on with them as they traveled. Now, RFIDs were turned into a networking technology by linking objects to the Internet through the tag.

In recent years automatic identification procedures have become very popular in many service industries, purchasing and distribution logistics industry, manufacturing companies and material flow systems. Some of the biggest retailers in the world - Albertsons, Metro, Target, Tesco, Wal-Mart- and the U.S. Department of Defense have said they plan to use RFID technologies to track goods in their supply chain [40],[41], [42], [43], [44].

The number of companies actively involved in the development and sale of RFID systems indicates that this market that should be taken seriously. Whereas global sales of RFID systems were approximately 900 million of dollars in the year 2000 it is estimated that, over the next five years, the market grows at a compound annual growth rate of almost $30 \%$, reaching $\$ 1.18$ billion in 2010 [45]. The RFID market therefore belongs to the fastest growing sector of the radio technology industry, including mobile phones and cordless telephones [46].

### 1.3.2. Classification of RFID systems

We can introduce different kinds of classification according to the feature we are considering.

An important feature of RFID systems is the power supply to the transponder. According to such aspect, transponders can be classified as passive or active. Passive transponders do not have an on-board battery and then all the power required to supply the transponder and to transmit data to the reader is generated by rectifying the RF power transmitted by the reader. Active transponders, instead, have an on-board battery to supply all or part (semi-passive) of the power required by the transponder.

The most important differentiation criterion for RFID systems is the physical coupling method, which strongly affects the achievable operating range. It is
possible to distinguish three different coupling methods, i.e. inductive coupling, electrical coupling and electromagnetic coupling. Each of them exploits a different physical principle and then can operate at different frequencies and can achieve different operating range.

Most RFID systems exploit the inductive coupling to transfer power and data between the reader and the transponder. In such systems, the coupling element is a coil that acts as antenna. Inductive-coupled transponders are almost always passive and then they draw the energy required for their operation by the reader. The reader coil generates a magnetic field that in part concatenates with the transponder's coil and an alternating voltage is generated at the transponder coil terminals. Such voltage is then rectified to generate the DC power required for the operation of the transponder. Usually, in order to boost the voltage generated at the transponder coil terminals, a capacitor is added in parallel with the antenna coil to create a resonant circuit at the transmission frequency of the tag-reader systems. The inductive coupling systems thus exploit the transformer-type coupling between the primary coil in the reader and the secondary coil in the transponders. In order to have such type of coupling, the transponder must be located in the near field of the transmitter coil antenna, that is the distance between the coils must be smaller than $0.16 \lambda$, where $\lambda$ is the wavelength associated to the transmission frequency. For such reason, inductive coupling systems can not operate at very high frequency otherwise the operating range would be drastically reduced. Since the efficiency of power transfer between the two coils is proportional to the operating frequency, the number of windings and their area, for a given efficiency, the higher the operating frequency of the RFID systems and the smaller is the size of the antenna coils in the reader and the transponder [33]. As a consequence, a trade-off must be found between operating range and sizes. They typically operate at 135 kHz or 13.56 MHz and the operating range is smaller than 1 m .

In electrically coupled RFID systems, the coupling element consists of a large metal plate, which acts as an electrode, and they typically are passive. By applying an alternating voltage at the reader's electrode, an alternating electrical field is generated which couples with the transponder's electrode, allowing power transfer from the reader to the transponder. Also in this case, a resonant circuit is created in the transponder to step up the voltage generated at the transponder's terminals. In order to ensure the capacitive coupling, the electrodes of transponder and reader must be close to each other. For such reason, such systems typically have operating range of few centimeters.

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 microwave range ( 2.45 GHz or 5.8 GHz ). Such systems can achieve an operating range of few meters, in the case of passive transponder, and larger than 15 m , in the case of active transponder. The main advantage is the possibility to achieve a higher operating range, which is required for the adoption of RFID systems in many
logistics and tracking applications. Moreover, since such systems have a high carrier frequency they can have a large bandwidth that allows transmission of large data volumes. Since the size of the coupling element is proportional to the wavelength, electromagnetic RFID systems can have smaller size than that of inductive coupling RFID systems.

Furthermore, passive RFID systems with electromagnetic coupling allow one to achieve operating ranges of some meters with no battery to be maintained. . The widespread adoption of passive electromagnetic coupling RFID systems strictly depends upon the possibility of extending the operating range to several meters by drastically reducing the power consumption of a transponder.

Another classification is done according to the possibility of writing information in the transponder. In read-only transponders, the identification code is set at the fabrication and it can not be modified anymore. Writeable transponders, instead, have an embedded EEPROM memory where the identification code and additional information is stored.

### 1.3.3. International standards

An important issue, which limits the widespread adoption of RFID systems, is that too many different standards still exist, i.e. EPC, several ISO standards, and proprietary standards. Such standards specify the communication protocol and parameters for air interface. The EPC standard was developed by Auto-ID Center and it is currently managed by EPC Global. EPC currently includes three different standards for RFID systems:

Class 0 for UHF RFID transponder [47];
Class 1 for 13.56 MHz - and UHF- RFID systems [48];
Class 1 Gen 2 for UHF RFID systems [49].
ISO is the International Standardization Organization has developed many standards for RFID systems according to the application they are using for. The standard ISO 18000 is one of the most widely adopted. It defines the parameters for air interface of transponders and it is divided into 7 parts according to the operation frequency of the RFID systems. More in detail, we have,

ISO 18000-1: generic parameters of air interface [50];
ISO 18000-2: parameters for air interface $<135 \mathrm{kHz}$ [51];
ISO 18000-3: parameters for air interface at 13.56 MHz [52];
ISO 18000-4: parameters for air interface at 2.45 GHz [53];
ISO 18000-5: parameters for air interface at 5.8 GHz [54];
ISO 18000-6: parameters for air interface at $860-930 \mathrm{MHz}$ [55];
ISO 18000-7: parameters for air interface at 433.92 MHz [56].
Besides such standards that define the communication protocol and air interface parameters we have also to mention the standard that defines frequency, power and channels that can be used without interfere with other existing communication standards. In Europe, in September 2004, ETSI defined the standard ETSI EN 302

208-1 that fixes at 2 W the maximum power level that can be transmitted by Radio Frequency Identification equipment operating in the band 865 MHz to 868 MHz [57]. Anyway, Italy is one of the last countries in Europe that has not adopted such standard yet. Indeed, Italy still adopts ERC/REC Recommendation 70-03, which, according to the national restrictions, fixes the maximum power, which can be transmitted by non-specific short range devices, is 25 mW ERP [58]. In US, in 2001, FCC defined the standard FCC - Part 15 that fixes at 4 W the maximum power that can be transmitted by radio frequency devices operating with a frequency larger than 916 MHz [59].

### 1.3.4. Passive UHF/Microwave RFID systems

Long range passive transponders ("tags") for RFID systems in the UHF or microwave frequency range do not have an on-board battery, and therefore must draw the power required for their operation from the electromagnetic field transmitted by the reader. The maximum power that can be transmitted by the reader is limited to 500 mW in Europe [57], according to the standard issued by ETSI, and 4 W in US [59], according to the standard issued by FCC. The RF energy radiated by the reader is used both to supply the digital section of the transponder and to allow data transmission from the tag to the reader through modulation of the backscattered radiation. If the transponder lies within the interrogation range of the reader, an alternating RF voltage is induced on the transponder antenna, which typical is a dipole or a patch antenna, and is rectified in order to provide a DC supply voltage for transponder operation. In addition, most of the passive and semi-passive RFID systems that operate in the UHF or microwave range exploit modulation of the backscattered radiation to transmit data from transponder to reader: while the reader transmits an unmodulated carrier, the data signal modulates the load of the transponder antenna in order to modulate the backscattered electromagnetic field, typically with ASK or PSK. Then the digital section is a very simple microprocessor or a finite state machine that must be able to manage the communication protocol, according to the standard.

Many commercial and research prototypes of passive RFID systems in the UHF and microwave frequency ranges have been presented in the last few years.
Several companies are involved in the development of passive UHF or microwave RFID transponder, such as Texas Instrument, STMicroelectronics, Symbol, ATMEL and Transcore. Commercial prototypes and their performance are shown in Table 1-I. Other research prototypes can be found in the literature [66], [67], [68]. Their performance is summarized in Table 1-II. A great interest thus exists both from the industrial and academic point of view. Anyway, the operating ranges achieved by commercial and research prototypes are still quite small because a large power consumption of the transponder. By assuming that the transponder's antenna is perfectly matched with the input of the transponder and that the operating range $r$ is limited by the input power of the transponder, the operating range, in a first approximation can be expressed as follows,

Table 1-I: Commercial prototypes of passive UHF/microwave RFID transponders.

| Company | Model | Operating frequency | Operating range | Data-rate | Memory size |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Symbol [60] | $\begin{aligned} & \text { RFX- } \\ & 6000 \end{aligned}$ | UHF | Read: 7.5m(US) <br> Write: 3m (US) | 1 kbps | 288 bits |
| Philips [61] | $\begin{aligned} & \text { UCODE } \\ & \text { HSL } \\ & \hline \end{aligned}$ | 2.45 GHz | $\begin{aligned} & 0.6 \mathrm{~m} \text { (EU) } \\ & 1.8 \mathrm{~m} \text { (US) } \end{aligned}$ | 40 kbps | 2048 bits |
| Philips [61] | UCODE HSL | UHF | $\begin{aligned} & 4 \mathrm{~m} \text { (EU) } \\ & 8.4 \mathrm{~m} \text { (US) } \end{aligned}$ | 40 kbps | 2048 bits |
| Philips [61] | UCODE <br> EPC G2 | UHF | 7 m (US) | 640 kbps | 512 bits |
| TI [62] | Gen 2 | UHF | N/A | 40 kbps | 128 bits |
| STM [63] | XRA 00 | UHF | N/A | 140 kbps | 128 bits |
| STM [63] | XRAG2 | UHF | N/A | 640 kbps | 432 bits |
| $\begin{aligned} & \text { ATMEL } \\ & \text { [64] } \\ & \hline \end{aligned}$ | $\begin{aligned} & \text { ATA559 } \\ & 0 \\ & \hline \end{aligned}$ | UHF | 15 m (US) | 60 kbps | 1000 bits |
| Transcore [65] | AT5110 | UHF | 3 m (US) | N/A | 120 bits |

Table 1-II: Research prototypes of passive UHF/microwave RFID transponder.

| Work | Operating <br> frequency | Operating <br> range | Data-rate | Memory <br> size |
| :--- | :--- | :--- | :--- | :--- |
| Curty [66] | 2.45 GHz | 12 m (US) | N/A | No |
| Nakamoto [67] | UHF | $4.3 \mathrm{~m}(\mathrm{US})$ | 40 kbps | 2000 bits |
| Karthaus [68] | UHF | $4.5 \mathrm{~m}(\mathrm{EU})$ <br> $9.25 \mathrm{~m}(\mathrm{US})$ | N/A | N/A |

$$
\begin{equation*}
r=\sqrt{\eta \frac{P_{E I R P}}{4 \pi P_{I N}} A_{e}}, \tag{1-6}
\end{equation*}
$$

where $A_{e}$ is the effective aperture of the transponder's antenna, $P_{I N}$ is the input power of the transponder, $\eta$ is the power efficiency of the transponder, $P_{\text {EIRP }}$ indicates the power at which an isotropic emitter would have to be supplied to generate the same radiation power of the reader antenna. By assuming an efficiency of $37 \%$, which is that achieved in some research prototypes [66], [67], to use a dipole antenna and an input power of the transponder equal to $1 \mu \mathrm{~W}$, the operating range achievable in Europe and US, at 2.45 GHz , is 5.3 m and 15 m , respectively; in the UHF band is 15 m and 40 m , respectively. As a consequence, it is clear enough that by reducing the power consumption of the transponder to about $1 \mu \mathrm{~W}$, the operating range of the transponder can be improved very much compared to present available prototypes and commercial devices. Such results can be achieved, at the system level, by a proper dimensioning of the tag-reader system for the
choice of the modulation depth and, at physical level, by designing very low power circuits.

### 1.4. References

[1] H. Ruegg, W. Thommen, P. Sauthier, "A saturation-controlled flip-flop for low voltage micropower systems," ISSCC Dig. Tech. Papers, San Francisco, USA, pp. 60-61, 1971.
[2] F. Leuenberger, E. Vittoz, "Complementary-MOS low-power low-voltage integrated binary counter," Proc. IEEE, Vol. 57, pp. 1528-1532, 1969.
[3] E. Vittoz, B. Gerber, F. Leuenberger, "Silicon-Gate Frequency Divider for the Electronic Wrist Watch," IEEE Journal of Solid State Circuits, Vol. SC-7, No. 2, pp. 100-104, 1972.
[4] C. Cojocaru et al., "A 43mW Bluetooth Transceiver with -91dBm Sensitivity," ISSCC Dig. Tech. Papers, San Francisco, USA, pp. 90-91, 2003.
[5] N. Filiol et al., "A 22mW Bluetooth Transceiver with Direct RF Modulation and On-chip IF Filters," ISSCC Dig. Tech. Papers, San Francisco, USA, pp. 202-203, 2001.
[6] A.P. Chandrakasan, S. Sheng, R.W. Brodersen, "Low-Power CMOS Digital Design," IEEE Journal of Solid State Circuits, Vol. 27, pp. 473-484, 1992.
[7] T.G. Noll, E. de Man, "Pushing the Performance Limits due to Power Dissipation of Future ULSI Chips," ISSCC Dig. Tech. Papers, San Francisco, USA, Vol. 28, pp. 10-17, 1993.
[8] J.C. Shah, S.S. Sapatnekar, "Wiresizing with buffer placement and sizing for power-delay tradeoffs," Proc. of VLSI Design, pp. 346-351, 1996.
[9] L. Benini, S.K. Shuklam, R.K. Gupta, "Architectural, system level and protocol level techniques for power optimization for networked embedded systems," Proc. of VLSI Design, pp. 18-21, 2005.
[10]F. Carbognani et al., " $42 \%$ power savings through glitch-reducing clocking strategy in a hearing aid application," Proc. of ISCAS, pp. 4-7, 2006.
[11]Y. Taur, T.H. Ning, "Fundamentals of Modern VLSI Devices", Cambridge University Press, Cambridge, United Kingdom, 1998, pp. 96.
[12]G.D. Wilk, R.M. Wallace, J.M. Anthony, "High-k gate dielectrics: Current status and materials properties and considerations," Journal of Applied Physics, Vol. 89, No. 10, pp. 5243-5275, 2001.
[13] S. Mutoh et al., "1-V Power Supply High-Speed Digital Circuit Technology with Multi-Threshold Voltage CMOS," IEEE Journal of Solid State Circuits, Vol. 30, No. 8, pp. 847-854, 1995.
[14]K. Suzuki et al., "A 300 MIPS/W RISC core processor with variable supplyvoltage scheme in variable threshold-voltage CMOS," Proc. of CICC, pp. 587590, 1997.
[15] S. Mukhopadhyay et al., "Gate leakage reduction for scaled devices using transistors stacking," IEEE Transactions on Very Large Scale Integration Systems, pp. 716-730, 2003.
[16]H. De Man, "Ambient Intelligence: Gigascale Dreams and Nanoscale Realities," ISSCC Dig. of Tech. Papers, pp. 29-35, 2005.
[17]F. Catthoor et al., "Code Transformations for Data Transfer and Storage Exploration Preprocessing in Multimedia Processor," ISSCC Design and Test, Vol. 18, No. 3, pp. 70-81, 2001.
[18]H. De Man et al., "Filling the Gap Between System Conception and Silicon/Software Implementation," ISSCC Dig. of Tech. Papers, pp. 158-159, 2002.
[19]E.A. Vittoz, "Low-Power Design: Ways to Approach the Limits," Proc. of IEEE ISCAS, pp. 14-18, 1994.
[20]R. Castello, P.R. Gray, "Optimal dynamic range integrators," IEEE Transactions on Circuits and Systems, Vol. CAS-32, pp. 865-876, 1985.
[21]B. Razavi, Design of Analog CMOS Integrated Circuits, McGraw Hill, New York, pp. 212-215, 2001.
[22] E.M. Blumenkrantz, "The analog floating point technique," IEEE Symp. on Low-Power Electronics Dig. of Tech. Papers, pp. 72-73, 1995.
[23]E.H. Callaway, Wireless Sensor Networks: Architectures and Protocols, Boca Raton, USA: Auerbach Publications, pp. 1-10, 2003.
[24]J.L. Schoeneman, H.A. Smartt, D. Hofer, "WIPP transparency project -container tracking and monitoring demonstration using the authenticated tracking and monitoring system (ATMS)," Waste Management Conference, Tucson, Arizona, 2000.
[25]C.C. Enz, N. Scolari, U. Yodprasit, "Ultra Low-Power Radio Design for Wireless Sensor Networks," IEEE International Workshop on RadioFrequency Integration Technology, Singapore, pp. 1-4, 2005.
[26]F. Boekhorst, "Ambient Intelligence, the Next Paradigm for Consumer Electronics: How will it Affect Silicon?," ISSCC Dig. of Tech. Papers, San Francisco, USA, 2002.
[27] T. Basten, L. Benini, A. Chandrakasan, M. Lindwer, J. Liu, R. Min, F. Zhao, "Scaling into Ambient Intelligence," Proc. DATE 2003, pages 76-81, 2003.
[28]L.J. Scotts, "VLSI Applications in Implantable Mediacal Electronics," IEDM Dig. Of Tech. Papers, pp. 9-14, 1989.
[29] J. Berkman, J. Prak, "Biomedical Microprocessor with Analog I/O," ISSCC Dig. of Tech. Papers, San Francisco, USA, pp. 168-169, 2002.
[30]L. Scotts, J. Miner, R. Baker, "An 8b microcomputer for implantable biomedical applications," ISSCC Dig. of Tech. Papers, San Francisco, USA, pp. 14-15, 2002.
[31]J. Ryan, K. Carroll, B. Pless, "A Four Chip Implantable Defibrillator/Pacemaker Chipset," CICC Digest of Technical Papers, pp. 7.6.17.6.4, 1989.
[32]F. Callias, F. Salchli, D. Girard, "A Set of Four ICs in CMOS Technology for a Programmable Hearing Aid," IEEE Journal of Solid State Circuits, Vol. 24, No. 2, 1989.
[33]K. Finkenzeller, RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, $2^{\text {nd }}$ ed, Wiley and Sons, 1999, pp. 117-126, pp. 143-148 and pp. 183-186.
[34]L. Barnette, "Ship Identification," Transactions of the IRE Professional Group on Communications Systems," Vol. 3, No. 1, pp. 65-66, 1955.
[35] G. Leopard, "The flight evaluation of aircraft antennas," IEEE Transactions on Antennas and Propagation," Vol. 8, No. 2, pp. 158-166, 1960.
[36] A.S. Palatnick, H.R. Inhelder, "Automatic vehicle identification system—Methods of approach," IEEE Transactions on Vehicular Technology, Vol. 19, No. 1, pp. 128-136, 1970.
[37]M.W. Cardullo, W.L. Parks, "Transponder Apparatus and System," US Patent 3713 148, 1973.
[38] J.P. Hanton, H.A. Leach, "Electronic Livestock Identification System," US Patent 4262 632, 1981.
[39]"The History of RFID Technology", RFID Journal. Available at: http://www.rfidjournal.com
[40]M. Roberti, "The second-largest supermarket chain in the United States met with suppliers and explained how it plans to roll out RFID technology", RFID Journal, available at: http://www.rfidjournal.com/article/articleview/1227/1/1.
[41]R. Wessel, "Metro Group to Roll Out RFID at up to 150 Sites", RFID Journal, available at: http://www.rfidjournal.com/article/articleview/2772.
[42]M. Roberti, "Target Testa RFID for Security", RFID Journal, available at: http://www.rfidjournal.com/article/articleview/1282/1/1.
[43]"Tesco RFID Rollout Starts in April", RFID Journal, available at: http://www.rfidjournal.com/article/articleview/658/1/1/
[44] Wal-Mart, "Continued Expansion of Radio Frequency Identification (RFID)", 2004.Available at: http://walmartstores.com/GlobalWMStoresWeb.
[45]A. Nathanson, "RFID READER MARKET WORTH \$1.18B IN 2010", Venture Development Corporation Press Center, 2006.
[46] T. Smith, "World mobile phone market growth to stall", 2006, available at:
http://www.channelregister.co.uk/2006/01/13/isuppli_mobile_phone_market_f orecast/
[47] 900 MHz Class 0 Radio-Frequency Identification Tag, February 2003: available at: http://www.epcglobalinc.org/standards/.
[48] $860 \mathrm{MHz}-930 \mathrm{MHz}$ Class I Radio-Frequency Identification Tag Radio Frequency \& Logical Communication Interface Specifications, November 2002: available at: http://www.epcglobalinc.org/standards/.
[49]EPC Radio-Frequency Identity Protocols Class-1 Generation-2 UHF RFID Protocols for Communications at $860 \mathrm{MHz}-960 \mathrm{MHz}$ : available at: http://www.epcglobalinc.org/standards/.
[50]ISO/IEC 18000-1, "Generic Parameters for the Air Interface for Globally Accepted Frequencies", 2004.
[51]ISO/IEC 18000-2, "Parameters for Air Interface Communications below 135 kHz", 2004.
[52] ISO/IEC 18000-3, "Parameters for Air Interface Communications at 13.56 MHz", 2004.
[53]ISO/IEC 18000-4, "Parameters for Air Interface Communications at 2.45 GHz", 2004.
[54] ISO/IEC 18000-5, "Parameters for Air Interface Communications at 5.8 GHz ", 2004.
[55] ISO/IEC 18000-6, "Parameters for Air Interface Communications at 860 to 930 MHz", 2004.
[56] ISO/IEC 18000-2, "Parameters for Air Interface Communications at 433.92 MHz", 2004.
[57]ETSI EN 302 208-1, "Electromagnetic compatibility and Radio spectrum Matters (ERM); Radio Frequency Identification Equipment operating in the band 865 MHz to 868 MHz with power levels up to 2 W ; Part 1: Technical requirements and methods of measurements", 2004.
[58]ERC/REC Recommendation 70-03, Appendix 3 - National Restrictions.
[59]Federal Communication Commission, Part 15 - Radio Frequency Devices, October 2001.
[60] Symbol Gen 2 RFX 6000 datasheet. Available at: http://www.symbol.com/gen2tags.
[61] Philips UCODE datasheet. Available at: http://www.nxp.com/products/identification/ucode/index.html
[62]Texas Instrument Gen 2 datasheet. Available at: http://www.ti.com/rfid/shtml/prod-trans.shtml.
[63]STMicroelectronics XRA datasheet. Available at: http://www.st.com/stonline/products/families/memories/rfid/rfid.htm
[64] ATMEL ATA 5590 datasheet. Available at: http://www.atmel.org/products/RFID.
[65] Transcore AT 5110 datasheet. Available at: http://www.transcore.com/wdtranscoreproducts.html
[66] J.-P. Curty, N. Joehl, C. Dehollain, M.J. Declercq, "Remotely Powered Addressable UHF RFID Integrated System," IEEE Journal of Solid State Circuits, Vol. 40, No. 11, pp. 2193-2202, 2005.
[67]H. Nakamoto et al., "A Passive UHF RFID Tag LSI with 36.6\% Efficiency CMOS-Only Rectifier and Current-Mode Demodulator in $0.35 \mu \mathrm{~m}$ FeRAM Technology," ISSCC Dig. Tech. Papers, San Francisco, USA, pp. 310-311, 2006.
[68] U. Karthaus, M. Fischer, "Fully Integrated Passive UHF RFID Transponder IC With 16.7- $\mu \mathrm{W}$ Minimum RF Input Power," IEEE Journal of Solid State Circuits, Vol. 38, No. 10, pp. 1602-1608, 2003.

## 2. Design Criteria and Architecture

### 2.1. Introduction

Long range passive transponders ("tags") for RFID systems do not have an onboard battery, and therefore must draw the power required for their operation from the electromagnetic field transmitted by the reader [1]. The RF energy radiated by the reader is used both to supply the digital section of the transponder and to allow data transmission from the tag to the reader through modulation of the backscattered radiation. If the transponder lies within the interrogation range of the reader, an alternating RF voltage is induced on the transponder antenna, and is rectified in order to provide a DC supply voltage for transponder operation. In order to further increase the supply voltage, an N -stage voltage multiplier is typically used, providing a DC output voltage, at constant input power, roughly $N$ times larger than that achievable with a single stage. In addition, most of the passive and semi-passive RFID systems that operate in the UHF or microwave range exploit modulation of the backscattered radiation to transmit data from transponder to reader: while the reader transmits a unmodulated carrier, the data signal modulates the load of the transponder antenna in order to modulate the backscattered electromagnetic field, typically with ASK or PSK [1].

It is apparent that the larger the modulation of the impedance seen by the antenna, the larger the modulation depth and the signal-to-noise ratio at the reader, but also the larger the mismatch, and therefore the smaller the DC power converted by the voltage multiplier.

In order to maximize the operating range, it is important to achieve a non trivial trade-off between the desired error probability at the reader, and the DC power available for supplying the transponder, which is also strongly dependent on the power efficiency of RF-DC conversion.

The maximization of the conversion efficiency requires the optimization of the voltage multiplier and of the power matching network, taking into account the nonlinear behavior of the voltage multiplier.

The architecture of a passive microwave RFID transponder is shown in Fig. 2-1. The coupling element is an antenna, which typically is a dipole or a patch antenna. A voltage multiplier converts the input alternating voltage into a DC voltage which is used by a series voltage regulator to provide the regulated voltage required for the correct operation of the transponder. The voltage multiplier is matched with the antenna in order to ensure the maximum power transfer from the transponder's


Fig. 2-1: Passive Transponder Architecture.
antenna to the input of the voltage multiplier. A backscatter modulator is used to modulate the impedance seen by the transponder's antenna, when transmitting. The RF section is then connected to the digital section, which typically is a very simple microprocessor or a finite state machine able to manage the communication protocol.

In this paper we present a set of design criteria for the RF section of passive transponders in the UHF and microwave frequency range referring to the architecture shown in Fig. 2-1, with the main objective of maximizing the operating range. We therefore focus on the optimization of the voltage multiplier and on its power matching to the antenna, and we derive a set of criteria that allow us to choose and optimize backscatter modulation in order to either maximize the operating range, once the data-rate is fixed, or maximize the data-rate, once the operating range is fixed.

In the rest of the paper, all numerical examples will refer to the 0.35 mm CMOS technology from AMS, but of course our considerations can be applied to any technology.

Our investigation will show that, for a passive RFID system compliant to European regulations in the 2.45 GHz or 868 MHz ISM frequency bands, the achievable operating range, considering a power consumption of the digital section of the transponder of $1 \mu \mathrm{~W}$, is larger than 3.4 m and 9.5 m , respectively. At the same time, we will show that, for a passive $2.45-\mathrm{GHz}$ RFID system, given an operating range of 3.4 meters, the achievable data-rate is about 17 kbps and for a passive $868-\mathrm{MHz}$ RFID system, given an operating range of 9.5 meters, the achievable data-rate is about 70 kbps . Considering the more permissive US regulations, the maximum achievable operating distances are 11 m in the 2.45 GHz frequency band, and 29 m at 916 MHz , considering a data-rate of some tens of kbps. The extremely low power consumption considered for the digital logic is achievable by using subthreshold logic schemes, given that a simple finite state


Fig. 2-2: $N$-stage voltage multiplier and cascaded series voltage regulator.
machine operating at a frequency smaller than 1 MHz is typically adequate to implement RFID protocols. However, such aspect is beyond the scope of the present paper and will not be discussed here.

### 2.2. Voltage Multiplier and Power Matching Network

In this section we will describe the design criteria for both the voltage multiplier and the power matching network, in order to maximize the power efficiency of the transponder, defined as the ratio between the RF power available at the transponder's antenna and the DC power at the output of the voltage multiplier available for supplying the transponder. As we will explain in the next section, the power efficiency of the transponder strongly affects the operating range of the tagreader system.

### 2.2.1. $\quad$-stage Voltage Multiplier

An N -stage voltage multiplier consists of a cascade of $N$ peak to peak detectors, as shown in Fig. 2-2 [2]. Let us suppose to apply, at the input of the voltage multiplier, a sinusoidal voltage, $V_{I N}$, with a frequency $f_{0}$ and an amplitude $V_{0}$. In order to ensure a small ripple in the output voltage $V_{U}$, the capacitors indicated with $C$ in Fig. 2-2 have to be dimensioned so that their time constant is much larger than the period of the input signal, that is, $I_{U} /\left(2 \pi C V_{U}\right) \ll f_{0}$, where $I_{U}$ is the DC output


Fig. 2-3: Simplified equivalent circuit of the considered diodes: a) substrate losses neglected; b) equivalent circuit including substrate losses.
current. In this way, the voltage across $C$ capacitors and the output voltage can be considered a DC voltage. As a consequence, in the high frequency analysis, it is possible to consider $C$ capacitors as short-circuits and therefore all diodes appear to lie directly in parallel or anti-parallel to the input. In this situation the input RF voltage entirely drops across the diodes. In the DC analysis, $C$ capacitors can be considered as open circuits, so that we have $2 N$ identical diodes in series with the output. The voltage $V_{d}$ that drops across each diode is therefore given by

$$
\begin{equation*}
V_{d}= \pm V_{0} \cos \left(\omega_{0} t\right)-\frac{V_{U}}{2 N} \tag{2-1}
\end{equation*}
$$

where the sign ' + ' is applied to diodes with even subscript (see Fig. 2-2) and the sign '-' is applied to diodes with odd subscript. We can represent the equivalent circuit of the diode as an ideal diode in parallel with a capacitance, $C_{D}$, as shown in Fig. 2-3a, neglecting diode series resistances. Indeed, since the DC power required by RFID passive transponders is quite low (in the order of few $\mu \mathrm{W}$ ), the DC output current of the voltage multiplier is very small leading to a negligible effect of the series resistance of the diodes. Such hypothesis was verified by circuit simulations. Thus the current $I_{d}$ in each diode is

$$
\begin{equation*}
I_{d}=I_{S}\left[\exp \left( \pm \frac{V_{0}}{V_{T}} \cos \left(\omega_{0} t\right)\right) \exp \left(-\frac{V_{U}}{2 N V_{T}}\right)-1\right]+C_{D} \frac{d V_{d}}{d t} \tag{2-2}
\end{equation*}
$$

where $I_{S}$ is the diode saturation current and $V_{T}$ is the thermal voltage. We can express the exponential of a co-sinusoidal function using the modified Bessel


Fig. 2-4: a) Required amplitude of the input voltage vs. the number of stages for an output power of $5 \mu \mathrm{~W}$ b) Required input power vs. the number of stages for an output power of $5 \mu \mathrm{~W}$, for an $I_{S}$ of 10 aA .
functions series expansion [3], as shown below,

$$
\begin{equation*}
\exp ( \pm x \cos (\omega t))=B_{0}( \pm x)+2 \sum_{n=1}^{\infty} B_{n}( \pm x) \cos (n \omega t) \tag{2-3}
\end{equation*}
$$

Since modified Bessel functions of odd (even) order are odd (even) [3], the DC current in each diode, which also is the DC current $I_{U}$ in the output load, is given by

$$
\begin{equation*}
I_{U}=I_{S}\left[B_{0}\left(\frac{V_{0}}{V_{T}}\right) \exp \left(-\frac{V_{U}}{2 N V_{T}}\right)-1\right] . \tag{2-4}
\end{equation*}
$$

As a consequence, the input-output characteristics of the $N$-stage voltage multiplier is intrinsically expressed by

$$
\begin{equation*}
\left(1+\frac{I_{U}}{I_{S}}\right) \exp \left(\frac{V_{U}}{2 N V_{T}}\right)=B_{0}\left(\frac{V_{0}}{V_{T}}\right) . \tag{2-5}
\end{equation*}
$$

The above equation can be easily solved by numerical iteration, yielding the monotonously decreasing behaviour of $V_{0}$ as a function of $N$ for fixed DC output voltage and power consumption plotted in Fig. 2-4a, for an $I_{S}$ of 10 aA , which is the $I_{S}$ for a minimum diode area in the technology we are considering. However, for large $N$ the curves almost saturate, since voltage multiplication is limited by the voltage drops on the diodes.

## Power Consumption

The average input power $P_{I N}$ required to obtain a given output voltage and power, can be calculated by summing up the average power, $P_{D}$, dissipated in each diode and the power, $P_{L}$, required by the load. Neglecting substrate losses, the average power dissipated in each diode is given by

$$
\begin{equation*}
P_{D}=\frac{1}{T} \int_{0}^{T} V_{d}(t) I_{d}(t) d t \tag{2-6}
\end{equation*}
$$

where $T$ is the period of the input voltage. By solving the integral with the expressions of $V_{d}(t)$ and $I_{d}(t)$, already given in (2-1) and (2-2), and taking into account the properties of the modified Bessel functions, it is possible to obtain:

$$
\begin{equation*}
P_{D}=I_{S} V_{0} B_{1}\left(\frac{V_{0}}{V_{T}}\right) \exp \left(-\frac{V_{U}}{2 N V_{T}}\right)-\frac{P_{L}}{2 N} . \tag{2-7}
\end{equation*}
$$

As a consequence, the average input power is given by,

$$
\begin{equation*}
P_{I N}=2 N P_{D}+P_{L}=2 N I_{S} V_{0} B_{1}\left(\frac{V_{0}}{V_{T}}\right) \exp \left(-\frac{V_{U}}{2 N V_{T}}\right) . \tag{2-8}
\end{equation*}
$$

Considering that $V_{0}$ is a function of $N$ from (2-5), in order to obtain the plot of $P_{I N}$ as a function of $N$, shown in Fig. 2-4b, for a fixed output power, it is necessary to solve the equation by a numerical iteration. From the plot shown in Fig. 2-4b, it is possible to note that the maximum power efficiency is obtained by using the minimum number of stages. In order to provide a criterion for the choice of the


Fig. 2-5: a) Required $Q$-factor of the matching network vs. the number of stages for an output power of $5 \mu \mathrm{~W}$ b) Average input power vs. diodes area for a single stage voltage multiplier with an output power of $5 \mu \mathrm{~W}$.
appropriate number of stages, we need to find a relation between the number of stages of the voltage multiplier and the $Q$ factor of the power matching network placed between the antenna and the voltage multiplier. We can define an equivalent input resistance of the voltage multiplier, considering the power consumption, as shown below,

$$
\begin{equation*}
R_{e q}=\frac{V_{0}^{2}}{2 P_{I N}} . \tag{2-9}
\end{equation*}
$$

Now, assuming to use an LC matching network, the $Q$-factor of the LC matching network is bound to the resistance transformation ratio [4] and its is given by

$$
Q=\sqrt{\frac{R_{e q}}{R_{A}}-1}
$$

where $R_{A}$ is the antenna resistance, neglecting antenna losses. Assuming that the antenna is a dipole (antenna resistance $72 \Omega[1]$ ), from (2-5), (2-9) and (2-10) we plot $Q$ as function of $N$ in Fig. 2-5a. It is clear that similar conclusions can be obtained using other types of antennas.

Although the highest efficiency is reached with only one stage, the choice of the number of stages has to be done taking into account the values of $Q$ that can be physically achievable, which, typically, are not larger than few tens. A possibility to reduce the required $Q$ is represented by the choice of antennas with higher radiation resistance, such as for example a 2 -wire or a 3-wire folded dipole. Once the number of stages is chosen following the above criteria, it is necessary to dimension the diodes in order to optimize the power efficiency of the voltage multiplier. Considering that $I_{S}$ is proportional to the diode area, by solving the equation $P_{I N}\left(I_{S}\right)$ for a fixed value of $N$ and for a certain value of the output voltage and power, one obtains the curve plotted in Fig. 2-5b. It is possible to note that the higher is the diode area and then the saturation current of the diodes and the better is the power efficiency of the voltage multiplier. But it is not possible to increase excessively the diode area, otherwise the diode capacitance would become comparable with the $C$ capacitances of the voltage multiplier. This situation would lead to a reduction of the efficiency because only a fraction of the input voltage would drop across the diodes. The previous consideration suggests that the best choice would be represented by Schottky diodes, which have higher saturation current compared to normal diodes, for a fixed area. Furthermore, since diodes in the voltage multiplier must have a switching time smaller than the period of the input signal, Schottky diodes have to be preferred because they are typically much faster than p-n diodes. However, we will focus our attention on p-n junction diodes that have the advantage of being available in a cheaper CMOS process.

## Power consumption in the presence of substrate losses

In order to take in consideration the substrate losses in the diodes, we can use the equivalent circuit shown in Fig. 2-3b, where $R_{S U B}$ and $C_{S U B}$ are the substrate parasitic resistance and capacitance, respectively [2]. By substituting the equivalent circuit of the diodes into the voltage multiplier, shown in Fig. 2-2, we can note that the power dissipation due to substrate losses in the diodes with odd subscript is zero because the voltage that drops across the series of $R_{S U B}$ and $C_{S U B}$ is approximately a DC voltage. Instead, the power dissipation due to substrate losses in the diodes with even subscript is given by

$$
P_{\text {DSUB }} \cong \frac{1}{2} V_{0}^{2} R_{S U B}\left(2 \pi f_{0} C_{S U B}\right)^{2},
$$



Fig. 2-6: Input power vs. Number of stages for an output voltage of 0.5 V , an output power of $5 \mu \mathrm{~W}$ and for three values of $R_{S U B}\left(\omega C_{S U B}\right)^{2}$
under the assumption that $2 \pi f_{0} R_{S U B} C_{S U B} \ll 1$ [2]. Thus, the average input power is obtained by summing up the expression (2-8) previously found and the power dissipated in the diodes due to the substrate losses, as shown below,

$$
P_{I N}=2 N I_{S} V_{i n} B_{1}\left(\frac{V_{0}}{V_{T}}\right) \exp \left(-\frac{V_{U}}{2 N V_{T}}\right)+N P_{D S U B} .
$$

We plotted $P_{I N}(N)$ for different values of $R_{\text {SUB }}\left(2 \pi f_{0} C_{\text {SUB }}\right)^{2}$ in Fig. 2-6. It is possible to note that, for increasing $R_{\text {SUB }}\left(2 \pi f_{0} C_{\text {SUB }}\right)^{2}$, the number of stages corresponding to the minimum input power is shifted to higher values. Once $R_{S U B}$ and $C_{S U B}$ are known, it is possible to find the optimum number of stages, in order to maximize the power efficiency of the voltage multiplier.

### 2.2.2. Input Equivalent Impedance

The equivalent input impedance of the voltage multiplier is constituted by the parallel of a resistance and a capacitance. From our previous considerations, the input capacitance of the voltage multiplier is the sum of the capacitances of all diodes. Since, in a voltage multiplier, diodes conduct for a very small fraction of the period of the input signal, when calculating the mean value of the diode capacitance we can neglect the diffusion capacitance. Using the SPICE model of a p-n junction [5], the depletion layer capacitance $C_{D}$ of a diode is given by

$$
C_{D}=\left\{\begin{array}{l}
\frac{C_{j 0} A}{\left(1-\frac{V d}{V \gamma}\right)^{m}} \quad V_{d}<(F C) V_{\gamma} \\
C_{j 0} A(1-F C)^{-1-m}\left[1-F C(1+m)+m \frac{V_{d}}{V_{\gamma}}\right] V_{d}>(F C) V_{\gamma}
\end{array},\right.
$$

where $C_{j 0}$ is the depletion capacitance at zero bias, $m$ is the grading coefficient, $V_{\gamma}$ is the built-in voltage, $F C$ is the forward bias depletion capacitance coefficient and $A$ is the area of the diode [5]. Since the diode capacitance is a function of the voltage applied to the diode, we can consider the average value of the diode capacitance within the diode voltage swing, as shown below,

$$
\begin{equation*}
\bar{C}_{D}=\frac{1}{2 V_{0}} \int_{-V_{0}-\frac{V_{0}}{2 N}}^{V_{0}-\frac{V_{0}}{2 N}} C_{D}\left(V_{d}\right) d V_{d} . \tag{2-14}
\end{equation*}
$$

Considering the average value of the capacitance of each diode, the equivalent input capacitance, $C_{e q}$, of the voltage multiplier, due to the diode capacitance, is given by

$$
\begin{equation*}
C_{e q}=2 N \bar{C}_{D} . \tag{2-15}
\end{equation*}
$$

The equivalent input resistance, $R_{e q}$, of the voltage multiplier is the resistance calculated from the power consumption with (2-9).

### 2.2.3. Power Matching Network

Since the power at the transponder antenna varies with the distance between the


Fig. 2-7: Power matching network.
reader and the transponder, power matching will be pursued in the condition of minimum power available at the antenna that still ensures correct operation of the transponder. Indeed, in the next section we will show that the transponder continues to work correctly when the power at the antenna increases even if power matching is lost. The power matching LC network is shown in Fig. 2-7, where $L=Q R_{A} / \omega_{0}$, $C=Q /\left(\omega_{0} R_{e q}\right)$ [4], and $Q$ is the quality factor of the LC network, obtained from (2-10). The backscatter modulator is for the moment removed, so that we can optimize power matching. The inductance $L^{\prime}$, shown in Fig. 2-7, is used in order to compensate the equivalent input capacitance $C_{e q}$ of the voltage multiplier ( $C$ and $L^{\prime}$ can be substituted by a single reactance, whose sign depends on their relative value). It is clear that, since $C_{e q}$ is only a time-averaged capacitance, the power matching is good provided that the variations of the input impedance of the voltage multiplier are small with respect to the average values. To verify the previous statement, we can calculate the impedance seen by the transponder's antenna in the worst matching condition, i.e., when the value of the input capacitance of the voltage multiplier is the farthest from its mean value. Referring to Fig. 2-7, the expressions of the real and imaginary part of the impedance, $Z$, seen by the equivalent voltage generator at the antenna, $V_{a n t}$, can be computed as,

$$
\begin{equation*}
\operatorname{Re}\{Z\}=\frac{R_{e q}}{Q^{2}+1}\left[1+\frac{Q^{2}+1}{1+\left(Q \pm Q^{\prime}\right)^{2}}\right] \tag{2-16}
\end{equation*}
$$

$$
\begin{equation*}
\operatorname{Im}\{Z\}=R_{e q}\left[\frac{Q\left(1+\left(Q \pm Q^{\prime}\right)^{2}\right)-\left(Q \pm Q^{\prime}\right)\left(Q^{2}+1\right)}{\left(1+Q^{2}\right)\left(1+\left(Q \pm Q^{\prime}\right)^{2}\right)}\right] \tag{2-17}
\end{equation*}
$$

where $Q^{\prime}=2 \pi f_{0} R_{e q} \Delta C$ and $\Delta C$ is the maximum variation of the input capacitance of the voltage multiplier with respect to its mean value. In order to ensure the correct operation of the power matching network, the real part $Z$ has to be twice the antenna's resistance and its imaginary part has to be zero. Such conditions are fulfilled if $Q^{\prime} \ll Q$. From (2-16) and (2-17) the following condition can be derived:

$$
\Delta C \ll \frac{Q}{2 \pi f_{0} R_{e q}}
$$

To estimate $\Delta C$ we can assume that the mean value of the diode capacitance is its value at the average voltage drop, i.e., $-V_{U} / 2 N$. The maximum value of the diode capacitance is obtained when $V_{d}=V_{0}-V_{U} / 2 N$. Using an $N$-stage voltage multiplier, we can assume that the output voltage is $2 N$ times the amplitude of the input voltage minus $2 N$ times the $V_{\gamma}$ of the diodes. As consequence we can write the amplitude of the input voltage as $V_{0}=V_{U} / 2 N+V_{\gamma}$

In this condition we can express $\Delta C$ as the difference between the maximum and the mean value of the input capacitance, i.e.,

$$
\begin{align*}
& \Delta C=C_{D}\left(V_{\gamma}\right)-C_{D}\left(-V_{U} / 2 N\right)= \\
& =2 N C_{j 0} A\left[\frac{1-F C(1+m)+m}{(1-F C)^{1+m}}-\left(1+\frac{V_{U}}{2 N V_{\gamma}}\right)^{-m}\right] .
\end{align*}
$$

On the one hand, once the output power and voltage are given, by substituting (2-19) into (2-18), we obtain the maximum value of the diode area that enables us to achieve power matching. On the other hand, once the output voltage and the minimum diode area are fixed, as allowed by the CMOS technology used, from (2-18) and (2-19) we can derive the maximum equivalent input resistance, $R_{\text {eqmax }}$, of the voltage multiplier and then, from (2-9) the minimum achievable output power, $P_{\text {OUTmin }}$, for correct power matching. In order to be able to achieve power matching with very low output power, we need to use diodes with a small parasitic capacitance, for a fixed minimum diode area, in order to reduce $\Delta C$.

Let $k$ be the ratio between $P_{\text {OUTmin }}$ and the maximum DC power required at the output of the voltage multiplier. Now we can suppose that for the application we


Fig. 2-8: a) Maximum Operating Range as function of $k$ for the two matching strategies described, b) Power Efficiency of a single stage voltage multiplier as function of the output power for three values of the output voltage.
are interested to, it is sufficient to have an output power $k$ times smaller than $P_{\text {OUTmin }}$. We can consider two options, in order to evaluate the best one for the optimization of the operating range.
A first option (case A) could be to dimension the voltage multiplier and the power matching network for an output power equal to $P_{\text {OUTmin }}$, although a smaller output power would be sufficient; this would lead to a worse power efficiency and then to a reduction of the operating range but it allows us to have a correct power matching. Indeed, in the case of power matching, the input power of the voltage multiplier can be written as,

$$
P_{I N}=\frac{P_{E I R P}}{4 \pi r^{2}} A_{e}
$$

where $A_{e}$ is the effective aperture of the transponder's antenna, $r$ is the distance between reader and transponder, and $P_{\text {EIRP }}$ indicates the power at which an isotropic emitter would have to be supplied to generate the same radiation power of the reader antenna ( $P_{\text {EIRP }}$ is limited by national regulations [7]): a reader antenna
with a gain $G_{T}$ may irradiate a maximum power $P_{\text {EIRP }} / G_{T}$ )
The input power of the voltage multiplier can also be written as in (2-8), considering an output power $P_{L}=P_{\text {OUTmin }}$. Equating (2-20) and (2-8), it is possible to plot the maximum operating range achievable as function of $k$.

A second option (case B) is to dimension the voltage multiplier considering an output power $P_{L}=P_{\text {OUTmin }} / k$; in such a condition, as already said, since the power required at the output of the voltage multiplier is smaller than $P_{\text {OUTmin }}$ (and then $R_{e q}$ is larger than $R_{\text {eqmax }}$ ), power matching could not be achieved. In order to recover power matching, we can put a resistance $R^{*}$, in parallel with the input of the voltage multiplier so that the input resistance of the voltage multiplier becomes smaller than $R_{\text {eqmax }}$. In such a case the input power is given by summing up the input power, $P_{I N}$, of the voltage multiplier, given by (2-8), and the power dissipated by $\mathrm{R}^{*}$, as shown below,

$$
\begin{equation*}
P_{I N}+\frac{V_{0}^{2}}{2 R^{*}}=\frac{P_{E I R P}}{4 \pi r^{2}} A_{e} . \tag{2-21}
\end{equation*}
$$

Substituting (2-8) into (2-21), for $P_{L}=P_{\text {oUT min }} / k$, it is possible to plot the maximum operating range achievable as function of $k$.

Referring to the AMS $0.35 \mu \mathrm{~m}$ IC technology with a minimum diode area equal to $1 \mu \mathrm{~m}^{2}$, and supposing to consider an output voltage of 2 V , an operating frequency of 2.45 GHz and a $Q$-factor of the $L C$ matching network of 10 , the minimum output power, $P_{\text {OUTmin }}$, that allows us to achieve the power matching is $12 \mu \mathrm{~W}$. From Fig. 2-8a, where the maximum operating ranges achievable with each option are plotted as a function of $k$, we can see that option B is to be preferred, even if some power is dissipated in the resistor $R$. Fig. 2-8a also tells us that there is no advantage in reducing the power consumption below $P_{\text {OUTmin }}$, since we cannot reduce the available power required at the antenna. Referring to case B, we plot in Fig. 2-8b the power efficiency of the voltage multiplier as function of the required DC output power. It is possible to note that for small output power, when a resistance is added in parallel with the input of the voltage multiplier, the smaller is the output voltage and the higher is the power efficiency obtained; for larger output power, instead, the larger is the output voltage and the better is the obtained power efficiency.

### 2.2.4. Non-Linear Effects

Because of the non-linear effects of the voltage multiplier, the current $I_{a n t}$ in the antenna also comprises components at frequencies that are integer multiples of the operation frequency. Indeed, the component of the input voltage of the voltage multiplier at the operating frequency $f_{0}$ generates an input current, $I_{i n}$, constituted by


Fig. 2-9: Equivalent circuit of the system antenna-tag to calculate the third harmonic of the antenna's current.
the odd harmonic components of $f_{0}$. In order to calculate the input current of the voltage multiplier, let us consider a single stage; the input current can be obtained by subtracting the currents in the two diodes across which a voltage $V_{d}$, given by (21 ), drops, and considering the diode model without substrate losses. Then, such current has to be multiplied by the number of stages in order to obtain the total input current, $I_{i n}$, of the N -stage voltage multiplier, and its expression is given by,

$$
\begin{equation*}
I_{\text {in }}=N I_{S}\left[\exp \left(\frac{V_{\text {in }}}{V_{T}}\right)-\exp \left(-\frac{V_{\text {in }}}{V_{T}}\right)\right] \exp \left(-\frac{V_{U}}{2 N V_{T}}\right)+2 N \bar{C}_{D} \frac{d}{d t} V_{i n} . \tag{2-22}
\end{equation*}
$$

Exploiting the modified Bessel function series expansion for the two exponential functions that appear in (2-22), the input current, generated by a sinusoidal input voltage at the operation frequency, can be written as shown below,

$$
\begin{align*}
I_{i n}= & 4 N I_{S} \exp \left(-\frac{V_{U}}{2 N V_{T}}\right) \sum_{n=1}^{\infty} B_{2 n-1}\left(\frac{V_{0}}{V_{T}}\right) \cos \left[(2 n-1) \omega_{0} t\right]-.  \tag{2-23}\\
& -2 N \bar{C}_{D} V_{0} \omega_{0} \sin \left(\omega_{0} t\right)
\end{align*}
$$

The amplitude of the fundamental of the current, $I_{\text {lant }}$, in the antenna is given by

$$
\begin{equation*}
\left|I_{\operatorname{lant}}\right|=I_{10} \sqrt{Q^{2}+1} \tag{2-24}
\end{equation*}
$$

where $I_{10}$ is the amplitude of the fundamental of the input current of the voltage
multiplier. Let us consider only the effect of the third harmonic $I_{3 i n}$ of $I_{i n}$. It is clear that the higher harmonics of the input voltage have the same qualitative behaviour, but they are more strongly suppressed by the power matching network, so that we will take into consideration only $I_{3 i n}$. Referring to the equivalent circuit shown in Fig. 2-9, we can calculate the amplitude of third harmonic $I_{3 a n t}$ of the current in the antenna, due to $I_{3 i n}$. From (2-23) we can obtain the amplitude of the third harmonic $I_{\text {3ant }}$ of the current in the antenna as

$$
\begin{equation*}
\left|I_{\text {sant }}\right|=\frac{3 \omega_{0} L^{\prime}\left(Q^{2}+1\right) I_{30}}{\sqrt{\left(R_{e q}-9 \omega_{0} L^{\prime} Q\right)^{2}+9\left(Q R_{e q}+\left(1-8 Q^{2}\right) \omega_{0} L^{\prime}\right)^{2}}} \tag{2-25}
\end{equation*}
$$

where $I_{30}$ is the amplitude of $I_{3 i n}$. Using (2-19) and using for the SPICE parameters $F C, \mathrm{~m}$ and $\mathrm{V}_{\gamma}$, the default values, which are $0.5,0.31$ and 0.69 V , respectively, we can find that for an output voltage close to one volt, as typical in passive RFID systems, $\Delta C \cong C_{e q}$. By substituting $C_{e q}$ to $\Delta C$ in (2-18), since the inductance $L^{\text {, }}$ resonates with $C_{e q}$, we find the following

$$
Q \omega_{0} L^{\prime} \gg R_{e q} .
$$

Using (2-24) and (2-26), we can rewrite (2-25) as follows,

$$
\begin{equation*}
\left|I_{\text {anat }}\right|=\frac{\sqrt{Q^{2}+1}}{\sqrt{64 Q^{4}-7 Q^{2}+1}}\left|I_{\text {lant }}\right| . \tag{2-27}
\end{equation*}
$$

By imposing that the amplitude of the third harmonic of the current in the antenna is fifty times smaller than the amplitude of the fundamental, we obtain that $Q$ has to be larger than 7 . If $Q$ is smaller than 7 , in order to verify the previous condition, we can place a parallel $L^{*} C^{*}$ network in parallel to the input of the voltage multiplier, which resonates at the operating frequency. This network has no effects with respect to the operating frequency but attenuates the harmonics of the current in the antenna. The value of this equivalent capacitance, $C_{n}$, for the n -th harmonic is given by,

$$
\begin{equation*}
C_{n}=\frac{n^{2} C^{*}}{n^{2}-1} . \tag{2-28}
\end{equation*}
$$

Using the parallel resonant $L^{*} C^{*}$ network, the new expression for the amplitude of the third harmonic of the current in the antenna, as function of the amplitude of the


Fig. 2-10: Quality factor of the parallel resonant $L^{*} C^{*}$ network as function of the quality factor of the power matching network.
fundamental, using (2-24), is given by

$$
\left|I_{3}\right|=\frac{\sqrt{Q^{2}+1}}{\sqrt{9\left(Q+\frac{8 Q^{*}}{9}\right)^{2}+\left(1-8 Q^{2}-8 Q Q^{*}\right)^{2}}}\left|I_{1}\right|
$$

where $Q^{*}=\omega_{0} R_{e q} C^{*}$. Imposing that the amplitude of the third harmonic must be at least fifty times smaller than the amplitude of the fundamental, we can derive $Q^{*}$ as function of $Q$. This relationship is shown in Fig. 2-10. Once $Q^{*}$ is chosen, the parallel resonant $L^{*} C^{*}$ network is dimensioned.

### 2.2.5. Matching when Conditions Vary

The dimensioning of the power matching network must be done for the maximum operating range when the power available at the terminals of the transponder's antenna is the minimum one that allows the transponder to operate correctly. Then it is important to analyze what happens when the transponder is moved closer to the reader and the input power of the voltage multiplier increases. As a consequence of such variations, the input equivalent resistance of the voltage
multiplier varies causing mismatch and so a part of the power available at the antenna's terminals is reflected. It is important to verify that the power that comes to the input of the voltage multiplier is sufficient to ensure its correct operation.

In order to keep the DC supply voltage constant, a series voltage regulator is placed at the output of the rectifier. We can assume that the current in the voltage regulator is much smaller than the current provided at its output and that the minimum voltage drop required across the regulator is only few tens of mV , as can be shown in practical implementations [6]. Therefore, we can consider the efficiency of the voltage regulator practically one at the maximum operating range, and for shorter distances we can assume that the input current of the voltage regulator is constant and equal to the DC current $I_{U}$ required by the load. In other words, the output of the rectifier sees the voltage regulator as an ideal DC current generator $I_{U}$. Once the power matching network was dimensioned in the condition of maximum operating range, as previously explained, from (2-5), for the minimum output power required for the correct operation of the transponder, we can obtain the amplitude of the input voltage, $V_{0}$, and so from (2-8) we obtain the input power of the rectifier. Then, once the input power of the voltage multiplier is calculated, we can derive the minimum amplitude of the voltage, $V_{\text {SMIN }}$, at the antenna, that ensures the correct operation of the transponder and its expression is given by,

$$
V_{S M I N}=\sqrt{8 R_{A} P_{I N}} .
$$

As already said, when increasing the voltage at the antenna with respect to its minimum value, given by (2-30), the input resistance of the voltage multiplier varies. In order to calculate the input resistance of the voltage multiplier as function of the voltage at the antenna, from the circuit of Fig. 2-7, we can calculate the input power of the voltage multiplier as function of the voltage at the antenna and of the input resistance of the voltage multiplier. Its expression is given by,

$$
P_{I N}=\frac{R_{e q}\left(Q^{2}+1\right)}{2\left(R_{A}\left(Q^{2}+1\right)+R_{e q}\right)^{2}} V_{S}^{2},
$$

where $V_{S}$ is the amplitude of the voltage at the antenna's terminals. The input power has also the expression shown in (2-8) and using (2-5) to substitute the


Fig. 2-11: Required input power vs. the amplitude of the antenna voltage for an output voltage of 2 V and an output power of $5 \mu \mathrm{~W}$.
exponential function that appears in (2-8), we obtain the equation shown below,

$$
\frac{\sqrt{Q^{2}+1}}{4 I_{S}\left(R_{A}\left(Q^{2}+1\right)+R_{e q}\right)\left(1+\frac{I_{U}}{I_{S}}\right)} V_{S}=\frac{B_{1}\left(\frac{V_{0}}{V_{T}}\right)}{B_{0}\left(\frac{V_{0}}{V_{T}}\right)}
$$

where $V_{0}$ can be derived from (2-31) using (2-9). From (2-32), by numerical iteration, it is possible to obtain the input resistance of the rectifier as a function of $V_{S}$ and $I_{U}$. Once such result is obtained, from (2-31), we can derive the input power of the rectifier as a function of $V_{S}$ and then the amplitude of the input voltage, $V_{0}$, of the voltage multiplier as a function of $V_{S}$. Then from (2-5) we can derive the output voltage of the rectifier as a function of $V_{S}$. Again assuming that the antenna is a dipole ( $R_{A}=72 \mathrm{ohm}$ ), the output voltage, $V_{U}$, is equal to 2 V , the maximum output power, $P_{L}$, is $5 \mu \mathrm{~W}$, we can calculate, in the condition of maximum operating range, the minimum antenna voltage and its value is 77 mV .


Fig. 2-12: Equivalent circuit of the antenna and the load represented by transponder.

We can plot the power, $P_{I N}$, at the input of the voltage multiplier and the power, $P_{I N_{-} A V}$, at the terminals of the antenna, when increasing the antenna voltage with respect to its minimum value, given by (2-30). Both quantities are plotted as a function of $V_{S}$ in Fig. 2-11. Of course, in the condition of maximum operating range, for which power matching is achieved, $P_{I N}=P_{I N_{-} A V}$. When the antenna's voltage increases, part of the power available at the terminals of the antenna, given by the difference between the two curves plotted on Fig. 2-11, is reflected by the power matching network because of the mismatch due to the variation of the input resistance of the voltage multiplier. Anyway, the input power of the voltage multiplier always is larger than the minimum value $P_{\text {MIN }}$ necessary to ensure the correct operation of the transponder.

### 2.3. BACKSCATTER MODULATOR

### 2.3.1. $A S K$ and PSK Backscatter Modulation

In this section, we briefly review ASK and PSK backscatter modulation, in order to identify the most appropriate choice for the task at hand. Indeed, referring to Fig. 2-12 and assuming a minimum scattering antenna, the amplitude of the backscattered power, $P_{B S}$, has the expression shown below,

$$
P_{B S}=\frac{P_{E I R P}}{4 \pi r^{2}} A_{e} \frac{4 R_{A}^{2}}{\left|R_{A}+j X \| R\right|^{2}} .
$$

The impedance seen by the antenna can be represented as a resistance $R$ in parallel with a reactance $X$, as shown in Fig. 2-12. If the antenna can not be
considered as a minimum scattering antenna, in order to calculate the backscattered power, in the second term of $(2-33)$ we have to add another term, which is the backscattered power when the antenna is left open and is independent of the antenna load; so that such term has no effect on the probability of error at the reader.

In the case of ASK modulation, we can assume that the impedance seen by the antenna is real $(X \gg R)$ and is modulated by the data signal between two values $R_{1}$ and $R_{2}$. In order to have equal mismatch in both states, it is sufficient to choose $R_{2}=R_{A}^{2} / R_{1}$; in such condition, in both states, the same power is transferred from the antenna to the load. Assuming $R_{2}>R_{1}$, in order to modulate the resistance seen by the antenna, we can use a switch, driven by the data signal, to connect a resistance $R_{M O D}$ in parallel with the input resistance $R_{2}$, of the transponder, in such a way that $R_{1}=R_{2} \| R_{\text {MOD }}$. When $R_{M O D}$ is not connected, the antenna sees a resistance $R_{2}$ and all the power $P_{I N 2}$ transferred from the antenna to the load can be used to supply the transponder; when the resistance $R_{M O D}$ is connected, the antenna sees a resistance $R_{l}$ and only a fraction $P_{I N l}$ of the power transferred from the antenna to the load can be used to supply the transponder, while the remaining part is dissipated on the resistance $R_{M O D}$. As a consequence, the power transferred from the antenna to the load remains constant in both states, but $P_{I N I}$ and $P_{I N 2}$ are different, and are given by

$$
\begin{equation*}
P_{I N 2}=P_{A V} \frac{4 R_{A} R_{2}}{\left(R_{A}+R_{2}\right)^{2}}, P_{I N 1}=P_{A V} \frac{4 R_{A} R_{1}}{\left(R_{A}+R_{1}\right)^{2}} \frac{R_{M O D}}{R_{M O D}+R_{2}}, \tag{2-34}
\end{equation*}
$$

where $P_{A V} \equiv P_{E I R P} A_{e} /\left(4 \pi r^{2}\right)$. It is possible to demonstrate that, except for the solution $R_{I}=R_{2}$ (i.e., $R_{M O D} \rightarrow \infty$ ), which would imply the absence of modulation, the equation $P_{I N 2}=P_{I N I}$ has no solution. This means that, when modulating, the tag can not be supplied with constant power. Using the condition $R_{2}=R_{A}^{2} / R_{1}$, the backscattered powers, $P_{B S I}$ and $P_{B S 2}$, when the impedance seen by the antenna are $R_{I}$ and $R_{2}$, respectively, read

$$
\begin{equation*}
P_{B S 2}=P_{A V} \frac{4 R_{A}^{2}}{\left(R_{A}+R_{2}\right)^{2}}, P_{B S 1}=\frac{R_{2}^{2}}{R_{A}^{2}} P_{B S 2} . \tag{2-35}
\end{equation*}
$$

We shall see later that the probability of error at the receiver depends on a unique quantity, an effective (or "modulated") power $P_{U}$, which is the power of the signal obtained from the demodulation of the voltage directly applied to the radiation resistance of the antenna, again dissipated - for convenience - on a resistance $R_{A}$.

The larger the value of $P_{U}$, the smaller the probability of error. In order to compare the performances of the ASK and PSK backscatter modulation, we can calculate $P \mathrm{U}$ for the ASK modulation, assuming to use a coherent receiver, perfectly equal to the one that will be used for PSK modulation (to be discussed later) except for the detector threshold. Without loss of generality, we can refer to unipolar RZ coding for both ASK and PSK modulation: if we transmit an alternating sequence of symbols ' 0 ' and ' 1 ', the demodulated signal is a square wave whose amplitude varies between $\left(V_{I}-V_{2}\right) / 2$ and zero. We therefore obtain $P_{U}=\left(V_{1}-V_{2}\right)^{2} / 8 R_{A}$. As a consequence, if the reader's antenna is perfectly matched, $P_{U}$ has the expression shown below,

$$
\begin{equation*}
P_{U}=\left(\sqrt{8 R_{A} P_{B S 1}}-\sqrt{8 R_{A} P_{B S 2}}\right)^{2} /\left(8 R_{A}\right)=P_{B S 2}\left(1-\frac{R_{2}}{R_{A}}\right)^{2} . \tag{2-36}
\end{equation*}
$$

In the case of PSK modulation, we must have $R=R_{A}$ so that the transponder is close to the matching condition, while $X$ is modulated with the data signal. In fact, referring to Fig. 2-12, the phase $\theta$ of the backscattered signal, proportional to the voltage $V_{B S}$ on the radiation resistance $R_{A}$ reads

$$
\theta=\angle V_{B S}=-a \tan \left\{\frac{R_{A} X}{R_{A}^{2}+2 X^{2}}\right\}
$$

If $X$ is modulated symmetrically with respect to zero, also is $\theta$, which implies that the power $P_{B S}$ reflected by the antenna and the power $P_{I N}$ transferred to the transponder remain constant during modulation and are given by

$$
\begin{gather*}
P_{B S}=P_{A V} \frac{4\left(R_{A}^{2}+X^{2}\right)}{R_{A}^{2}+4 X^{2}},  \tag{2-38}\\
P_{I N}=P_{A V} \frac{4 X^{2}}{R_{A}^{2}+4 X^{2}} . \tag{2-39}
\end{gather*}
$$

At the receiver, in order to demodulate the PSK signal we have to use a coherent receiver. As it will be clearer in section IV, supposing to transmit an alternating sequence of symbols ' 0 ' and ' 1 ', the demodulated signal is a square wave whose amplitude varies between two states, which are zero and $A \sin (\theta)$ where $A$ is the amplitude of the signal received at the reader, determined by the backscattered


Fig. 2-13: Effective power $P_{U}$ at the reader and power $P_{I N}$ transferred to the transponder for the ASK and PSK modulations, as a function of the reflection coefficient.
power, and $\theta$ is the modulation depth, given by (2-37). Following the reasoning used for the ASK modulation, the power $P_{U}$ is given by,

$$
P_{U}=4 P_{B S} \sin ^{2}(\theta)
$$

In order to compare the performance of ASK and PSK modulation [2], we can plot in Fig. 2-13, $P_{U} / 4 P_{A V}$, for both modulations, as function of the reflection coefficient $\rho$, defined as $\rho \equiv\left|\left(Z_{I N}-R_{A}\right) /\left(Z_{I N}+R_{A}\right)\right|$, where $Z_{I N}$ is the input impedance of the transponder $Z_{I N}=R \| j X$. Fig. 2-13 also shows $P_{I N} / P_{A V}$, considering in the case of the ASK modulation the average of $P_{I N I}$ and $P_{I N 2}$, assuming that the two states have the same likelihood. It is possible to note that for a given $\rho$ the PSK backscatter modulation ensures a larger $P_{I N}$ but the ASK backscatter modulation ensures a larger $P_{U}$. Since, as we will see later, the most critical aspect limiting the operating range is represented by the input power of the transponder, the PSK backscatter
modulation is to be preferred.
Furthermore, PSK backscatter modulation allows us to provide a constant power supply to the transponder during modulation.
Let us note for both ASK and PSK modulations, a coherent receiver would be required at the reader to filter away the unmodulated carrier backscattered by unwanted obstacles (clutter) which is orders of magnitude larger than the backscattered modulated signal to which it adds at the receiver. In such a way, the clutter can be separated by simply multiplying the received signal by the locally synthesized carrier. The alternative option to filter away the clutter would be represented by the use of subcarrier backscatter modulation, that would allow to use a simpler incoherent receiver, at least for the ASK modulation. However, such solution would have the serious drawback of implying a very large increase of the modulator switching frequency, and therefore of the transponder power consumption, leading to a significant reduction of the operating range.

From now on we will take in consideration the PSK backscatter modulation.

### 2.3.2. PSK Backscatter Modulator

Most of PSK backscatter modulators [2], [8], [9], independently of their implementation, allow modulation of their output capacitance with the input signal. Referring to Fig. 2-1, $L$ is used to make the imaginary part of the admittance seen at the output of the modulator symmetric with respect to zero. So $L$ has to be chosen to resonate with the mean value of the capacitance seen from the output of the modulator when the input signal varying and is therefore,

$$
\begin{equation*}
L=\frac{1}{\left(2 \pi f_{0}\right)^{2}\left(\frac{C_{V 1}+C_{V 2}}{2}\right)}, \tag{2-41}
\end{equation*}
$$

where $f_{0}$ is the operating frequency and $C_{V I}, C_{V 2}$ are the output capacitances of the modulator when the input signal is high and low, respectively. The reactance $X$ is therefore given by

$$
\begin{equation*}
X=\frac{2}{\left(2 \pi f_{0}\right)\left|C_{V_{1}}-C_{V_{2}}\right|} \tag{2-42}
\end{equation*}
$$

An integrated circuit implementation of a PSK backscatter modulator for passive radio frequency identification (RFID) transponders is proposed in the following. Such modulator offers a significant reduction of the power consumption with respect to other schemes already presented in the literature. Furthermore, the


Fig. 2-14: Scheme of the proposed PSK backscatter modulator.
topology of the proposed modulator allows us to control its output resistance so that only a negligible fraction of the active power at the antenna goes to the modulator. The most important requirements for such modulators are low power consumption, since they are used in passive systems, and small area occupation on the chip to maintain their cost as low as possible.

## Circuit description

The scheme of the proposed modulator is shown in Fig. 2-14. Transistor $\mathrm{M}_{1}$ is a MOS varactor that operates in inversion or in cutoff, depending on the input signal, causing the variation of the capacitance seen at the output of the modulator. Transistor $\mathrm{M}_{2}$, instead, does not affect the output capacitance, since it has a small width with respect to $\mathrm{M}_{1}$, but determines the resistance at the output of the modulator, which is, essentially, its drain-source resistance. As a consequence, the channel length of $\mathrm{M}_{2}$ has to be large enough so that the output resistance of the modulator is much larger than the antenna resistance. Such choice ensures that only a negligible fraction of the power at the antenna goes to the modulator, as required for the correct operation of the transponder. CoUT is the capacitance seen at the output of the modulator and is due to the interconnections, the antenna and the input capacitance of the other stages which the modulator is connected to. As it will be clearer later, the capacitance $C_{I N}$ has to be larger than the gate-source and gatedrain capacitance of $\mathrm{M}_{1}$, in order not to degrade the variation of the output capacitance of the modulator. Once the value of $C_{I N}$ is chosen, as previously described, the two transistors $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$ of the inverter have to be dimensioned to fix the switching time of the varactor so that the channel bandwidth occupation of the backscattered signal complies with the requirements set by the standard.


Fig. 2-15: Schemes of the comparison PSK backscatter modulators: a) from [1]; b) modified from [2].

Referring to Fig. 2-14 and to the BSIM3v3 transistor model [5], the output capacitance, $C_{V}$, of the modulator is given by,

$$
C_{V}=C_{D B}+C_{J D}+C_{J S}+C_{D M 2}+\left(C_{G S}+C_{G D}\right) \|\left(C_{I N}+C_{G B}+C_{I N V}\right),
$$

where $C_{X Y}$ is the capacitance seen between the terminals X and Y of $\mathrm{M}_{1}, C_{J D}$ and $C_{J S}$ are the drain-bulk and source-bulk junction capacitances, $C_{D M 2}$ is the capacitance seen from the drain of $\mathrm{M}_{2}$ to ground, $C_{I N V}$ is the capacitance seen from the output of the inverter formed by $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$. When the input signal varies, only the capacitances $C_{G S}, C_{G D}, C_{D B}$ and $C_{G B}$ vary, while the other capacitances remain constant. Thus, it is clear that, if $C_{I N}$ is not sufficiently larger than $C_{G S}$ and $C_{G D}$, the variation of the output capacitance is drastically reduced, since it predominantly depends upon the gate-source and gate-drain capacitances. In such condition, the variation $\Delta C$ of the output capacitance of the modulator can be well approximated as

$$
\begin{equation*}
\Delta C \cong C_{D B}+\left(C_{G S}+C_{G D}\right)\left\|C_{I N}-\left(C_{G S o v}+C_{G D o v}\right)\right\| C_{I N} \tag{2-44}
\end{equation*}
$$

where $C_{G D o v}$ and $C_{G S o v}$ are the overlap gate-drain and gate-source capacitances of $M_{1}$. Referring to Fig. 2-14, in each period of the input signal, since the drain and the source of $\mathrm{M}_{1}$ are always pulled down to ground, we do not need to charge and discharge the output capacitance and the equivalent drain-ground and sourceground capacitances. Thus, the power $P_{M O D}$ dissipated by the modulator is given by,

$$
\begin{equation*}
P_{M O D} \cong\left(C_{G S}+C_{G D}+C_{I N}\right) f V_{D D}^{2} \tag{2-45}
\end{equation*}
$$



Fig. 2-16: Area occupation on the chip vs. modulation of the output capacitance $\Delta C$.

From (2-44) it is clear that the choice $C_{I N} \gg C_{G S}+C_{G D}$ maximises the modulation depth but leads to high power consumption when modulating. Since in passive RFID systems it is very important to maintain the power consumption as small as possible, the value of the capacitor $C_{I N}$ has to be chosen as to minimize the power consumption for a given modulation depth. Since all the capacitances of $\mathrm{M}_{1}$ are proportional to the channel width, $W$, of $\mathrm{M}_{1}$, therefore from (2-44), for a given modulation depth, we can derive $W\left(C_{I N}\right)$. Substituting such relation in (2-45) and deriving the power with respect to $C_{I N}$, we can find the value of $C_{I N}$ that allow us to minimize the power consumption and that ensures the modulation depth previously fixed.

## Comparison with other topologies

Now, we can compare the performances of the proposed modulator with that of two other schemes appeared in the literature, on the basis of an implementation with a typical $0.35 \mu \mathrm{~m}$ CMOS process from AMS, and a supply voltage of 0.6 V , compatible with subthreshold operation of the digital section. In the modulator shown in Fig. 2-15a [1], the capacitor $C_{M O D}$ is dimensioned to obtain the desired variation $\Delta C$ of the output capacitance. The great disadvantage of such scheme is


Fig. 2-17: Power consumption of modulators shown in Fig. 2-14 and Fig. 2-15 vs. the variation of the output capacitance for an operating frequency of 2.45 GHz and 868 MHz .
the area occupation on the chip, especially when a very low supply voltage is used, as can happen in passive RFID systems. From Fig. 2-16 it is possible to verify that the area occupation on the chip of the modulator, shown in Fig. 2-15a, is much larger than that one of the proposed modulator. The modulator shown in Fig. 2-15b is a modified version of that proposed in [2]: here the modulation is obtained by exploiting the output capacitance of the inverter formed by $\mathrm{M}_{1}$ and $\mathrm{M}_{4}$ rather than the capacitance of a varicap placed between the input and the output of the inverter, as done in [2]. In such a way, although the operation principle is not changed, we have a reduced number of components (two inverters and one of the two coupling capacitors are removed) and therefore a reduced power consumption and area occupation for the same modulation depth. Anyway, also with such improvement, the modulator shown in Fig. 2-15b occupies an area on the chip between two and three times larger than our modulator, as can be seen in Fig. 2-16. In such modulator transistors $\mathrm{M}_{1}$ and $\mathrm{M}_{4}$ determine the output capacitance while $\mathrm{M}_{2}$ and $\mathrm{M}_{3}$ allow us to increase the output resistance of the modulator. In order to obtain a given modulation of the output capacitance and minimum area occupation, $\mathrm{M}_{1}$ must have minimum width while $\mathrm{M}_{4}$ is dimensioned to obtain the desired $\Delta C$. Transistors $\mathrm{M}_{2}$ and $\mathrm{M}_{3}$ have minimum width, in order to obtain a large output resistance in both states. The great disadvantage is that the capacitance seen at the output of the
modulator has to be charged and discharged in each period of the input signal; as a consequence the power consumption for a given variation of the output capacitance of the modulator shown in Fig. 2-15b is much larger than that of the proposed modulator, as can be seen in Fig. 2-17. Similar conclusions are drawn also from comparison at higher supply voltages.

The choice of $\Delta C$ must be done in order to maximize the operating range, achieving a trade off between minimum probability of error at the reader and maximum power transfer to the transponder, as will be shown later. Analysis show that a range of 3.4 m can be achieved at 2.45 GHz with $\Delta C=130 \mathrm{fF}$ and a range of 9.5 m can be achieved at 868 MHz with $\Delta C=174 \mathrm{fF}$, for a data rate of 40 kbps and a DC power consumption of the transponder of one microwatt.

### 2.4. Modulation Depth and Maximum Operating Range

In this section we will describe the criteria for choosing the modulation depth in order to maximize the operating range of the tag-reader system. When modulating the impedance seen by the transponder's antenna, it is necessary to ensure that the power at the input of the voltage multiplier is larger than the minimum required for its correct operation and the probability of error at the receiver is smaller than a given value, required for the correct receiving. A trade off has to be found between the two conditions in order to maximize the operating range.

### 2.4.1. Transponder input power

Referring to Fig. 2-12, the power $\mathrm{P}_{I N}$ transferred from the antenna to the input of the transponder is given by (2-39). In order to ensure the correct operation of the transponder $\mathrm{P}_{I N}$ must be larger than the minimum power required for the transponder operation, consisting of the sum of the power, $\mathrm{P}_{\text {MOD }}$, dissipated by the modulator and the power, $\mathrm{P}_{\text {DIG }}$, dissipated by the digital section, divided by the efficiency, $\eta$, previously calculated. As consequence, the following relation has to be fulfilled:

$$
\frac{P_{E I R P}}{4 \pi r^{2}} A_{e} \frac{4 X^{2}}{R_{A}^{2}+4 X^{2}}>\frac{1}{\eta}\left(P_{M O D}+P_{D I G}\right) .
$$

From (2-46) it is evident that the larger the power efficiency of the transponder and the larger is the range of values of $X$ that satisfy (2-46). Equation (2-46) gives a first condition that has to be fulfilled by $X$, in order to ensure the correct operation of the tag-reader system. The other condition will be given by the probability of error at the receiver.

### 2.4.2. Probability of Error at the Reader

## Received Signal at the Reader's Antenna

As already said, the signal received by the reader's antenna has two components: a PSK backscattered signal and an un-modulated carrier. The PSK backscattered signal, $\mathrm{x}_{B S}(\mathrm{t})$, coming from the transponder's antenna has a carrier frequency equal to the operating frequency of the RFID system and a phase, which belongs to a group of two values symmetric with respect to zero and with an absolute value, $\theta$, given by (2-37). The amplitude of such signal can be calculated from the amplitude of the power backscattered by the transponder's antenna, taking in consideration the free space attenuation. Assuming that the reader's antenna is perfectly matched with the reader, the amplitude, A , of the signal, $\mathrm{x}_{B S}(\mathrm{t})$, is given by the expression shown below,

$$
A=\frac{A_{e}}{2 \pi r^{2}} \sqrt{\frac{2 R_{A}\left(R_{A}^{2}+X^{2}\right)}{R_{A}^{2}+4 X^{2}} P_{E I R P}} .
$$

So $\mathrm{x}_{B S}(\mathrm{t})$ reads

$$
x_{B S}(t)=\sum_{i=0}^{\infty} \operatorname{Arect}\left(\frac{t-T / 4-i T / 2}{T / 2}\right) \cos \left(2 \pi f_{0} t+\alpha_{i}\right),
$$

where $\alpha_{i} \in\{-\theta, \theta\}, 1 / T$ is the data-rate and $\operatorname{rect}(t / T)$ is a square impulse with an amplitude equal to 1 , a duration $T$ and it is centered at $t=0$. The un-modulated carrier, $\mathrm{x}_{U M}(\mathrm{t})$, that comes back to the reader's antenna, is a sinusoidal voltage with an amplitude, $B$, and a phase, $\beta$, completely unknown:

$$
x_{U M}(t)=B \cos \left(2 \pi f_{0} t+\beta\right) .
$$

The complete received signal, $x(t)$, at the reader's antenna is obtained by summing up the two components. Since the un-modulated carrier $\mathrm{x}_{U M}(\mathrm{t})$ has a power level much larger than the PSK backscattered signal, in order to generate the local carrier required to down convert the received signal, it is not possible to use a PLL because it would reproduce the un-modulated carrier. A possible solution is to transmit, at the beginning of each data package, a preamble, which is, for example, a sequence of symbols ' 1 ', known to the reader and to use an adaptive system that varies the phase of the locally generated carrier until the signal at the output of the filter has the correct amplitude. In this situation the local oscillator is synchronized with the


Fig. 2-18: PSK Receiver Architecture (In the insert: elementary impulse when transmitting ' 0 ' and ' 1 ').
received PSK signal and generates a sinusoidal voltage at the operating frequency with a phase equal to the mean value of the phase of the received PSK signal. Assuming to use an unipolar RZ coding, as will be described later, if we transmit, as already said, a sequence of symbols ' 1 ', the carrier generated locally has a zero phase; using an oscillator with quadrature outputs, it is possible to generate the oscillations for the two branches of the receiver. The receiver has to be able to remove the un-modulated signal, which would introduce an error in the demodulation due to its random phase.

## Receiver Architecture

The scheme of the PSK receiver is shown in Fig. 2-18 [4]. We consider an AWGN (Additive White Gaussian Noise) input noise, $w(t)$. Since the phase of the backscattered signal, $\mathrm{x}_{B S}(\mathrm{t})$, belongs to a group with two values symmetrical with respect to zero, the lower branch of the receiver has no effect since the cosinusoidal function is an even function and cannot detect the phase variation. By multiplying the input signal, $x(t)$, by the locally generated carrier, in quadrature with the received signal and considering only the low frequency component, which will be the only one to survive after filtering, we obtain that the signal, $y(t)$, at the input of the filter, $\mathrm{g}_{R}(t)$, is given by,

$$
\begin{equation*}
y(t)=\sum_{i=0}^{\infty} \operatorname{Arect}\left(\frac{t-T / 4-i T / 2}{T / 2}\right) \sin \left(\alpha_{i}\right)+B \sin (\beta) \tag{2-50}
\end{equation*}
$$

The receiver must be able to remove the un-modulated signal, which would introduce an error in the demodulation due to the random phase. To this aim, the baseband receiver filter must have a zero-average impulse response, to block the un-modulated carrier. Since the impulse response of the filter has to be equal to the elementary impulse of one of the two symbols, the previous condition imposes an appropriate choice of the data coding, in which at least a symbol has zero mean value. The most common codes, such as Manchester, unipolar RZ coding, DBP, Miller coding, verify such condition [1]. Now, in order to obtain numerical results, we can choose one of the previous coding schemes and calculate the probability of error after choosing a proper receiving filter. As an example, without loss of generality, we consider a unipolar RZ coding: in such a case, the elementary impulse, $p_{l}(t)$, when transmitting a bit ' 1 ', is obtained by transmitting a symbol $\theta$ and a symbol $-\theta$, in sequence. Instead, the elementary impulse, $p_{0}(t)$, when transmitting a bit ' 0 ', is obtained by transmitting two symbols $-\theta$, in sequence. The two elementary impulses are shown in the insert of Fig. 2-18. Now we choose the impulse response of the filter, $g_{R}(t)$, with the same waveform of the elementary impulse associated to the transmission of ' 1 '. Since the filter, $g_{R}(t)$, amplifies or attenuates the signal and the noise at the same manner, we can suppose that the amplitude of the impulse response of the filter is equal to $1 / T$. In this condition the frequency response, $G_{R}(f)$, of the filter is given by,

$$
\begin{equation*}
G_{R}(f)=\frac{1}{2} \sin c\left(\frac{f T}{2}\right) \exp \left(-j \omega \frac{T}{4}\right)\left[\exp \left(-j \omega \frac{T}{2}\right)-1\right] \tag{2-51}
\end{equation*}
$$

where $\sin c(x)=\sin (\pi x) /(\pi x)$. Once fixed the impulse response of the filter we can derive the expression of the signal $z(t)$, at the output of the filter, when transmitting ' 1 ' and ' 0 '. The expressions of the signals, $z_{l}(t)$ and $z_{0}(t)$, at the output of the filter when transmitting ' 1 ' and ' 0 ', respectively are given by,

$$
\begin{equation*}
z_{1}(t)=p_{1}(t) \otimes g_{R}(t), z_{0}(t)=p_{0}(t) \otimes g_{R}(t) . \tag{2-52}
\end{equation*}
$$

By sampling the signal, at the output of the filter at time instants $t_{k}$ multiple of $T$, the signal at the input of the detector is zero, when transmitting ' 0 ', and $A \sin (\theta)$, when transmitting ' 1 '. Since the un-modulated carrier provides a constant component at the input of the filter, exactly as it occurs when transmitting ' 0 ', after
the convolution and the sampling it gives a zero signal. Supposing that the transmissions of ' 0 ' and ' 1 ' have the same likelihood, the detector's threshold can be chosen in the middle between $z_{l}\left(t_{k}\right)$ and $z_{0}\left(t_{k}\right)$, according to the MAP criterion [4]. Supposing that the noise has a zero mean value and a standard deviation $\sigma$, the probability of error, $P_{e 0}$, when transmitting ' 0 ' and the probability of error, $P_{e l}$, when transmitting ' 1 ' are given by,

$$
P_{e 0}=P_{e 1}=\operatorname{erf}\left(\frac{A / 2 \sin (\theta)}{\sigma}\right)=\operatorname{erf}\left(\frac{\sqrt{R_{A} P_{U}}}{\sqrt{2} \sigma}\right)
$$

where $\operatorname{erf}(x)$ is the error function. Since the two symbols ' 0 ' and ' 1 ' have the same likelihood the total probability of error is equal to each probability of error when transmitting ' 0 ' and ' 1 '.

Following the same procedure in the case of a Manchester coding scheme and choosing the receiving filter in the same manner, one obtains, after the sampling, for the two symbols the two levels $\pm A \sin (\theta)$, and then a smaller probability of error.

In the case of ASK backscatter modulation, using the receiver shown in Fig. 2-18 and the unipolar RZ coding, we would obtain the following expression for the total error probability $P_{e}$

$$
P_{e}=\operatorname{erf}\left(\frac{V_{1}-V_{2}}{4 \sigma}\right)=\operatorname{erf}\left(\frac{\sqrt{R_{A} P_{U}}}{\sqrt{2} \sigma}\right) .
$$

As already said and as evident from (2-53) and (2-54), the parameter $P_{U}$ calculated in section III is the only parameter of the backscatter modulator affecting $P_{e}$.

## Noise Spectral Density

We can suppose that the noise at the input of the receiver is due to the thermal noise of the antenna, followed by an amplifier with a noise figure, $F$, and with a gain sufficiently high to allow us to neglect the noise figure of the following stages. The thermal noise of the antenna and the noise of the first stage, the expression of the noise spectral density, $N_{\text {oth }}$, is given by,

$$
\begin{equation*}
N_{0 t h}=2 F k_{B} T R_{A}, \tag{2-55}
\end{equation*}
$$

where $k_{B}$ is the Boltzmann constant and $T$ is the absolute temperature. If the input noise is a white Gaussian noise with a power spectral density equal to $N_{0 t h}$, its
power spectral density after filtering is obtained by multiplying the power spectral density at the input by the square absolute value of the frequency response, $G_{R}(f)$, of the filter. As consequence, the standard deviation $\sigma$ of the noise at the output of the filter is obtained by integrating the power spectral density as shown below,

$$
\begin{equation*}
\sigma^{2}=N_{0 t h} \int_{-\infty}^{+\infty}\left|G_{R}(f)\right|^{2} d f \tag{2-56}
\end{equation*}
$$

We can now consider the phase noise due to the local oscillator. The carrier generated by the frequency synthesizer is not perfectly monochromatic, due to the phase noise. The power spectrum density of the phase noise goes down, initially, as $1 / f^{3}$ due to the flicker noise of the devices in the oscillator, then as $1 / f^{2}$ due to the thermal noise of the devices in the oscillator, with a typical corner frequency of some hundreds of kHz [10].

Since the reader uses the same oscillator to generate the transmitted carrier and the local oscillation, the phase noise that affects the received signal and the local oscillation, used to downconvert the received signal, are generated by the same oscillator but in different time instants with a certain delay $\Delta T$. As a consequence, after the downconversion the low frequency component is given by,

$$
\begin{equation*}
V(t)=\left.V_{I N}(t) V_{O L}(t)\right|_{\text {Low }}=A \sin [\theta+\varphi(t)-\varphi(t-\Delta T)] . \tag{2-57}
\end{equation*}
$$

Now we have to study the variance of $\varphi(t)-\varphi(t-\Delta T)$. Its power spectral density $S(f)$ is given by,

$$
S(f)=S_{\varphi}(f) \left\lvert\, 1-\exp (-j \omega \Delta T)^{2}=4 S_{\varphi}(f) \sin ^{2}\left(\frac{\omega}{2} \Delta T\right)\right.
$$

In general we can write the power spectral density $S_{\varphi}(f)$ of the phase noise $\varphi(t)$ as sum of a $1 / f^{3}$ component and a $1 / f^{2}$ component, imposing that the two components are equal at the corner frequency $f_{c}$. The expression of $S_{\varphi}(f)$ is shown below,

$$
S_{\varphi}(f)=\frac{A}{f^{2}}+\frac{A f_{c}}{f^{3}} .
$$

By substituting the two components of $S_{\varphi}(f)$ in (2-58) and integrating in
frequency, we can calculate the variance $\sigma_{w}^{2}$ associated to the $1 / f^{2}$ component of the power spectral density of the phase noise and the variance $\sigma_{1 / f}^{2}$ associated to the $1 / f^{3}$ component of the power spectral density of the phase noise. In the case of open-loop VCO, from [11] we can derive $\sigma_{w}^{2}$, whose expression is given by,

$$
\begin{equation*}
\sigma_{w}^{2}=4 \pi^{2} A \Delta T \quad\left[\mathrm{rad}^{2}\right] \tag{2-60}
\end{equation*}
$$

As one can expect, the power of the jitter associated to the $1 / f^{2}$ component is proportional to the delay. In the case of open-loop VCO, an approximate expression of the power of the jitter associated to the $1 / f^{3}$ component is given in [12] and its expression is,

$$
\begin{equation*}
\sigma_{1 / f}^{2}=4 \pi^{2} \alpha^{2} A f_{c} \Delta T^{2}\left[\mathrm{rad}^{2}\right] \tag{2-61}
\end{equation*}
$$

where $\alpha$ is a dimensionless parameter typically close to 5 [12]. Now we can calculate the variance of $\varphi(t)-\varphi(t-\Delta T)$ by summing up the $1 / f^{2}$ and the $1 / f^{3}$ components of the power spectral density of the phase noise. Now we can define $\varphi$ as the maximum value of $\varphi(t)-\varphi(t-\Delta T)$, given by $\varphi=\sqrt{2\left(\sigma_{w}^{2}+\sigma_{1 / f}^{2}\right)}$. In the case of RFID systems, since the operating range is few meters (especially for the systems operating at 2.45 GHz ), the delay is small with respect to $1 / f_{c}$ and then the variance of the jitter is mainly due to the $1 / f^{2}$ component. Such a result was obtained by considering an open loop VCO. In the case of a closed loop PLL, the power spectral density of the phase noise is filtered by the transfer function of the PLL, and the variance is even smaller, so that the open loop VCO represents the worst case.

The phase jitter causes a shift of the constellation of symbols, leading to an error in the detection. Indeed, in presence of a phase jitter, $\varphi$, the level associated to the transmission of a symbol ' 0 ' remains zero and the level associated to the transmission of a symbol ' 1 ' becomes $[A \sin (\theta) \cos (\varphi)]$ rather than $A \sin (\theta)$. According to the MAP criterion, since the two symbols have the same likelihood, the threshold of the detector is chosen in the middle of the two levels in absence of phase noise and then the probability of error when transmitting ' 0 ' is given by (2-53), while the probability of error when transmitting ' 1 ' is given by,

$$
\begin{equation*}
P_{e l}=e r f\left(\frac{A \sin (\theta)(2 \cos (\varphi)-1)}{2 \sigma}\right) \tag{2-62}
\end{equation*}
$$

The total probability of error is given by the mean value of the probabilities of error when transmitting ' 0 ' and ' 1 '. In order to ensure reasonable receiver performance, the total probability of error has to be smaller than a given probability of error, $P_{e}{ }^{*}$, as shown in the expression below,

$$
\begin{equation*}
P_{e}=\frac{1}{2}\left\{e r f\left(\frac{A \sin (\theta)(2 \cos (\varphi)-1)}{2 \sigma}\right)+e r f\left(\frac{A \sin (\theta)}{2 \sigma}\right)\right\}<P_{e}^{*} . \tag{2-63}
\end{equation*}
$$

Since $A$ and $\theta$ are functions of $X$, (2-63) gives us another condition to find the values of $X$ that ensures acceptable operation of the transponder-reader system.
Considering that the RFID system has an operating frequency of 2.45 GHz or 868 MHz , the maximum allowed $P_{\text {EIRP }}$ for the reader is 500 mW according to European regulations [7]. We also assume that the transponder antenna is a $\lambda / 2$ dipole, which has a radiation resistance of $72 \Omega$ and an effective aperture of $0.13 \lambda^{2}$ [1]. Furthermore, using a single stage voltage multiplier, the power efficiency is about $15 \%$ [13]. Then we assume that the supply voltage is 0.6 V , and the power $P_{D I G}$ is equal to $1 \mu \mathrm{~W}$. We also assume that $T=300 \mathrm{~K}, R_{A}=72$ ohm and $F=5 \mathrm{~dB}, N_{\text {oth }}$ is equal to $1.88 \mathrm{nV}^{2} / \mathrm{Hz}$ and $P_{e}{ }^{*}=10^{-3}$. In such conditions we can plot the values of $X$ that satisfy (43) and (60), as function of the distance $r$ between reader and transponder. The diagram is shown in Fig. 2-19 for different data rates: pairs of $X$ and $r$ that satisfy (2-46) and (2-63) lie between the two curves for a fixed data-rate. In particular, the values of $X$ that satisfy (2-46) have, as lower boundary, the continuous curve corresponding to the considered datarate; On the other hand, the values of $X$ that satisfy (2-63) have, as upper boundary, the dashed curve corresponding the considered datarate. From Fig. 2-19, it is clear that by reducing the data-rate, the range for a fixed $X$ becomes larger because the power required by the transponder, for the modulation and the digital section, decreases and because the bandwidth of the receiving filter becomes smaller leading to a smaller noise power at its output and then to a smaller error probability.

From Fig. 2-19, it is possible to note that by a proper choice of the modulation depth, the operating range, for a passive RFID system, which works in the microwave range, is larger than 4 meters, while for a passive RFID system, which works in the UHF frequency range is larger than 11 meters. It is also possible to find the maximum data-rate, once the operating range is fixed. For example we canchoose an operating range of 4 meters and we can plot the two curves that satisfy (2-46) and (2-63) as function of the datarate, in order to obtain the value


Fig. 2-19: $X$ values that satisfy the $(2-46)$ and (2-63) as function of the distance between reader and transponder for three different data rates for an operating frequency of 868 MHz (top) and 2.45 GHz (bottom).


Fig. 2-20: $X$ values that satisfy the $(2-46)$ and $(2-63)$ as function of the frequency of the data signal: a) for an operative range of 3.4 meters and an operating frequency of $2.45 \mathrm{GHz}, \mathrm{b}$ ) for an operative range of 9.5 meters and an operating frequency of 868 MHz .
of $X$ that allow us to maximize the data-rate for the chosen operating range. Such a diagram is shown in Fig. 2-20, for an operating frequency of 2.45 GHz and 868 MHz : the region with acceptable values of $X$, is the one comprised between the two curves. It is possible to note that, for an operating range of 3.4 meters and for an operating frequency of 2.45 GHz , the maximum data rate is about 17 Kbps ; while, for an operative range of 9.5 meters and for an operating frequency of 868 MHz , the maximum frequency of the data signal is about 70 kbps . If US regulations are considered, which provide a maximum allowed $P_{E R P}$ of 4 W , the same considerations lead to a maximum operating range of about 11 m at 2.45 GHz and 29 m at 916 MHz , for a data-rate of few tens of kbps.

### 2.5. References

[1] K. Finkenzeller, RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, $2^{\text {nd }}$ ed, Wiley and Sons, 1999, pp. 117-126, pp. 143-148 and pp. 183-186.
[2] U. Karthaus, M. Fischer, "Fully Integrated Passive UHF RFID Transponder IC With $16.7-\mu \mathrm{m}$ Minimum RF Input Power," IEEE Journal of Solid-State Circuits, vol. 38, issue 10, pp. 1602-1608, October 2003.
[3] M. Abramowitz, I. A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, $9^{\text {th }}$ ed, New York, USA: Dover, 1972, pp. 358-364.
[4] B. Razavi, RF Microelectronics, Upper Saddle River, NJ: Prentice Hall PTR, 1997, pp. 74-93.
[5] SPICE MOS Model BSIM3v3. Available: http://www.device.eecs.berkeley.edu/~bsim3.
[6] G. De Vita, G. Iannaccone, "Ultra-Low-Power RF Section of a Passive Microwave RFID Transponder in $0.35 \mu \mathrm{~m}$ BiCMOS," Proc. ISCAS 2005, Kobe, Japan, May 2005.
[7] CEPT REC 70-03 Annex 1, ETSI EN 330 220-1.
[8] M. Kossel, "Microwave Backscatter Modulation Systems", 2000 IEEE MTT-S Int. Microwave Symp. Dig., Boston, MA, June 2000, pp 11-16.
[9] M. Kossel, "An Active Tagging System Using Circular Polarization Modulation," IEEE Trans. Microwave Theory and Tech., vol. 47, issue 12, pp. 2242-2248, December 1999.
[10]F. Svelto, R. Castello, "A Bond-Wire Inductor-MOS Varactor VCO Tunable from 1.8 to 2.4 GHz ," IEEE Transactions on Microwave Theory and Techniques, Vol. 50, issue 1, January 2002.
[11]U.K. Moon, K. Mayaram, J.T. Stonick, "Spectral Analysis of Time-Domain Phase Jitter Measurements," IEEE Transactions on Circuits and Systems, Vol. 49, issue 5, May 2002.
[12]C. Liu, J.A. McNeill, "Jitter in Oscillators with 1/f Noise Sources," Proc. 2004 Int. Symp. On Circuits and Systems, Vancouver, Canada, May 2004, pp. 773776.
[13] G. De Vita, G. Iannaccone, "Design Criteria for the RF section of Long Range passive RFID Systems," Proc. NORCHIP 2004, Oslo, Norway, November 2004, pp. 107-110.

## 3. Voltage Reference

### 3.1. Introduction

Low voltage and extreme low power are essential design requirements for circuits and systems to be deployed in a pervasive electronics scenario, where battery replacement can be very costly or where other scarce energy scavenging techniques are used. This leads to a strong demand for circuit building blocks operating with low supply voltage and sub microwatt power. Among them, voltage reference generators are used in almost all analog and digital systems to generate a DC voltage independent of the supply voltage and of temperature variations and they are preferentially implemented with a standard CMOS process for compatibility with the rest of the system. A common way to generate a reference voltage is to use a bandgap voltage reference, which can be implemented in any standard CMOS technology by exploiting the parasitic vertical BJTs [1], [2]. Bandgap voltage references typically provide a voltage around 1.25 V and then require an even larger supply voltage. Sub-1 V Operation can be enabled by using resistive subdivision methods [2].
Other voltage references are based on the availability of transistors with two different threshold voltages in the same CMOS technology. Such feature can be obtained by using a selective channel implant [3], [4], by using different materials for the gate stack [5], by doping differently the polysilicon gates [6]. Such solutions can not be implemented in a standard CMOS technology because they require additional fabrication steps. Other types of voltage references, implemented with a standard CMOS technology, are based on a weighted difference between the gatesource voltages of two MOS transistors [7], [8], [9] but they can not usually operate in the sub-1 V regime.

Design solutions of this type typically compensate the temperature dependence of the mobility only at the reference temperature, leading to a degraded temperature coefficient when moving away from the reference temperature [1], [2], [7], [8]. Moreover, they use one or more resistances leading to a large area occupation on the chip.

In this section we present some CMOS voltage references, which exploit the MOS characteristics in the saturation and in the subthreshold regions, obtaining a perfect cancellation of the effect of the temperature dependence of the carrier mobility for any temperature with very low-power consumption. The CMOS

a

b

Fig. 3-1: a) Simplified circuit of a conventional voltage reference based on the difference between the gate-source voltages of two MOS transistors, b) Simplified circuit of the voltage reference proposed in [8].
voltage reference, presented in subsection 3.4, also compensates second order effects, such as the body effect and channel length modulation, achieving very good temperature compensation. Finally, the CMOS voltage reference proposed in subsection 3.5 , is able to operate with a supply voltage smaller than 1 V ensuring, at the same time, the complete suppression of the temperature dependence of the mobility and the compensation of the second order effects on the temperature coefficient. Such design is thus suitable for low-voltage and low-power operation.

### 3.2. Overview of CMOS-Based Voltage Reference

A simplified circuit of a conventional voltage reference generator based on the difference between the gate-source voltages of two MOS transistors is shown in Fig. 3-1a. Both transistors work in the saturation region. The I-V characteristic of a MOS transistor can be well approximated by using (3-1).

$$
\begin{equation*}
I_{D}=\frac{\mu C_{o x}}{2} \frac{W}{L}\left(V_{G S}-V_{t h}\right)^{2}\left(1+\lambda V_{D S}\right), \tag{3-1}
\end{equation*}
$$

where $\mu$ is the electron mobility in the channel, $V_{t h}$ is the threshold voltage, $m$ is the subthreshold swing parameter, $\lambda$ is the channel length modulation coefficient, $W$ and $L$ are the channel width and length, respectively. By using (3-1) and by neglecting the channel length modulation effect $(\lambda=0)$, the reference voltage $V_{\text {REF }}$ has the expression shown below [9],

$$
\begin{equation*}
V_{R E F}=V_{G S 2}-V_{G S 1}=V_{t h 2}-V_{t h 1}+\sqrt{2 I}\left(\sqrt{\frac{1}{k_{2}}}-\sqrt{\frac{1}{k_{1}}}\right) \tag{3-2}
\end{equation*}
$$

where $I$ is the current indicated in Fig. 3-1a.
The voltage reference generators, already reported in the literature, and based on the difference between the gate-source voltages of two MOS transistors [7], [8], [10] do not allow us to achieve a perfect suppression of the temperature dependence of the mobility and of the bias current. Therefore, a low bias current is used so that the temperature coefficient of the reference voltage is dominated by the threshold voltages of the two MOS transistors. As a consequence, the reference voltage is well approximated by the difference between the threshold voltages of the two MOS transistors ( $V_{\text {REF }} \cong V_{\text {th2 }}-V_{\text {th } 1}$ ). In order to be able to use such a circuit as a voltage reference generator, we need: $i$ ) additional fabrication steps to implement multi-threshold voltage MOS transistors to generate a non zero reference voltage; ii) good process control to make the temperature coefficients of the two MOS transistors as close as possible. However, the temperature coefficient of the two MOS transistors is affected by the temperature dependence of the mobility, whose effect is not cancelled but just attenuated by choosing a low biasing current.

Another design based on the difference between the gate-source voltages of two MOS transistors has been presented in the literature [8]. A simplified circuit of such design is shown in Fig. 3-1b. The current $I$, indicated in Fig. 3-1b, is a PTAT current. By using (3-1) with $\lambda=0$, the expression of the reference voltage is,

$$
\begin{equation*}
V_{R E F}=\left(1+\frac{R_{1}}{R_{2}}\right) V_{G S n}-\left|V_{G S p}\right|, \tag{3-3}
\end{equation*}
$$

where $R_{1}$ and $R_{2}$ are the resistances shown in Fig. 3-1b, $V_{g s n}$ and $V_{g s p}$ are the gatesource voltages of $\mathrm{M}_{\mathrm{n}}$ and $\mathrm{M}_{\mathrm{p}}$. Thanks to the use of a weighted difference between the gate-source voltages of two MOS transistors, a standard CMOS process can be used. As shown in [8], the resistances ratio is chosen in order to compensate the temperature dependence of the threshold voltage while the ratios of the channel width to the channel length of the two transistors $M_{n}$ and $M_{p}$ are chosen in order to compensate the mobility temperature dependence only at the reference temperature. As a consequence, notwithstanding the advantage of using a standard CMOS process, the temperature dependence of mobility on the output reference voltage is still not completely cancelled, degrading the temperature coefficient when moving away from the reference temperature. A measured temperature coefficient of $36.9 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ is achieved. Furthermore, the minimum supply voltage is larger than 1.4 V , preventing a low voltage application.

### 3.3. Proposed Voltage Reference 1

In this sub-section we present a voltage reference generator, which can be implemented in any standard CMOS technology, based on the weighted gate-source voltage difference between two NMOS transistors. The proposed design leads to a perfect cancellation of the effect of the temperature dependence of the carrier mobility on the output regulated voltage, for any temperature, achieving very good temperature compensation.

### 3.3.1. Operating Principle of the Proposed Reference Voltage Generator

Let us consider the active load of the proposed voltage reference generator, shown in Fig. 3-2. Assuming that the current in $\mathrm{R}_{1}$ and $\mathrm{R}_{2}$ is negligible with respect to the current in $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$, and that both $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$ work in the saturation region with current $M I_{0}$, the output reference voltage is given by

$$
\begin{equation*}
V_{R E F}=\left(1+\frac{R_{1}}{R_{2}}\right) V_{G S 8}-V_{G S 7}=\frac{R_{1}}{R_{2}} V_{t h 8}+\sqrt{2 M I_{0}}\left[\left(1+\frac{R_{1}}{R_{2}}\right) \frac{1}{\sqrt{k_{8}}}-\frac{1}{\sqrt{k_{7}}}\right], \tag{3-4}
\end{equation*}
$$



Fig. 3-2: Proposed Voltage Reference Circuit 1.
where $V_{t h 8}$ is the threshold voltage of $\mathrm{M}_{8}$, assumed to be equal to that of $\mathrm{M}_{7}$ (negligible body effect). As is evident from (3-4), the temperature coefficient of the output regulated voltage consists of a first component due to the temperature dependence of the threshold voltage and a second component due to the temperature dependence of the mobility and of the bias current. Since $k_{7}$ and $k_{8}$ are proportional to the mobility, in order to completely cancel the effect of the temperature dependence of the mobility on the output regulated voltage, the bias current must in turn be proportional to mobility.
As a first approximation we can consider that the threshold voltage of an NMOS transistor linearly decreases with temperature, as shown below,

$$
\begin{equation*}
V_{t h}(T)=V_{t h}\left(T_{0}\right)-K_{t 1}\left(T-T_{0}\right), \tag{3-5}
\end{equation*}
$$

where $K_{\mathrm{t} 1}$ is a BSIM3v3 coefficient that models the temperature dependence of the threshold voltage, $T$ is the absolute temperature and $T_{0}$ is the absolute temperature at which $K_{\mathrm{t} 1}$ is calculated. Since the threshold voltage has a negative linear temperature coefficient, the second component of (3-4) must have a positive linear temperature coefficient to allow us to achieve the temperature compensation. As a consequence, once the complete cancellation of the temperature dependence of the mobility is achieved, as previously explained, we require that the bias current has a square dependence on the temperature so that the second component of (3-4) has a positive linear temperature coefficient. To summarize, in order to achieve the temperature compensation with a perfect cancellation of the temperature dependence of mobility for any temperature, we need a bias current proportional to mobility and to the temperature squared, that is $I_{0} \propto \mu(T) T^{2}$. We have found a solution to generate such a current based on MOS transistors in the saturation and in the subthreshold region, as will be explained in the next section.

### 3.3.2. Circuit Description

The proposed voltage reference generator is shown in Fig. 3-2. It consists of a circuit that generates a current $I_{0}$ almost independent of the supply voltage $V_{D D}$; such current is then amplified and injected into an active load to generate the reference voltage.

## Current Generator Circuit

In the current generator circuit a current mirror imposes equal currents in the two branches of the circuit. Since a simple CMOS current mirror is used to obtain a larger dynamic range, in order to reduce the channel length modulation effect of $\mathrm{M}_{5}$ and $\mathrm{M}_{6}$, which causes a mismatch between the currents in the two branches, the channel lengths of the two transistors $\mathrm{M}_{5}$ and $\mathrm{M}_{6}$ are chosen large enough. Transistors $M_{1}$ and $M_{2}$ are biased in the subthreshold region, while transistors $M_{3}$ and $\mathrm{M}_{4}$ work in the saturation region. Such behavior is achieved through careful biasing: the gate-source voltage of $\mathrm{M}_{3}\left(\mathrm{M}_{4}\right)$ must be larger than the gate-source voltage of $\mathrm{M}_{1}\left(\mathrm{M}_{2}\right)$. As a consequence, since the four transistors have the same drain current, the $W / L$ ratio of $\mathrm{M}_{1}\left(\mathrm{M}_{2}\right)$ has to be larger than that of $\mathrm{M}_{3}\left(\mathrm{M}_{4}\right)$. The IV characteristics of a MOS transistor in the subthreshold region can be approximated by

$$
\begin{equation*}
I_{D}=\mu C_{o x} V_{T}^{2} \frac{W}{L} \exp \left(\frac{V_{G S}-V_{t h}}{m V_{T}}\right)\left[1-\exp \left(-\frac{V_{D S}}{V_{T}}\right)\right], \tag{3-6}
\end{equation*}
$$

where $V_{T}$ is the thermal voltage. Referring to Fig. 3-2, the voltage at the gates of $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ can be written by following both the path from the gate to ground through $\mathrm{M}_{2}$ and $\mathrm{M}_{4}$, and that from the gate to ground through $\mathrm{M}_{1}$ and $\mathrm{M}_{3}$. By equating the two expressions for the voltage at the gates of $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$, obtained from the two paths, previously described, one can obtain

$$
\begin{equation*}
V_{G S 1}+V_{G S 3}=V_{G S 2}+V_{G S 4} . \tag{3-7}
\end{equation*}
$$

Using (3-1) to derive the gate-source voltages of $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$ and (3-6) to derive the gate-source voltages of $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ in the case of a drain-source voltage much larger than the thermal voltage $V_{T}$, and assuming that the currents in the two branches of the current circuit generator are equal, (3-7) becomes
$V_{t h 1}+m V_{T} \ln \left(\frac{I_{0}}{\mu C_{o x} V_{T}^{2} W_{1} / L_{1}}\right)+V_{t h 3}+\sqrt{\frac{2 I_{0}}{k_{3}}}=V_{t h 2}+m V_{T} \ln \left(\frac{I_{0}}{\mu C_{o x} V_{T}^{2} W_{2} / L_{2}}\right)+V_{t h 4}+\sqrt{\frac{2 I_{0}}{k_{4}}},(3-8)$
where, $I_{0}$ is the current shown in Fig. 3-2. In (3-8), we have $V_{t h 3}=V_{t h 4}$. In addition, the overdrive of $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$ is very low so that $V_{g s 3}$ and $V_{g s 4}$ differ by only few tens of mV and therefore the threshold voltages of $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ are increased by practically the same amount due to the body effect. By assuming $V_{t h 1}=V_{t h 2}$, we can derive the expression for the current $I_{0}$ as

$$
\begin{equation*}
I_{0}=\frac{m^{2} V_{T}^{2} k_{4}}{2}\left(\frac{N}{N-1}\right)^{2} \ln ^{2}\left(\frac{W_{2} / L_{2}}{W_{1} / L_{1}}\right), \tag{3-9}
\end{equation*}
$$

where $N=\sqrt{k_{3} / k_{4}}$. As it is evident from (3-9), the proposed design allows us to generate a bias current that is linearly dependent on the mobility, through $k_{4}$, and dependent on the temperature squared through $V_{T}$. As already said in the previous section, such a current allows us to obtain the temperature compensation with a perfect cancellation of the effect of the temperature dependence of the mobility. Furthermore, the proposed configuration of the current generator allows us not to use a resistance to generate the current $I_{0}$, as it usually happens in such kind of circuits [8], [11]. This is particularly important if the current to be generated has to be very small, to drastically reduce the power consumption of the voltage reference circuit. In such case a large resistance would be required causing a large area occupation on the chip. Moreover, since the two transistors $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ operate in the subthreshold region, the effect of channel length modulation is negligible
provided that their drain-source voltages are sufficiently larger than the thermal voltage, as evident from (3-6). On the other hand, the channel length modulation effect of $M_{3}$ and $M_{4}$ is negligible because they are long-channel devices, and their drain-source voltages have very small variations when the supply voltage is varied, because they are equal to the gate-source voltages, which, for small variation of $I_{0}$, are almost constant. Almost all the variation of the supply voltage drops on the drain-source voltages of $\mathrm{M}_{2}$ and $\mathrm{M}_{5}$. In order to drastically reduce the channel length modulation effect of $\mathrm{M}_{5}$, its channel length has to be quite large. In such a way, the channel length modulation effect of all transistors in the current circuit generator of Fig. 3-2 is quite low leading to good performances in terms of line sensitivity and PSRR.
Since the reference voltage generator has two stable states, corresponding to the current given by (3-9) and to zero current, a start-up circuit is used to ensure that the former stable state is achieved. If $I_{0}$ is zero, it provides a start-up current to change the stable state.

## Active Load

The active load used to generate a reference voltage with a low temperature drift consists of two NMOS transistors biased by a current obtained by mirroring the current $I_{0}$ and amplifying it by a factor $M$, as shown in Fig. 3-2. Both transistors $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$ operate in the saturation region. The output voltage reference has the expression shown below,

$$
\begin{equation*}
V_{R E F}=\left(1+\frac{R_{1}}{R_{2}}\right) V_{G S 8}-V_{G S 7} . \tag{3-10}
\end{equation*}
$$

where $R_{1}$ and $R_{2}$ are shown in Fig. 3-2. Since $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$ operate in the saturation region, from (3-1), we can derive the expressions for $V_{G S 7}$ and $V_{G S 8}$ as ,

$$
\begin{equation*}
V_{G S 7}=V_{t h 7}+\sqrt{\frac{2 M I_{0}}{k_{7}}} ; V_{G S 8}=V_{t h 8}+\sqrt{\frac{2 M I_{0}}{k_{8}}} . \tag{3-11}
\end{equation*}
$$

Since $I_{0}$ is linearly dependent on the mobility, as evident from (3-9), and also $k_{7}, k_{8}$ are linearly dependent on mobility, both $V_{G S 7}$ and $V_{G S 8}$ turn out to be independent of mobility. As a consequence, the output regulated voltage, and then its temperature coefficient, are independent of electron mobility. Then, thanks to the particular topology chosen, the effect of the temperature dependence of mobility on the output voltage is completely suppressed, allowing us to achieve a very good temperature
coefficient, as shown by experimental results.
As will be clearer later, in order to ensure the correct temperature compensation, it is necessary that most of the bias current $M I_{0}$ flows through the transistors $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$ rather than through the resistances $\mathrm{R}_{1}$ and $\mathrm{R}_{2}$. As a consequence, once we fix the maximum resistance values acceptable for a reasonable area occupation on the chip, we determine the minimum value of the bias current that ensures the correct operation of the voltage reference generator, from the condition

$$
\begin{equation*}
M I_{0} \gg \frac{V_{D S 8}+V_{G S 7}}{R_{1}+R_{2}} \tag{3-12}
\end{equation*}
$$

Equation (3-12) limits the minimum power consumption of the voltage reference generator. In order to set the biasing current $M I_{0}$ as small as possible, the two transistors $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$ were dimensioned so that i) $V_{G S 7}$ is very close to the threshold voltage but sufficient to ensure the operation of $\mathrm{M}_{7}$ in the saturation region and ii) $V_{D S 8}$ is as small as possible to ensure the operation of $\mathrm{M}_{8}$ in the saturation region. In such a way, the numerator of the second member of (3-12) is as small as possible allowing us to minimize the bias current for given values of the two resistances. The resistances are implemented by using high resistive poly to minimize the area occupation on the chip.

### 3.3.3. Supply Voltage Dynamic Range

The minimum supply voltage is imposed by the current generator circuit. In particular, we have to ensure that $\mathrm{M}_{2}$ has a drain-source voltage of at least 100 mV so that the effect of the drain-source voltage in (3-6) and then the channel length modulation of $\mathrm{M}_{2}$ can be neglected. Consequently, the following expression has to be satisfied,

$$
\begin{equation*}
V_{D D}>\left|V_{G S 6}\right|+V_{D S 2 M I N}+V_{G S 4}, \tag{3-13}
\end{equation*}
$$

which implies that the supply voltage must be larger than 1.5 V in the AMS $0.35 \mu \mathrm{~m}$ CMOS process. The maximum supply voltage is imposed by the maximum drain-source voltage allowed for MOS transistors, as shown below,

$$
\begin{equation*}
V_{D D}<\left|V_{D S S M A X}\right|+V_{G S 1}+V_{G S 3} . \tag{3-14}
\end{equation*}
$$

Since in the AMS $0.35 \mu \mathrm{~m}$ CMOS process the maximum value for the drainsource voltage of a MOS transistor is 3.3 V , the maximum value of the supply voltage is about 4.3 V .

### 3.3.4. Temperature Compensation

As a first approximation we can consider that the threshold voltage of an NMOS transistor linearly decreases with the temperature, as shown in (3-5). In the case of the AMS $0.35 \mu \mathrm{~m}$ CMOS IC technology, the parameter $K_{t l}$ is $0.33 \mathrm{mV} /{ }^{\circ} \mathrm{C}$ for an NMOS transistor and $0.45 \mathrm{mV} /{ }^{\circ} \mathrm{C}$ for a PMOS transistor. Let us define $h$ as,

$$
\begin{equation*}
h=M m^{2} \frac{W_{4}}{L_{4}}\left(\frac{N}{N-1}\right)^{2} \ln ^{2} \frac{W_{2} / L_{2}}{W_{1} / L_{1}} . \tag{3-15}
\end{equation*}
$$

By using (3-15), the current $M I_{0}$ that flows in the active load can be written as,

$$
\begin{equation*}
M I_{0}=\frac{\mu C_{o x} V_{T}^{2}}{2} h . \tag{3-16}
\end{equation*}
$$

As a consequence, assuming that (3-12) is fulfilled and then that $M I_{0}$ is the drain current of $\mathrm{M}_{7}$ and $\mathrm{M}_{8,}$, by using (3-11), the gate-source voltages of $\mathrm{M}_{7}$ and $\mathrm{M}_{8}, V_{G S 7}$ and $V_{G S 8}$, can be expressed as shown below

$$
\begin{equation*}
V_{G S 7}=V_{t h 7}+\sqrt{\frac{h}{W_{7} / L_{7}}} V_{T} ; V_{G S 8}=V_{t h 8}+\sqrt{\frac{h}{W_{8} / L_{8}}} V_{T} . \tag{3-17}
\end{equation*}
$$

From (3-10) and (3-17), we can derive the following expression for the reference voltage,

$$
\begin{equation*}
V_{R E F}=\left(1+\frac{R_{1}}{R_{2}}\right)\left(V_{t h 8}+\sqrt{\frac{h}{W_{8} / L_{8}}} V_{T}\right)-V_{t h 7}-\sqrt{\frac{h}{W_{7} / L_{7}}} V_{T} . \tag{3-18}
\end{equation*}
$$

If $V_{\text {REF }}$ is small enough to minimize the body effect on $\mathrm{M}_{7}$ and large enough to bias $\mathrm{M}_{8}$ in the saturation region, we can assume $V_{t h 7}=V_{t h 8}$. By differentiating (3-18) with respect to the temperature and taking into account (3-5), one obtains

$$
\frac{\partial V_{R E F}}{\partial T}=\left(1+\frac{R_{1}}{R_{2}}\right)\left(-K_{t 1}+\sqrt{\frac{h}{W_{8} / L_{8}}} \frac{k_{B}}{q}\right)+K_{t 1}-\sqrt{\frac{h}{W_{7} / L_{7}}} \frac{k_{B}}{q},(3-19)
$$

where $k_{B}$ is the Boltzmann constant and $q$ is the electron charge. By setting (3-19) to zero, we obtain

$$
\begin{equation*}
\frac{R_{1}}{R_{2}}=\frac{\sqrt{h} \frac{k_{B}}{q}\left(\frac{1}{\sqrt{W_{8} / L_{8}}}-\frac{1}{\sqrt{W_{7} / L_{7}}}\right)}{K_{t 1}-\sqrt{h} \frac{k_{B}}{q} \frac{1}{\sqrt{W_{8} / L_{8}}}} . \tag{3-20}
\end{equation*}
$$

Therefore, if (3-20) is satisfied, we obtain that the temperature coefficient (3-19) is zero for any temperature. It is clear that this is true within the approximation done in (3-5) and the simplified transistor characteristics expressions (3-1) and (3-6) for the operation of a MOS transistor in the saturation and subthreshold regions. Since the temperature dependence of the threshold voltage is not perfectly linear, a temperature dependent error will appear at the output of the reference voltage generator as we move away from the reference temperature at which the coefficient $K_{\mathrm{t} 1}$ was computed.
Once the two transistors $M_{7}$ and $M_{8}$ are dimensioned for the minimum value of the bias current, given by (3-12), from (3-20) we can derive the ratio between the two resistance values, in order to achieve the best temperature compensation.
From (3-19), it is evident that, as already said, the temperature coefficient of the proposed voltage reference generator does not depend upon the mobility. Then, thanks to the particular topology used both for the current generator circuit and the active load, the effect of the temperature dependence of mobility is suppressed, achieving a low temperature coefficient.

If we take into consideration an error in the resistor ratio $R_{1} / R_{2}$, from (3-19) it is clear that when the resistor ratio increases (decreases) with respect to its nominal value the temperature coefficient becomes positive (negative). Anyway, since in our design the value of the resistor ratio is much smaller than unity, an error on the resistance ratio is drastically reduced by summing it to the unity in (3-19). This allows us not to use trimming procedures.


Fig. 3-3: Die Photograph (core).

### 3.3.5. Line Sensitivity

Because of the channel length modulation effect, when the supply voltage varies, the bias current $I_{0}$ varies as well, causing the variation of the output reference voltage. By substituting the expressions of $V_{G S 7}$ and $V_{G S 8}$ given in (3-11), in (3-10), we can derive the following expression for the output regulated voltage,

$$
\begin{equation*}
V_{R E F}=\left(1+\frac{R_{1}}{R_{2}}\right) V_{t h 8}-V_{t h 7}+\sqrt{M I_{0}}\left[\left(1+\frac{R_{1}}{R_{2}}\right) \sqrt{\frac{2}{k_{8}}}-\sqrt{\frac{2}{k_{7}}}\right] . \tag{3-21}
\end{equation*}
$$

As shown in (3-21), only one term of the output reference voltage weakly depends on the bias current (via a square root). Furthermore, such term has a very small weight in (3-21), due to the fact that the overdrives of the two transistors M7 and M8 are very small and then the gate-source voltages are very close to the threshold voltage.

### 3.3.6. Experimental Results

The proposed voltage reference has been implemented with AMS $0.35 \mu \mathrm{~m}$ CMOS technology. The die photograph is shown in Fig. 3-3. Measurements show that the proposed voltage reference generator generates a mean reference voltage of about 168 mV with a variation of $\pm 2.3 \mathrm{mV}$ at room temperature when the supply voltage varies from 1.5 V to 4.3 V , as shown in Fig. 3-4a. Fig. 3-5 shows the output voltage dependence on temperature for different values of the supply voltage. The measured temperature coefficient over a temperature range from $0^{\circ} \mathrm{C}$ to $80{ }^{\circ} \mathrm{C}$ at $V_{D D}=2 \mathrm{~V}$ and $V_{D D}=3 \mathrm{~V}$ is $25 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ and increases to 37 and $39 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ at


Fig. 3-4: Experiments: a) Output Voltage vs. Supply voltage at room temperature, b) PSRR at room temperature and for a supply voltage of 2 V .


Fig. 3-5: Measured output voltage vs. temperature for 4 values of the supply voltage.
$V_{D D}=4.3 \mathrm{~V}$ and $V_{D D}=1.5 \mathrm{~V}$, respectively, corresponding to the maximum and minimum allowed supply voltage. At $80^{\circ} \mathrm{C}$ the current drawn at the maximum supply voltage is $2.4 \mu \mathrm{~A}$ and at the minimum supply voltage is $1.5 \mu \mathrm{~A}$. At room temperature, instead, the current drawn at the maximum supply voltage is $2.1 \mu \mathrm{~A}$ and at the minimum supply voltage is $1.2 \mu \mathrm{~A}$. The power supply rejection ratio, without any filtering capacitor, is -65 dB at 100 Hz and -57 dB at 10 MHz , for the smallest supply voltage. At larger supply voltage the power supply rejection ratio reduces to -74 dB at 100 Hz and to -59 dB at 10 MHz , as shown in Fig. 3-4b. The occupied chip area is $0.08 \mathrm{~mm}^{2}$.

### 3.4. Proposed Voltage Reference 2

In this sub-section we present a second voltage reference generator, which can be implemented in any standard CMOS technology, based on the weighted gate-source voltage difference between two NMOS transistors. The proposed design leads to a perfect cancellation of the effect of the temperature dependence of the carrier mobility, for any temperature, and to a compensation of the second order effects, achieving very good temperature compensation. With respect to the similar solution proposed in the previous sub-section, here we implement the circuit without using any resistance and therefore reducing the area occupation on the chip by almost one order of magnitude, and exhibiting a power consumption at least one order of magnitude smaller than all other solutions presented in the literature.

### 3.4.1. Circuit Description

The proposed voltage reference generator is shown in Fig. 3-6. The operating principle of such voltage reference is exactly equal to the previous one. Indeed, by assuming that all transistors of the active load, shown in Fig. 3-6, work in the saturation region, the output reference voltage is given by

$$
\begin{equation*}
V_{R E F}=V_{t h}+\left[\frac{1}{\sqrt{k_{8}}}\left(1+\sqrt{\frac{W_{10} / L_{10}}{W_{9} / L_{9}}}\right)-\frac{1}{\sqrt{k_{7}}}\right] \sqrt{2 I_{0}} \tag{3-22}
\end{equation*}
$$

As it is evident from (3-22), the temperature coefficient of the output reference voltage consists of a first component due to the temperature dependence of the threshold voltage and of a second component due to the temperature dependence of mobility and of the bias current. Since $k_{7}$ and $k_{8}$ are proportional to mobility, a bias current proportional to mobility would completely suppress the effect of the temperature dependence of mobility on the output reference voltage. Once the temperature dependence of the mobility has been suppressed, we require that the


Fig. 3-6: Proposed Voltage Reference Circuit.
bias current has a square dependence on the temperature, to compensate the temperature dependence of the threshold voltage. As a consequence, also in this case, we need $I_{0} \propto \mu(T) T^{2}$ and then the same current generator circuit, proposed in the voltage reference of section 2.4 , is used.

The active load, instead, used to generate a reference voltage with a low temperature drift, is different than the previous design, and consists of two NMOS transistors biased by mirroring the current $I_{0}$ and an active voltage divider formed by $\mathrm{M}_{9}$ and $\mathrm{M}_{10}$, as shown in Fig. 3-6. All transistors in the active load operate in the saturation region. The output voltage reference has the expression shown below,

$$
\begin{equation*}
V_{R E F}=\left(1+\sqrt{\frac{W_{10} / L_{10}}{W_{9} / L_{9}}}\right) V_{G S 8}+\left(1-\sqrt{\frac{W_{10} / L_{10}}{W_{9} / L_{9}}}\right) V_{t h}-V_{G S 7} . \tag{3-23}
\end{equation*}
$$

Since $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$ operate in the saturation region, from (3-1), the expressions for $V_{G S 7}$ and $V_{G S 8}$ are,

$$
\begin{equation*}
V_{G S 7}=V_{t h}+\sqrt{\frac{2 I_{0}}{k_{7}}} ; V_{G S 8}=V_{t h}+\sqrt{\frac{2 I_{0}}{k_{8}}} \tag{3-24}
\end{equation*}
$$

Since $I_{0}$ is linearly dependent on mobility, as evident from (3-24), and then $k_{7}, k_{8}$ are linearly dependent on mobility as well, the output reference voltage and then its temperature coefficient, are independent of electron mobility.
As will be clearer later, in order to ensure the correct temperature compensation, it is necessary that most of the bias current $I_{0}$ flows through the transistors $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$ rather than through $\mathrm{M}_{9}$ and $\mathrm{M}_{10}$. By imposing a current in $\mathrm{M}_{8}$ much larger than that in $\mathrm{M}_{10}$, the following relationship must be fulfilled,

$$
\begin{equation*}
\frac{k_{8}}{2}\left(V_{G S 8}-V_{t h}\right)^{2} \gg \frac{k_{10}}{2}\left(V_{G S 10}-V_{t h}\right)^{2} . \tag{3-25}
\end{equation*}
$$

Equation (3-25) is satisfied for $W_{8} / L_{8} \gg W_{10} / L_{10}$.
In the voltage reference generator proposed in the previous sub-section a resistive voltage divider is used in the active load rather than the active divider. Since only a negligible fraction of the biasing current must flow across the voltage divider, very large resistances are needed in the case of small biasing current, causing an area occupation much larger than that it would be needed with an active divider. Moreover, for a given area occupation, the use of an active divider allows us to drastically reduce the bias current and then the power consumption required to verify the condition that almost all the bias current flows through $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$. As a consequence, the solution of using an active voltage divider allows us to drastically reduce the power consumption by about one order of magnitude compared to that proposed in sub-section 2.3. At the same time, the area occupation on the chip is drastically reduced.

### 3.4.2. Temperature Compensation

For the current $I_{0}$ we can use the same expression proposed in (3-9). As a consequence, assuming that (3-25) is fulfilled and then that $I_{0}$ is the drain current of $\mathrm{M}_{7}$ and $\mathrm{M}_{8,}$, by using (3-24), the gate-source voltages of $\mathrm{M}_{7}$ and $\mathrm{M}_{8}, V_{G S 7}$ and $V_{G S 8}$, can be expressed as

$$
\begin{equation*}
V_{G S 7}=V_{t h}+\sqrt{\frac{h}{W_{7} / L_{7}}} V_{T} ; V_{G S 8}=V_{t h}+\sqrt{\frac{h}{W_{8} / L_{8}}} V_{T} . \tag{3-26}
\end{equation*}
$$

The term $h$ is defined as in (3-15) considering that in this case $M=1$. From (3-23) and (3-26), we can derive the following expression for the reference voltage,

$$
\begin{equation*}
V_{R E F}=V_{t h}+\left[\left(1+\sqrt{\frac{W_{10} / L_{10}}{W_{9} / L_{9}}}\right) \sqrt{\frac{h}{W_{8} / L_{8}}} V_{T}-\sqrt{\frac{h}{W_{7} / L_{7}}} V_{T}\right] \tag{3-27}
\end{equation*}
$$

By differentiating (3-27) with respect to the temperature and taking into account (3-5), one obtains

$$
\frac{\partial V_{R E F}}{\partial T}=-K_{t 1}+\left(1+\sqrt{\frac{W_{10} / L_{10}}{W_{9} / L_{9}}}\right) \sqrt{\frac{h}{W_{8} / L_{8}}} \frac{k_{B}}{q}-\sqrt{\frac{h}{W_{7} / L_{7}}} \frac{k_{B}}{q} .
$$

By equating (3-28) to zero, we obtain

$$
\begin{equation*}
\sqrt{\frac{W_{10} / L_{10}}{W_{9} / L_{9}}}=\frac{K_{t 1}+\sqrt{h} \frac{k_{B}}{q}\left(\frac{1}{\sqrt{W_{7} / L_{7}}}-\frac{1}{\sqrt{W_{8} / L_{8}}}\right)}{\sqrt{h} \frac{k_{B}}{q} \frac{1}{\sqrt{W_{8} / L_{8}}}} \tag{3-29}
\end{equation*}
$$

Therefore, if (3-29) is satisfied, we obtain that the temperature coefficient (3-28) is zero for any temperature. It is clear that this is true within the approximation done in (3-5). Since the temperature dependence of the threshold voltage is not perfectly linear, a temperature dependent error will appear at the output of the reference voltage generator as we move away from the reference temperature.

### 3.4.3. Second order effects on the temperature coefficient

## Channel length modulation effect

In the calculation of the current $I_{0}$ in (3-9) the channel length modulation effect was neglected. In order to take it into account, we have to consider the I-V characteristic of a MOS in the saturation region, given by (3-1), for $\lambda \neq 0$.
Since transistors $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ work in the subthreshold region, the drain current is exponentially dependent on the drain source voltage and then, if their drain source voltage is 3 or 4 times larger than the thermal voltage the channel length modulation effect can be neglected. For transistors $M_{3}$ and $M_{4}$, instead, we can use
(3-1). In such condition, using the same procedure used to calculate (3-9), the current $I_{0}$ is

$$
\begin{equation*}
I_{0}=\frac{m^{2} V_{T}^{2} k_{4}}{2}\left(\frac{N}{N / \sqrt{1+\lambda V_{D S 4}}-1 / \sqrt{1+\lambda V_{D S 3}}}\right)^{2} \ln ^{2}\left(\frac{W_{2} / L_{2}}{W_{1} / L_{1}}\right) . \tag{3-30}
\end{equation*}
$$

By dimensioning the channel length of $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$ large enough, we can assume that $\lambda V_{D S 3} \ll 1$ and $\lambda V_{D S 4} \ll 1$. In such condition and by using first order Taylor series expansion, the current $I_{0}$ can be rewritten as follows

$$
\begin{equation*}
I_{0} \cong \frac{\mu C_{o x}}{2} h V_{T}^{2}\left(1+\frac{\lambda}{2} \frac{N V_{D S 4}-V_{D S 3}}{N-1}\right)^{2} \cong \frac{\mu C_{o x}}{2} h V_{T}^{2}\left(1+\frac{\lambda}{2} V_{G S 4}\right)^{2} . \tag{3-31}
\end{equation*}
$$

In our design, in order to minimize the current $I_{0}$ and then the power consumption the factor $N$ was set large enough. In such condition and by considering that $V_{D S 4}=V_{G S 4}>V_{D S 3}$, the second approximation of (3-31) can be done. By using (3-23), (3-24) and (3-31), the reference voltage can be rewritten as

$$
V_{R E F}=V_{R E F 0}+\left(V_{R E F 0}-V_{t h}\right) \frac{\lambda}{2} V_{G S 4},
$$

where $V_{\text {REF } 0}$ is the reference voltage calculated for $\lambda=0$ and given by (3-27). By differentiating (3-32) with respect to the temperature and considering that $V_{G S 4}$ is very close to the threshold voltage, to ensure a large dynamic range, if (3-29) is verified, the temperature coefficient is

$$
\begin{equation*}
\frac{\partial}{\partial T} V_{R E F} \cong \frac{\lambda}{2} K_{t 1}\left(2 V_{t h}-V_{R E F 0}\right) . \tag{3-33}
\end{equation*}
$$

As a consequence, the best choice to suppress the effect of the channel length modulation effect on the temperature coefficient is to set $V_{\text {REF } 0}=2 V_{t h} \cong 0.9 \mathrm{~V}$.

### 3.4.4. Body effect

The body effect on the transistors of the current generator circuit of Fig. 3-6 can be neglected because $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$ have the source terminal grounded and the source-
bulk voltages of $M_{1}$ and $M_{2}$ are very close to each other since they are equal to the gate-source voltages of $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$, which are in turn both very close to the threshold voltage to ensure a large dynamic range. As a consequence, the body effect is almost equal for $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ and then, in the calculation of the current $I_{0}$, has no effect since the two threshold voltages are subtracted. In order to take into account the body effect of $\mathrm{M}_{7}$ and $\mathrm{M}_{9}$, we can model the bulk dependency of the threshold voltage as

$$
\begin{equation*}
V_{t h}=V_{t h 0}+K_{1}\left(\sqrt{2 \phi_{S}-V_{B S}}-\sqrt{2 \phi_{S}}\right), \tag{3-34}
\end{equation*}
$$

where $V_{t h 0}$ is $V_{t h}$ for $V_{B S}=0, K_{l}$ is a process constant, $\varphi_{S}$ is the surface potential of the MOS transistor and $V_{B S}$ is the bulk-source voltage. In such condition, the output reference voltage can be rewritten as,

$$
V_{R E F}=V_{R E F 0}+\left(V_{t h 9}-V_{t h 0}\right)-\left(V_{t h 7}-V_{t h 0}\right) .
$$

By using (3-34) to write $V_{t h 7}$ and $V_{t h 9}$, if (3-29) is verified the temperature coefficient can be written as follows

$$
\frac{\partial}{\partial T} V_{R E F}=K_{1} \frac{\partial \phi_{S}}{\partial T}\left(\frac{1}{\sqrt{2 \phi_{S}-V_{B S 9}}}-\frac{1}{\sqrt{2 \phi_{S}-V_{B S 7}}}\right)
$$

As a consequence, the best choice to suppress the body effect on the temperature coefficient is to set $V_{B S 9}=V_{B S 7}$. Thanks to the compensation both of the temperature dependence of the mobility and of the second order effects, the temperature coefficient of the proposed voltage reference generator is much lower than that of other designs that exploits the same principle based on the difference between the gate-source voltages of two MOS transistors [8], [10], as will be shown later.

### 3.4.5. Power Supply Rejection Ratio

Because of the small power consumption, the power supply rejection ratio of the proposed circuit becomes a critical issue. Indeed, for small drain currents the transconductance $g_{m}$ of a MOS transistor becomes very small degrading the PSRR.


Fig. 3-7: Circuit for the calculation of the PSRR at DC.
In order to recovery the PSRR, the drain-source resistance $r_{0}$ of MOS transistors must be set large enough to be able to obtain a high intrinsic gain $g_{m} r_{0}$. In the case of a MOS in the subthreshold region, if $V_{D S} \gg V_{T}$, in virtue of the exponential dependence on the drain-source voltage, the drain-source resistance becomes very large. In the case of a MOS in the saturation region, from (3-1), $r_{d}=1 /\left(\lambda I_{D}\right)$; a large resistance is achieved with a small drain current and a large channel length. In order to calculate the PSRR at DC, we can use the circuit for small signal analysis shown in Fig. 3-7. From such circuit we can write,

$$
\begin{equation*}
\frac{i_{O U T}}{v_{D D}} \cong \frac{1}{r_{05}}\left(\frac{1}{g_{m 2}} \frac{g_{m 1} g_{m 3}}{g_{m 1}+g_{m 3}}-\frac{g_{m 5}}{g_{m 6}}\right)^{-1}, \tag{3-37}
\end{equation*}
$$

where, $g_{m i}$ is the transconductance of transistor $\mathrm{M}_{\mathrm{i}}$ and $r_{05}$ is the drain-source resistance of $\mathrm{M}_{5}$. In order to achieve a small variation of the generated current when the supply voltage varies, the PMOS transistor $\mathrm{M}_{5}$ must have a sufficiently long channel and the terms in the parenthesis must be as large as possible. From the circuit of Fig. 3-7 we can also calculate,


Fig. 3-8: Circuit for the calculation of the PSRR at high frequency.

$$
\begin{equation*}
\frac{v_{R E F}}{v_{D D}}=\left(1-\frac{g_{m 8}}{g_{m 7}} \frac{g_{m 10}}{g_{m 9}+g_{m 10}}\right) \frac{g_{m 9}+g_{m 10}}{g_{m 8} g_{m 9}}\left(\frac{1}{r_{011}}+\frac{g_{m 11}}{g_{m 6}} \frac{i_{O U T}}{v_{D D}}\right) \tag{3-38}
\end{equation*}
$$

As a consequence, in order to achieve a small PSRR, the channel length of transistor
$\mathrm{M}_{11}$ must be set large enough to obtain a large drain-source resistance. As clear from (3-37) and (3-38), the PSRR does not depend on absolute $g_{m}$ values, but only on $g_{m}$ ratios, which can be set as large as required to have a small PSRR also for very small bias current.
In order to calculate the PSRR at high frequency, we can use the circuit of Fig. 3-8. At high frequency, the drain-source resistances and the voltage-controlled current generators of MOS transistors can be neglected and the circuit for small signal analysis at high frequency reduces to a capacitive network. In virtue of the Miller effect on the gate-drain capacitance of $M_{11}$, the capacitance from the gate of $M_{11}$ to ground is much larger than that from the gate of $\mathrm{M}_{11}$ to the supply voltage terminal. As a consequence, for a small signal analysis at high frequency the gate of $\mathrm{M}_{11}$ can be considered grounded. From the circuit of Fig. 3-8 we can derive,

$$
\frac{v_{R E F}}{v_{D D}} \cong \frac{C_{G S 7}}{C_{G S 7}+C_{G D 8}} \frac{C_{D S 11}}{C_{D S 11}+C_{V}},
$$

where, $C_{V} \equiv C_{G D 11}+\left(C_{G S 10}+C_{G S 8}\right) \|\left(C_{G S 7} \| C_{G D 8}+C_{G S 9}\right)$ and $C_{G D i}, C_{G S i}, C_{D S i}$ are the


Fig. 3-9: Die Photograph of the Voltage Reference Generator (core).
gate-drain, gate-source and drain-source capacitances of the transistor $\mathrm{M}_{\mathrm{i}}$. Since the drain-source capacitance is much smaller than the other capacitances, the PSRR at high frequency is very small, as will be shown in the next section.

### 3.4.6. Experimental Results

The proposed voltage reference has been implemented in a standard $0.35 \mu \mathrm{~m}$ CMOS process. The die photograph of the core circuit is shown in Fig. 3-9. Measurements show that the voltage reference generates a mean reference voltage of about 891.1 mV with a variation, at room temperature, of $0.46 \% / \mathrm{V}$, when the supply voltage varies from 1.5 V to 4.3 V , as shown in Fig. 3-10a. The power supply rejection ratio, without any filtering capacitor, is -59 dB at 100 Hz and -52 dB at 10 MHz , as shown in Fig. 3-10b. Fig. 3-11 shows the output voltage dependence on temperature for different values of the supply voltage. The measured temperature coefficient at $V_{D D}=2 \mathrm{~V}$ and $V_{D D}=3 \mathrm{~V}$ is $15 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ and 12 $\mathrm{ppm} /{ }^{\circ} \mathrm{C}$, respectively, and increases to $18 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ at $V_{D D}=4.3 \mathrm{~V}$ and $22 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ at $V_{D D}=1.5 \mathrm{~V}$. At room temperature the current drawn is 110 nA at the maximum supply voltage ( 4.3 V ) and 80 nA at the minimum supply voltage $(1.5 \mathrm{~V})$. At $80^{\circ} \mathrm{C}$ the absorbed current is 130 nA at 4.3 V and 90 nA at 1.5 V .


Fig. 3-10: Experiments: a) Output Voltage vs. Supply voltage at room temperature, b) PSRR at room temperature and for a supply voltage of 3 V .


Fig. 3-11: Measured output voltage vs. temperature for 4 values of the supply voltage.


Fig. 3-12: Proposed voltage reference circuit.

### 3.5. Proposed Voltage Reference 3

Both the voltage references proposed in the previous two sub-sections are not suitable for low-voltage operation because the minimum supply voltage that ensures their correct operation is larger than 1.5 V . The MOS-based voltage reference proposed in this subsection can achieve a very good temperature coefficient thanks to a perfect suppression of the temperature dependence of the mobility and to compensation of the second order effects and, at the same time, allows low voltage operation with a minimum supply voltage smaller than 1 V .

### 3.5.1. Circuit Description

The proposed voltage reference generator is shown in Fig. 3-12. The operating principle of such voltage reference is the same as the previous one. Indeed, by assuming that all transistors of the active load, shown in Fig. 3-12, work in the saturation region, the output reference voltage is given by,

$$
V_{R E F}=V_{t h 10}+\sqrt{\frac{2 I_{0}}{k_{10}}}
$$

where $I_{0}$ is the bias current of $\mathrm{M}_{10}$. As it is evident from (3-40), the temperature coefficient of the output reference voltage consists of a first component due to the temperature dependence of the threshold voltage and of a second component due to the temperature dependence of mobility and of the bias current. Since $k_{10}$ is proportional to mobility, a bias current proportional to mobility would completely suppress the effect of the temperature dependence of mobility on the output reference voltage. Once the temperature dependence of the mobility has been suppressed, we require that the bias current has a square dependence on the temperature, to compensate the temperature dependence of the threshold voltage. As a consequence, also in this case, we need $I_{0} \propto \mu(T) T^{2}$.

A circuit formed by transistors numbered from $\mathrm{M}_{1}$ to $\mathrm{M}_{8}$ generates a current $I_{0}$ as independent as possible of the supply voltage $V_{D D}$. Such current is then injected into the diode-connected NMOS transistor $\mathrm{M}_{10}$. The temperature dependence of $I_{0}$ is compensated by the temperature dependence of the gate-source voltage of $\mathrm{M}_{10}$ generating a temperature compensated reference voltage $V_{R E F}$.

## Current generator circuit

The core of the current generator circuit is represented by transistors $M_{1}-M_{4}$, which determine the value of the current $I_{0}$, whereas transistors $\mathrm{M}_{5}$ and $\mathrm{M}_{6}$ impose equal current $I_{I}$ in $\mathrm{M}_{1}$ and $\mathrm{M}_{3}$ and transistors $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$ impose equal current $I_{0}$ in $M_{2}$ and $M_{4}$. Transistors $M_{1}$ and $M_{3}$ (indicated Fig. 3-12 with a symbol with a thicker line for the gate) are 5V-NMOS transistors with a threshold voltage of 0.7 V ; all the other transistors are 3.3V-MOS transistors with a threshold voltage of 0.45 V and -0.75 V for NMOS and PMOS, respectively. The two different threshold voltages allow us to bias $\mathrm{M}_{1}$ and $\mathrm{M}_{3}$ in the subthreshold region and, at the same time, to bias $M_{2}$ and $M_{4}$ in the saturation region. Such behavior is achieved by setting the gate-source voltages of $\mathrm{M}_{1}, \mathrm{M}_{2}$ and $\mathrm{M}_{3}, \mathrm{M}_{4}$ to a value between 0.45 V and 0.7 V . The I-V characteristics of an NMOS transistor that operates in the saturation and in the subthreshold region can be approximated by (3-1) and (3-6), respectively.
The gate-source voltages of $\mathrm{M}_{1}$ and $\mathrm{M}_{2}\left(\mathrm{M}_{3}\right.$ and $\left.\mathrm{M}_{4}\right)$ are identical and can be extracted from (3-1) and (3-6) by considering $M_{1}$ and $M_{3}$ in subthreshold with drain current $I_{1}$ and $\mathrm{M}_{2}$ and $\mathrm{M}_{4}$ in saturation with a drain current $I_{0}$. Then, by enforcing $V_{G S 1}=V_{G S 2}$ and $V_{G S 3}=V_{G S 4}$, we have

$$
\begin{align*}
& V_{t h 1}+m V_{T} \ln \left(\frac{I_{1}}{\mu C_{o x} V_{T}^{2} W_{1} / L_{1}}\right)=V_{t h 2}+\sqrt{\frac{2 I_{0}}{\mu C_{o x} W_{2} / L_{2}}},  \tag{3-41}\\
& V_{t h 3}+m V_{T} \ln \left(\frac{I_{1}}{\mu C_{o x} V_{T}^{2} W_{3} / L_{3}}\right)=V_{t h 4}+\sqrt{\frac{2 I_{0}}{\mu C_{o x} W_{4} / L_{4}}},
\end{align*}
$$

where we have neglected channel length modulation $(\lambda=0)$ and have set the term between square brackets in (3-6) to unity. Obviously, since the source terminals of all NMOS transistors are grounded, the body effect plays no role so that $V_{t h 1}=V_{t h 3}$ and $V_{t h 2}=V_{t h 4}$. By subtracting (3-41) from (3-42), we can extract the expression of the current $I_{0}$ :

$$
\begin{equation*}
I_{0}=\frac{\mu C_{o x} W_{4} / L_{4}}{2(N-1)^{2}} m^{2} V_{T}^{2} \ln ^{2}\left(\frac{W_{3} / L_{3}}{W_{1} / L_{1}}\right), \tag{3-43}
\end{equation*}
$$

where we define $N=\sqrt{\left(W_{4} / L_{4}\right) /\left(W_{2} / L_{2}\right)}$.

## Active load

The active load consists of a diode-connected NMOS transistor, $\mathrm{M}_{10}$. The current previously generated, given by (3-43), is then injected into the diode connected transistor $\mathrm{M}_{10}$, in order to generate a temperature compensated reference voltage. $\mathrm{M}_{10}$ operates in the saturation region and then by using (3-1) and (3-43), we can derive the output voltage $V_{\text {REF }}$

$$
\begin{equation*}
V_{R E F}=V_{t h 10}+\frac{m V_{T}}{N-1} \sqrt{\frac{W_{4} / L_{4}}{W_{10} / L_{10}}} \ln \left(\frac{W_{3} / L_{3}}{W_{1} / L_{1}}\right) \tag{3-44}
\end{equation*}
$$

The proposed configuration of the voltage reference generator allows us to generate the reference voltage without using any resistance, that are conversely used in similar types of circuits [2], [8]. This is particularly important in the case of an ultra-low-power voltage reference generator because a very large resistance would be necessary to generate the small required current $I_{0}$ (some tens of nA). As a consequence, the proposed circuit topology allows us to drastically reduce the area occupation on the chip, as will be shown later from comparison with the literature.

Since the reference voltage generator has two stable states, corresponding to the current given by (3-43) and to zero current, a start-up circuit (formed by $\mathrm{M}_{1 \mathrm{~S}}-\mathrm{M}_{3 \mathrm{~S}}$ ) is used to ensure that the former stable state is achieved.

### 3.5.2. Design Consideration

## Channel length modulation effect

Since transistors $\mathrm{M}_{2}, \mathrm{M}_{3}, \mathrm{M}_{5}, \mathrm{M}_{8}$ and $\mathrm{M}_{10}$ are diode-connected, almost all the variation of the supply voltage drops on the drain-source voltages of transistors $\mathrm{M}_{6}$, $\mathrm{M}_{7}, \mathrm{M}_{9}$ of the current mirrors, and on the drain-source voltages of transistors $\mathrm{M}_{1}$, $\mathrm{M}_{4}$. As a consequence, in order to drastically reduce the channel length modulation effect, the channel length of all the transistors in the current mirrors and of $\mathrm{M}_{4}$ must be quite large and the drain-source voltage of $\mathrm{M}_{1}$, which operates in the subthreshold region, must be much larger than $V_{T}$ so that the $V_{D S}$ dependence of the current in (3-6) becomes negligible.

## Minimum power consumption dimensioning

The minimum power consumption of the proposed voltage reference generator and then the minimum acceptable value of the bias current $I_{0}$ is imposed by $\mathrm{M}_{2}, \mathrm{M}_{4}$ and $\mathrm{M}_{10}$, which must operate in the saturation region. By assuming that $W_{4} / L_{4}>W_{2} / L_{2}$, if $\mathrm{M}_{4}$ operates in the saturation region with $V_{G S 4}>V_{t h 4}$ then $\mathrm{M}_{2}$, which has the same drain current, will work in the saturation region as well. In such condition, the minimum current $I_{0}$ can be evaluated by imposing that $\mathrm{M}_{4}$ operates in the saturation region with $V_{G S 4}=V_{t h 4}$. The minimum currents $I_{0 M I N}$ and $I_{I M I N}$ have thus the following expressions,

$$
\begin{align*}
& I_{0 M I N}=\frac{\mu C_{o x} W_{2} / L_{2}}{2} m^{2} V_{T}^{2} \ln ^{2}\left(\frac{W_{3} / L_{3}}{W_{1} / L_{1}}\right), \\
& I_{1 M I N}=\mu C_{o x} V_{T}^{2}\left(W_{3} / L_{3}\right) \exp \left(-\frac{V_{t h 3}-V_{t h 4}}{m V_{T}}\right) . \tag{3-46}
\end{align*}
$$

As clear from (3-45) and (3-46), in order to achieve small power consumption, we have to choose small $k_{2}$ and $k_{3}$. In order to ensure the operation of $\mathrm{M}_{10}$ in the saturation region when $I_{0}=I_{0 M I N}, k_{10}$ must be smaller than $k_{4}$.

Sensitivity to process variations
From (3-44) we can derive that the reference voltage generated by the proposed voltage reference is not process independent, as usually happens in bandgap references. By neglecting matching errors, the sensitivity of the reference voltage is mainly due to the accuracy of the threshold voltage of the diode-connected NMOS transistor $\mathrm{M}_{10}$. In order to achieve a low sensitivity to process variations, as usually required in such kind of applications, we have to minimize the variations of the threshold voltage of $\mathrm{M}_{10}$. Since the variations of the threshold voltage are inversely proportional to $\sqrt{W L}$, the channel length and width of $\mathrm{M}_{10}$ must be set large enough. In such a way an acceptable sensitivity to process variations is achieved, as will be shown later. In any case we have to expect that the standard deviation of the reference voltage generated by the proposed circuit is larger than that of a bandgap reference. On the other hand, bandgap references use one or more operational amplifier [1], [2], which consumes a large power because they have to guarantee a loop gain much larger than unity, and they usually require some resistances, which occupy a large area on the chip. The proposed voltage reference, instead, is a very simple circuit, just 10 transistors, and does not use neither an operational amplifier nor resistances and then we expect it to have a power consumption and an area occupation on the chip much smaller than those of bandgap references.

### 3.5.3. Dynamic Range

The minimum supply voltage is imposed by the current generator circuit. In particular, we have to ensure that $\mathrm{M}_{5}$ operates in the saturation region with $V_{G S 5}<V_{t h 5}\left(V_{t h 5}=-0.75 \mathrm{~V}\right)$ and that $\mathrm{M}_{1}$ has a drain-source voltage of at least 100 mV so that the $V_{D S}$ dependence of the current in $\mathrm{M}_{1}$ can be neglected. Consequently, the following expression has to be satisfied,

$$
\begin{equation*}
V_{D D}>\left|V_{G S S}\right|+V_{D S 1 M I N} . \tag{3-47}
\end{equation*}
$$

Then the supply voltage must be larger then 0.9 V in the AMS $0.35 \mu \mathrm{~m}$ CMOS process. Such voltage is also sufficient to ensure the operation of $M_{4}$ and $M_{8}$ in the saturation region. The maximum supply voltage is imposed by the maximum drainsource voltage allowed for MOS transistors, as shown below,

$$
V_{D D}<\left|V_{\text {DSQMAX }}\right|+V_{R E F} .
$$

Since in the AMS $0.35 \mu \mathrm{~m}$ CMOS process the maximum value for the drain-source voltage of a MOS transistor is 3.3 V , the maximum value of the supply voltage is about 4 V .

### 3.5.4. Temperature Compensation

As a first approximation we can consider that the threshold voltage of an NMOS transistor decreases linearly with the temperature, as shown in (3-5). By differentiating (3-44) with respect to the temperature and taking into account (3-5), one obtains

$$
\frac{\partial V_{R E F}}{\partial T}=-K_{t n}+\frac{m}{N-1} \frac{k_{B}}{q} \sqrt{\frac{W_{4} / L_{4}}{W_{10} / L_{10}}} \ln \left(\frac{W_{3} / L_{3}}{W_{1} / L_{1}}\right) .
$$

As clear from (3-49), the temperature coefficient is independent of the temperature dependence of the carrier mobility. Indeed, in virtue of the topology used, a perfect suppression of the temperature dependence of the mobility is achieved; this leads to a smaller temperature coefficient compared to cases in which the temperature dependence of the mobility is compensated only at the reference temperature [8], degrading the temperature coefficient when moving away from the reference temperature. By setting (3-49) to zero, we obtain the condition

$$
\sqrt{\frac{W_{4} / L_{4}}{W_{10} / L_{10}}}=\frac{K_{t n}(N-1)}{m \frac{k_{B}}{q} \ln \left(\frac{W_{3} / L_{3}}{W_{1} / L_{1}}\right)} .
$$

Therefore, if (3-50) is satisfied, we obtain that the temperature coefficient (3-49) is zero for any temperature. It is clear that this is true within the approximation done in (3-5) and the simplified transistor characteristics (3-1) and (3-6). Such first order condition for a zero temperature coefficient has been calculated by neglecting any second order effects.

### 3.5.5. Second order effects on the temperature coefficient

## Channel length modulation effect

In the calculation of the current $I_{0}$ in (3-43) the channel length modulation effect was neglected. In order to take it into account, the I-V characteristic of a MOS in the saturation region can be written as in (3-1) for $\lambda \neq 0$. Since transistors $M_{1}$ and $M_{2}$ work in the subthreshold region, the drain current is exponentially dependent on the
drain source voltage and then, if their drain source voltage is 3 or 4 times larger than the thermal voltage the channel length modulation effect can be neglected. For transistors $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$, instead, we can use (3-1). In such condition, using the same procedure used to calculate (3-43), the current $I_{0}$ is

$$
\begin{equation*}
I_{0}=\frac{m^{2} V_{T}^{2} k_{4}}{2}\left(\frac{1}{N-1 / \sqrt{1+\lambda V_{D S 4}}}\right)^{2} \ln ^{2}\left(\frac{W_{3} / L_{3}}{W_{1} / L_{1}}\right) \tag{3-51}
\end{equation*}
$$

By dimensioning the channel length of $\mathrm{M}_{4}$ large enough, we can assume that $\lambda V_{D S 4} \ll 1$. In such condition and by using first order Taylor series expansion of $1 / \sqrt{1+x} \cong 1-x / 2$ and of $1 /(1+x)^{2} \cong 1-2 x$, the current $I_{0}$ can be rewritten as follows

$$
I_{0} \cong I_{0 N O M}\left(1-\lambda \frac{V_{D S 4}}{N-1}\right)
$$

where $I_{O N O M}$ is the current $I_{0}$ when the channel length modulation effect is neglected $(\lambda=0)$. By using (3-40) and (3-52) and by using first order Taylor series expansion of $\sqrt{1-x} \cong 1-x / 2$, the reference voltage can be rewritten as

$$
\begin{equation*}
V_{R E F}=V_{R E F 0}-\left(V_{R E F 0}-V_{t h 10}\right) \frac{\lambda}{2} \frac{V_{D S 4}}{N-1}, \tag{3-53}
\end{equation*}
$$

where $V_{\text {REFO }}$ is the reference voltage calculated for $\lambda=0$ and given by (3-44). By using first order Taylor series expansion of $1 /(1-x) \cong 1+x$, the drain-source voltage $V_{D S 4}=V_{D D}-\left|V_{G S 8}\right|$ of $\mathrm{M}_{4}$ can be expressed by,

$$
V_{D S 4}=\left[V_{D D}-\left|V_{t h p}\right|-\left(V_{R E F 0}-V_{t h}\right) \sqrt{\frac{W_{10} / L_{10}}{W_{8} / L_{8}}}\right]\left[1+\frac{\lambda}{2} \frac{\left(V_{R E F 0}-V_{t h}\right)}{N-1} \sqrt{\frac{W_{10} / L_{10}}{W_{8} / L_{8}}}\right],(3-54)
$$

where $V_{t h p}$ is the threshold voltage of a PMOS transistor. By substituting (3-54) in (3-53), by assuming that (3-50) is satisfied, the temperature coefficient of the reference voltage has the following expression,

$$
\frac{\partial V_{R E F}}{\partial T}=-\frac{\lambda K_{t n}}{N-1}\left[V_{D D}-\left|V_{t h 8}\right|-\left(V_{R E F 0}-V_{t h 10}\left(2 \sqrt{\frac{W_{10} / L_{10}}{W_{8} / L_{8}}}-\frac{K_{t p}}{K_{t n}}\right)\right]\right.
$$

where $K_{t p}$ is the threshold voltage temperature coefficient of a PMOS transistor and second order terms have been neglected. By equating (3-55) to zero, we can derive the $\left(W_{10} / L_{10}\right) /\left(W_{8} / L_{8}\right)$ that allows us to achieve a zero temperature coefficient and its expression is given by:

$$
\begin{equation*}
\sqrt{\frac{W_{10} / L_{10}}{W_{8} / L_{8}}}=\frac{K_{t p}}{2 K_{t n}}+\frac{V_{D D}-\left|V_{t h 8}\right|}{2\left(V_{R E F 0}-V_{t h 10}\right)} \tag{3-56}
\end{equation*}
$$

Eq. (3-56) has been calculated for a $V_{D D}$ in the middle of the supply voltage range $\left(V_{D D}=2.45 \mathrm{~V}\right)$. From (3-56) it is evident that if $V_{D D}>2.45 \mathrm{~V}$, the temperature coefficient is negative, otherwise it is positive.

## Body effect

In the proposed voltage reference there is no body effect because the bulk terminals of all NMOS transistors are grounded and the bulk terminals of all PMOS transistors are connected to the supply voltage $V_{D D}$.
Thanks to the topology used and to a proper dimensioning the second order effects are compensated leading to very good temperature coefficient. As in the design proposed in sub-section 2.4, we expect a low temperature coefficient with the advantage that the proposed voltage reference allows sub- 1 V operation.

### 3.5.6. Experimental Results

The proposed voltage reference has been implemented with AMS $0.35 \mu \mathrm{~m}$ CMOS process. The die photograph is shown in Fig. 3-13. Measurements show that the proposed voltage reference generates a mean reference voltage of about 670 mV with a variation of 5.67 mV at room temperature, when the supply voltage varies from 0.9 V to 4 V , as shown in Fig. 3-14a. The power supply rejection ratio, without any filtering capacitor, is -47 dB at 100 Hz and -40 dB at 10 MHz , for the smallest supply voltage. At larger supply voltage the power supply rejection ratio decreases to -53 dB at 100 Hz and to -42 dB at 10 MHz , as shown in Fig. 3-14b. Fig. 3-15 shows the output voltage dependence on temperature for different values


Fig. 3-13: Die Photograph of the Voltage Reference Generator (core)


Fig. 3-14: Experiments: a) Output Voltage vs. Supply voltage at room temperature, b) PSRR at room temperature and for a supply voltage of 2 V .


Fig. 3-15: Measured output voltage vs. temperature for 4 different values of the supply voltage.
of the supply voltage. The measured temperature coefficient at $V_{D D}=2 \mathrm{~V}$ and $V_{D D}=3 \mathrm{~V}$ is $10 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ and $13 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$, respectively, and increases to 18 and 20 $\mathrm{ppm} /{ }^{\circ} \mathrm{C}$ at $V_{D D}=4 \mathrm{~V}$ and $V_{D D}=0.9 \mathrm{~V}$, respectively, corresponding to the maximum and minimum supply voltage. At $80{ }^{\circ} \mathrm{C}$ the current drawn at the maximum supplyvoltage is 70 nA and at the minimum supply voltage is 50 nA . At room temperature, instead, the current drawn at the maximum supply voltage is 55 nA and at the minimum supply voltage is 40 nA . The current drawn from the supply voltage as function of the supply voltage for different values of the temperature is shown in Fig. 3-16. In order to evaluate the sensitivity of the reference voltage to process variations, 20 different samples of the same batch have been tested: the mean value of the reference voltage is 670 mV and the standard deviation is $3.1 \%$. As expected the standard deviation is worse than that of a bandgap reference, which usually is in the order of $1 \%$ [1], [2]. The occupied chip area is $0.045 \mathrm{~mm}^{2}$. A comparison with best performing published voltage reference circuits fabricated with a standard CMOS process is shown in Table 3-I. It can be noted that the voltage reference


Fig. 3-16: Current drawn from the supply voltage vs. supply voltage for different values of the temperature.
generators proposed in [12], [13], [14], described in the sub-sections 3.3, 3.4 and 3.5 , have a temperature coefficient much smaller than that of the others design already reported in the literature, where the temperature dependence of the mobility is compensated only at the reference temperature. Moreover, the voltage reference generators proposed in [13], [14], described in the sub-sections 3.4 and 3.5, have a temperature coefficient much smaller than that of the voltage reference proposed in [12] and described in sub-section 3.3, where second order effects are not compensated. Moreover, the voltage reference proposed in [14] has the minimum supply voltage and by large the smallest power consumption, in the tens of nW range. The PSRR and the line sensitivity are comparable to other solutions already presented in the literature. Moreover, from Table 3-I we can derive that the power consumption and the area occupation on the chip of the voltage reference proposed in [14] are much smaller, by several orders of magnitude in the case of the power consumption and by more than one order of magnitude in the case of the area occupation on the chip, than those of the bandgap reference proposed in [2].

Table 3-I: Comparison with voltage reference generators available in the literature.

|  | De Vita et <br> al. [14] | Leung et <br> al. [8] | Leung et <br> al. [2] | De Vita <br> et al. [12] | De Vita et <br> al. [13] |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Technology | $\begin{aligned} & 0.35 \mu \mathrm{~m} \\ & \text { CMOS } \end{aligned}$ | $\begin{aligned} & 0.6 \mu \mathrm{~m} \\ & \mathrm{CMOS} \end{aligned}$ | $\begin{aligned} & 0.6 \mu \mathrm{~m} \\ & \text { CMOS } \end{aligned}$ | $\begin{aligned} & \hline 0.35 \mu \mathrm{~m} \\ & \text { CMOS } \end{aligned}$ | $0.35 \mu \mathrm{~m}$ <br> CMOS |
| Supply Voltage (V) | 0.9 to 4 | 1.4 to 3 | $\begin{gathered} 0.98 \text { to } \\ 1.5 \\ \hline \end{gathered}$ | 1.5 to 4.3 | 1.5 to 4.3 |
| Supply <br> Current ( $\mu \mathrm{A}$ ) | $\begin{gathered} \hline 0.04 @ 0.9 \mathrm{~V} \\ 0.055 @ 4 \mathrm{~V} \end{gathered}$ | $<9.7$ | <18 | $\begin{aligned} & 1.5 @ 1.5 \mathrm{~V} \\ & 2.4 @ 4.3 \mathrm{~V} \end{aligned}$ | $\begin{aligned} & \text { 0.08@1.5V } \\ & 0.11 @ 4.3 \mathrm{~V} \end{aligned}$ |
| $\mathrm{V}_{\text {ref }}$ | 670 mV | 309.3 mV | 603 mV | 168 mV | 891.1 mV |
| TC (ppm/ ${ }^{\circ} \mathrm{C}$ ) | 10 | 36.9 | 15 | 25 | 12 |
| Line Sensitivity | 0.27 \%/V | 0.08 \%/V | 0.73 \%/V | 0.95 \%/V | 0.46 \%/V |
| PSRR <br> @ 100 Hz <br> @10 MHz | $\begin{array}{r} \mathrm{V}_{\mathrm{DD}}=0.9 \mathrm{~V} \\ -47 \mathrm{~dB} \\ \\ -40 \mathrm{~dB} \end{array}$ | $\begin{array}{r} \mathrm{V}_{\mathrm{DD}}=1.4 \mathrm{~V} \\ -47 \mathrm{~dB} \\ -20 \mathrm{~dB} \end{array}$ | $\begin{gathered} \mathrm{V}_{\mathrm{DD}}=0.98 \mathrm{~V} \\ -44 \mathrm{~dB} \\ -17 \mathrm{~dB} \end{gathered}$ | $\begin{array}{r} \mathrm{V}_{\mathrm{DD}}=1.5 \mathrm{~V} \\ -65 \mathrm{~dB} \\ -57 \mathrm{~dB} \end{array}$ | $\begin{gathered} \mathrm{V}_{\mathrm{DD}}=1.5 \mathrm{~V} \\ -59 \mathrm{~dB} \\ -52 \mathrm{~dB} \end{gathered}$ |
| $\begin{aligned} & \hline \begin{array}{l} \text { Die area } \\ \left(\mathrm{mm}^{2}\right) \end{array} \\ & \hline \end{aligned}$ | 0.045 | 0.055 | 0.24 | 0.08 | 0.015 |

### 3.6. Conclusion

A series of low-voltage, extreme low-power voltage reference generators implemented in AMS $0.35 \mu \mathrm{~m}$ CMOS has been presented. The design conditions to minimize the power consumption and the temperature coefficient are described in detail. The complete suppression of the temperature dependence of mobility in a wide temperature range, the compensation of the channel length modulation effect on the temperature coefficient, and the elimination of the body effect have allowed us to obtain a very small temperature coefficient down to $10 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$. The minimum supply voltage down to only 0.9 V and the maximum quiescent current of only few tens of nA leads to a total absorbed power in the decananowatt range for all proposed designs, that makes the presented circuits very attractive for nanopower applications.

### 3.7. References

[1] B.S. Song, P.R. Gray, "A precision curvature-compensated CMOS bandgap reference," IEEE Journal of Solid State Circuits, vol. DC-18, pp. 634-643, December 1983.
[2] K.N. Leung, P.K.T. Mok, "A sub-1 V $15 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ CMOS Bandgap Voltage Reference without requiring Low Threshold Voltage Device," IEEE Journal of Solid State Circuits, vol. 37, pp. 526-530, April 2002.
[3] R.A. Blauschild, P.A. Tucci, R.S. Muller, R.G. Meyer, "A new NMOS Temperature Stable Voltage Reference," IEEE Journal of Solid State Circuits, vol. SC-13, pp. 767-774, December 1978.
[4] H. Tanaka, Y. Nakagome, J. Etoh, E. Yamasaki, M. Aoki, K. Miyazawa, "Sub$1 \mu \mathrm{~A}$ Dynamic Reference Voltage Generator for battery-operated DRAMs," IEEE Journal of Solid State Circuits, vol. 29, pp. 448-453, April 1994.
[5] M.C. Tobey, D.J. Gialiani, P.B. Askin, "Flat-Band Voltage Reference", U.S. Patent 3975 648, August 1976.
[6] H.J. Oguey, B. Gerber, "MOS Voltage Reference based on polysilicon gate work function difference," IEEE Journal of Solid State Circuit, vol. SC-15, pp. 264-269, June 1980.
[7] K.N. Leung, P.K.T. Mok, K.C. Kwok, "CMOS Voltage Reference," US Patent 6441 680, August 2002.
[8] K.N. Leung, P.K.T. Mok, "A CMOS Voltage Reference Based on Weighted $\Delta \mathrm{V}_{\mathrm{GS}}$ for CMOS Low-Dropout Linear Regulators," IEEE Journal of Solid State Circuits, vol. 38, pp. 146-150, January 2003.
[9] B.-S. Song and P. R. Gray, "Threshold-Voltage Temperature Drift in IonImplanted MOS Transistors," IEEE Journal of Solid State Circuits, SC-17, pp. 291-298, 1982.
[10]H.-J. Song and C.-K. Kim, A Temperature-Stabilized SOI Voltage Reference Based on Threshold Voltage Difference Between Enhancement and Depletion NMOSFET's, IEEE Journal of Solid State Circuits, 28 (1993) 671-677.
[11]H. Banba et al., A CMOS Bandgap Reference Circuit with Sub-1V Operation, IEEE Journal of Solid State Circuits, 34 (1999) 670-674.
[12] G. De Vita, G. Iannaccone, "Ultra-Low-Power, Temperature Compensated Reference Voltage Generator," Proc. of CICC, San Jose, USA, September 2005, pp 751-754.
[13] G. De Vita, G. Iannaccone, P. Andreani, "A $300 \mathrm{nW}, 12 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ Voltage Reference in a Digital $0.35 \mu \mathrm{~m}$ CMOS Process," Proc. of VLSI Symposium on Circuits, Honolulu, USA, June 2006, pp. 100-101.
[14] G. De Vita, G. Iannaccone, "A Sub-1 V, $10 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$, Nanopower Voltage Reference Generator," Proc. of ESSCIRC, Montreaux, Switzerland, pp. 307310, September 2006.

## 4. POWER SUPPLY

### 4.1. Introduction

Long range passive transponders ("tags") for RFID systems must draw the power required for their operation from the electromagnetic field transmitted by the reader [1]. If the transponder lies within the interrogation range of the reader, an alternating RF voltage is induced on the transponder antenna, and is subsequently rectified in order to provide a DC supply voltage for the digital section and to allow data transmission from the tag to the reader through modulation of the backscattered radiation. The practical adoption in the commercial arena of passive UHF and Microwave RFID (Radio Frequency IDentification) transponders for identification and ambient intelligence applications [2] is strongly affected by their operating range. The deployment and the cost of the ensemble of readers are obviously dependent on it, but also the very nature of the application and the user's experience can be deeply transformed by increasing the maximum distance at which a transponder can be addressed.

Their operating range is determined by the power at the transponder antenna required for its operation and by the maximum power transmitted by the reader, which is limited by European regulations to just 500 mW in the frequency ranges of interest to us ( $868 / 916 \mathrm{MHz}$ and 2.45 GHz ) [3]. A focused effort in the attempt to drastically reduce power consumption of the transponder is therefore required. In order to achieve an operating range of several meters, as required by many applications, it is necessary to reduce the power consumption of both the analog and digital sections of the tag to the $\mu \mathrm{W}$ or sub- $\mu \mathrm{W}$ regime, and to optimize the power efficiency of the voltage multiplier and the modulation depth [4].

Therefore, the voltage regulator must be able to supply a DC power of only few $\mu \mathrm{W}$, must have a very small quiescent current, in the order of few tens of nA, and must be able to provide the regulated voltage with a voltage drop as small as possible, in order to minimize the amplitude of the unregulated voltage required for the correct operation of the transponder.

To our knowledge, there are no examples in the literature of series regulators with similar specifications, probably because the application we envisage involves an extremely low power. Indeed, the lowest power voltage regulators presented in literature provide output currents of some tens of mA and have quiescent currents of a few tens of $\mu \mathrm{A}$ [5], [6]. As can be seen, they would absorb the whole power budget of a passive long distance RFID transponder.


Fig. 4-1: Passive Transponder Architecture.

### 4.2. Architecture of a Passive RFID Transponder

The architecture of a complete passive RFID transponder is shown in Fig. 4-1. The coupling element is typically a dipole or a patch antenna. A voltage multiplier converts the input alternating voltage into a DC voltage which is used by a series voltage regulator to provide the regulated voltage required for the correct operation of the transponder. The voltage multiplier is matched with the antenna in order to ensure the maximum power transfer from the transponder antenna to the input of the voltage multiplier. A backscatter modulator is used to modulate the impedance seen by the transponder antenna during transmission. The RF section is then connected to the digital section, which typically is a very simple microprocessor or a finite state machine running the communication protocol.
The input voltage $V_{\text {DDlow }}$, which has to provide almost all the power required for transponder operation, is generated by a single stage voltage multiplier, which is shown to provide the maximum conversion efficiency [4]. For the digital section, that we shall not discuss here, we consider an implementation based on subthreshold CMOS logic, with a standard $0.35 \mu \mathrm{~m}$ CMOS process, a regulated supply voltage $V_{R E G}=0.6 \mathrm{~V}$, and a clock frequency $<1 \mathrm{MHz}$, which enables sub- $\mu \mathrm{W}$ power consumption. In subthreshold circuits the supply voltage is scaled down below the threshold voltage and then the load capacitances are charged and discharged by subthreshold leakage currents. Leakage currents are order of magnitude lower than
drain currents in the strong inversion regime, so there is a significant limit on the maximum performance of subthreshold circuits. Therefore, traditionally, subthreshold circuits have been used for applications that require ultra-low power dissipation with moderate circuit performance, as in the case of passive RFID transponders. Recently, subthreshold circuits are becoming popular in emerging embedded applications, such as wireless sensor networks, where the key metric is minimizing energy dissipation rather than high speed performance.
The I-V characteristic of a MOS transistor in the subthreshold region is given by,

$$
\begin{equation*}
I_{D}=\mu C_{o x} V_{T}^{2} \frac{W}{L} \exp \left(\frac{V_{G S}-V_{t h}}{m V_{T}}\right)\left[1-\exp \left(-\frac{V_{D S}}{V_{T}}\right)\right], \tag{4-1}
\end{equation*}
$$

Because of the exponential relationship, the drain current and then the performance of the subthreshold digital circuit, i.e. switching time and oscillation frequency, are very sensitive to temperature and threshold voltage variations.
The propagation delay $t_{p}$ of a CMOS inverter is given by,

$$
\begin{equation*}
t_{p}=\frac{C}{I_{D}} \frac{V_{D D}}{2}, \tag{4-2}
\end{equation*}
$$

where $C$ is the load capacitance, $I_{D}$ is given by (4-1) and $V_{D D}$ is the supply voltage. If $V_{D D}$ was constant, as happens in CMOS circuits operating in the strong inversion regime, the large variations of $I_{D}$ due to temperature or process variations would strongly affect the performance of the subthreshold digital circuit.
For such reason, in this chapter, first, we present a voltage regulator with a negative temperature coefficient that compensates the temperature dependence of the subthreshold current, ensuring that the current in the logic gates, and therefore the delays and the power consumption of the subthreshold logic, and the frequency of the clock generator, are almost independent of temperature. Then, we present a voltage regulator that provides a supply voltage $V_{D D}$ so that the ratio $t_{p}$ is as independent as possible of temperature and process variations. In such a way, the propagation delay and then the performance of the subthreshold digital circuit becomes almost insensitive to temperature and process variations.

Moreover, a power consumption of few $\mu \mathrm{W}$ is required in order to achieve an operating range of the tag-reader system of several meters [4], which would allow us to significantly broaden the possible range of applications and reduce system deployment costs. In order to ensure the correct operation of the voltage regulator, a supply voltage $V_{D D l o w}$ few tens of mV larger than $V_{R E G}$ is sufficient. However, such voltage would not be sufficient to ensure the correct operation of the error amplifier and to generate the reference voltage $V_{\text {REF }}$. As a consequence, a second stage is


Fig. 4-2 : Circuit of the voltage regulator.
added to generate the higher voltage $V_{\text {DDhigh }}$ that has only to provide the very small power required for the error amplifier and the voltage reference generator. In such a way, power efficiency can be maximized.

In section 4.3 we propose a temperature compensated series voltage regulator. In section 4.4 we present a very low-power voltage regulator with a negative temperature coefficient to reduce the temperature sensitivity of subthreshold circuits. Finally, in section 4.5 we present a voltage regulator that keeps the delay of a logic gate constant with a low sensitivity to temperature and process variations.

### 4.3. Temperature compensated voltage regulator

### 4.3.1. Circuit description

The proposed series voltage regulator is shown in Fig. 4-2. The series voltage regulator consists of a differential amplifier that compares a fraction of the output voltage with the reference voltage - generated by the voltage reference generator previously described in section 3.4- and produces an error signal that drives the gate of an NMOS transistor, in order to keep the output voltage constant and equal to the reference voltage [5], [6].

A variation of the supply voltage affects the output regulated voltage through the bias current $I_{0}$ and through the transistor $\mathrm{M}_{\mathrm{U}}$. Then, we can calculate the two
components separately. Assuming the drain voltage of $\mathrm{M}_{\mathrm{U}}$ constant, the effect of the variation of the bias current on the output regulated voltage is, at DC,

$$
\begin{equation*}
\left.\frac{\Delta V_{R E G}}{\Delta I_{0}}\right|_{D C} \cong \frac{\frac{r_{d n} \| r_{d p}}{2} \frac{g_{m U} R_{L}}{1+g_{m U} R_{L}}}{1+\frac{r_{d n} \| r_{d p}}{2} g_{m} \frac{g_{m U} R_{L}}{1+g_{m U} R_{L}}} \cong \frac{1}{g_{m}} \tag{4-3}
\end{equation*}
$$

where $r_{d n} \mathrm{e} r_{d p}$ are the drain-source resistances of the NMOS and PMOS transistors of the differential amplifier, $g_{m}$ and $g_{m U}$ are the transconductances of the sourcecoupled transistors ( $M_{3}$ and $M_{4}$ in Fig. 4-2) and of the pass transistor $M_{U}$, respectively, and $R_{L}$ is the load resistance. Since the bias current is very small, in order to ensure very small power consumption, the transconductance of the sourcecoupled transistors is small causing a strong dependence of the regulated voltage on the bias current. As a consequence, in order to obtain a good PSRR, the differential amplifier is biased with a current almost independent of the supply voltage, obtained by mirroring the reference current $I_{0}$, generated in the voltage reference generator.

Assuming the bias current constant, instead, the effect of the variations of the supply voltage $V_{D D}$ on the output regulated voltage is given by

$$
\begin{equation*}
\left.\frac{\Delta V_{R E G}}{\Delta V_{D D}}\right|_{D C}=\frac{\frac{R_{L} / r_{d U}}{1+g_{m_{U}} R_{L}+R_{L} / r_{d U}}}{1+\frac{r_{d n} \| r_{d n}}{2} g_{m}} \cong \frac{2}{g_{m} r_{d n} \| r_{d p}} \frac{1}{g_{m U} r_{d U}} . \tag{4-4}
\end{equation*}
$$

To minimize the expression in (4-4) a high gain of the differential amplifier and of the output transistor $\mathrm{M}_{\mathrm{U}}$ is required.

It is also important to ensure a high PSRR at the operation frequency of the reader-transponder system ( 2.45 GHz or $868 / 916 \mathrm{MHz}$ ), in order to drastically attenuate the effect of the ripple superimposed on $V_{D D}$. In order to achieve such aim, a capacitance $C_{u}$ of 5 pF is put at the output of the series voltage regulator. The stability is a critical aspect for such kind of circuits because the pole associated with the output node could be quite low affecting the stability of the feedback amplifier. Indeed, since the output pole is given by $g_{m u} / C_{u}$, for small load current the output pole becomes quite small. For such reason a frequency compensation is required to avoid oscillation in the feedback amplifier, and is obtained by placing a capacitance $C_{c}$ of 50 pF at the output of the differential amplifier, in order to ensure that the pole associated with the capacitance $C_{c}$ is much lower than the output pole also for
small load current, so that the amplifier has a dominant pole. With that choice of the capacitance $C_{c}$ the output pole is about 80 Hz and is more than two orders of magnitude lower than the output pole for a load current one hundredth of the maximum value $(8 \mu \mathrm{~A})$. In order to reduce the area occupation on the chip, the capacitances $C_{c}$ and $C_{u}$ are implemented by using stack capacitors, which exploit the parallel between the poly2/polyl capacitance and poly1/n-well capacitance.

The minimum and maximum $V_{D D}$ required for the correct operation of the series voltage regulator are imposed by the reference voltage generator and are

$$
\begin{aligned}
& V_{D D M I N}=\left|V_{G S 6}\right|+V_{D S 2 M I N}+V_{G S 4} \cong 1.5 \mathrm{~V}, \\
& V_{D D M A X}=\left|V_{D S 5 M A X}\right|+V_{G S 1}+V_{G S 3} \cong 4.3 \mathrm{~V} .
\end{aligned}
$$

In the typical case in which $V_{D D}$ is provided by a voltage multiplier, a voltage limiter to 4.3 V and a large capacitance ( 500 pF ) are placed at the output of the voltage multiplier. The former ensures that $V_{D D}<V_{D D M A X}$, the latter keeps $V_{D D}$ almost constant when the RF signal at the input of the voltage multiplier is ASK modulated [7].

### 4.3.2. Temperature Coefficient

The output regulated voltage is a partition of the reference voltage through the voltage divider formed by $\mathrm{M}_{13}$ and $\mathrm{M}_{14}$ and its expression is given by

$$
\begin{equation*}
V_{R E G}=\sqrt{\frac{k_{13}}{k_{14}}}\left(V_{\text {REF }}-V_{t h n}\right)+V_{R E F}+\left|V_{\text {thp }}\right|, \tag{4-5}
\end{equation*}
$$

where, $V_{t h n}$ and $V_{t h p}$ are the threshold voltages of an NMOS and PMOS, respectively. By differentiating (4-5) with respect to the temperature and by setting it to zero, one can obtain a condition for the temperature coefficient of the reference voltage $V_{R E F}$ :

$$
\begin{equation*}
\frac{\partial V_{R E F}}{\partial T}=\frac{K_{t p}-\sqrt{k_{14} / k_{13}} K_{t n}}{1+\sqrt{k_{14} / k_{13}}}, \tag{4-6}
\end{equation*}
$$

where $K_{t n}$ and $K_{t p}$ are the temperature coefficients of the threshold voltage of the NMOS and PMOS, respectively. For the voltage reference generator, described in section 3.4, the temperature coefficient of the output reference voltage is given by,

$$
\begin{equation*}
\frac{\partial V_{R E F}}{\partial T}=-K_{t 1}+\left(1+\sqrt{\frac{W_{10} / L_{10}}{W_{9} / L_{9}}}\right) \sqrt{\frac{h}{W_{8} / L_{8}}} \frac{k_{B}}{q}-\sqrt{\frac{h}{W_{7} / L_{7}}} \frac{k_{B}}{q} \tag{4-7}
\end{equation*}
$$



Fig. 4-3: Die Photograph of the Series Voltage Regulator (core).


Fig. 4-4: a) Output voltage vs. $V_{d d}$ for an output current of $2 \mu \mathrm{~A}, \mathrm{~b}$ ) Output voltage vs. Output current for $V_{d d}=3 \mathrm{~V}$.

From (4-6) and (4-7), we can derive the ratio $\left(W_{10} / L_{10}\right) /\left(W_{9} / L_{9}\right)$ that allows us to achieve the temperature compensation.

### 4.3.3. Experimental Results

The proposed voltage regulator was successfully implemented in AMS $0.35 \mu \mathrm{~m}$ CMOS technology. The die photograph is shown Fig. 4-3. The measurements show that the mean output regulated voltage is 600.1 mV with a line sensitivity of $0.19 \% / \mathrm{V}$. The output regulated voltage as function of $V_{D D}$ is plotted in Fig. 4-4a. Furthermore, the output voltage varies from 599.6 mV to 600.26 mV when the output current varies from 0 to $8 \mu \mathrm{~A}$, leading to an equivalent DC output resistance of about $82.5 \Omega$ (a few thousand times smaller than the equivalent load resistance). The output regulated voltage as a function of the output current is plotted in Fig. $4-4 \mathrm{~b}$. The measurements also show a DC PSRR of -52 dB and a PSRR at 2.45 GHz of - 65 dB . At room temperature the quiescent current for $V_{D D}=1.5 \mathrm{~V}$ and $V_{D D}=4.3 \mathrm{~V}$ is about 90 nA and 120 nA , respectively. At $80^{\circ} \mathrm{C}$, instead, the quiescent current for $V_{D D}=1.5 \mathrm{~V}$ and $V_{D D}=4.3 \mathrm{~V}$ is about 120 nA and 150 nA , respectively. Fig. 4-5a shows the output voltage as function of the temperature. The temperature coefficient for $V_{D D}=2 \mathrm{~V}$ and $V_{D D}=3 \mathrm{~V}$ is $27 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ and $23 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$, respectively, and increases to $35 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ both for $V_{D D}=1.5 \mathrm{~V}$ and $V_{D D}=4.3 \mathrm{~V}$. Furthermore, as shown in Fig. 4-5b, when applying a load current pulse from 0 to $8 \mu \mathrm{~A}$, a settling time of $420 \mu \mathrm{~s}$ and a maximum variation of the output regulated voltage of 47 mV are obtained. In order to evaluate how fast the proposed series voltage regulator reacts to a variation of the RF power from no RF signal to a large value, a supply voltage pulse from 0 to 3 V was applied: measurements show that a settling time of about 1.5 ms is required. After such settling time, the storage capacitance is charged and the supply voltage remains almost constant during reception of the ASK modulated signal. The occupation of area on the chip is $0.1 \mathrm{~mm}^{2}$.

### 4.4. Negative-Temperature coefficient voltage regulator

The circuit used to generate the reference voltage is shown in Fig. 4-6. It consists of a circuit that generates a current $I_{0}$ almost independent of the supply voltage, which is in turn injected into a diode, to generate the reference voltage [8]. In the current generator, in order to ensure that the currents in the two branches are as close as possible when the supply voltage $V_{D D h i g h}$ varies, a cascode current mirror is used. We choose to use a bipolar rather than a CMOS cascode current mirror because in


Fig. 4-5: a) Output Voltage vs. temperature, b) Load current pulse response.


Fig. 4-6: Schematic of the voltage regulator.
our IC technology PMOS transistors have a too large threshold voltage ( -0.75 V ) leading to a strong reduction of the dynamic range of the voltage regulator. Assuming identical currents in the two branches, the current $I_{0}$, in Fig. 2, has the expression shown below,

$$
\begin{equation*}
I_{0}=\frac{2}{R^{2}}\left(\frac{1}{\sqrt{k_{1}}}-\frac{1}{\sqrt{k_{2}}}\right)^{2} \tag{4-8}
\end{equation*}
$$

where $k_{i}=\mu_{n} C_{o x} W_{i} / L_{i}$ for transistor $\mathrm{M}_{\mathrm{i}}$. In order to show the importance of using a cascode current mirror in the circuit that generates the current $I_{0}$, let us for a moment assume to use instead a simple current mirror: considering the Early effect of bipolar transistors, the ratio $m$ of the output to the input current of the current mirror would be given by

$$
\begin{equation*}
m=\left(1+\frac{V_{D D}-V_{G S 1}}{V_{A}}\right) /\left(1+\frac{V_{\gamma}}{V_{A}}\right) \tag{4-9}
\end{equation*}
$$

where $V_{A}$ is the Early voltage of the bipolar transistors and $V_{G S I}$ is the gate-source voltage of $\mathrm{M}_{1}$. In such a condition the current would have the expression shown below,

$$
\begin{equation*}
I_{0}=\frac{2}{R^{2}}\left(\sqrt{\frac{m}{k_{1}}}-\frac{1}{\sqrt{k_{2}}}\right)^{2} \tag{4-10}
\end{equation*}
$$

When the supply voltage varies between its minimum and maximum values, the ratio $m$ would vary between $m_{M A X}$ and $m_{M I N}$, causing a variation of the current $I_{0}$. As a consequence, the relative variation of the current due to the Early effect would be given by

$$
\begin{equation*}
\frac{\Delta I_{0}}{I_{0}}=\frac{\left(\sqrt{m_{M A X}}-\sqrt{k_{1} / k_{2}}\right)^{2}-\left(\sqrt{m_{M I N}}-\sqrt{k_{1} / k_{2}}\right)^{2}}{\left(1-\sqrt{k_{1} / k_{2}}\right)^{2}} \tag{4-11}
\end{equation*}
$$

As the supply voltage has large variations, we would have a large relative variation of the reference current leading to an unacceptable PSRR of the series voltage regulator. Therefore, we prefer to use a cascode current mirror even if it leads to a reduction of the dynamic range. Since the voltage reference generator has to consume only a negligible fraction of the total power dissipated by the transponder, the current $I_{0}$ has to be set in the order of few nA. In order to take into account the
channel length modulation we can use for the drain current of a MOS transistor in saturation region the expression shown below,

$$
I_{D}=\frac{k}{2}\left(V_{G S}-V_{t h}\right)^{2}\left(1+\lambda V_{D S}\right)
$$

where $\lambda$ is the channel length modulation coefficient approximately inversely proportional to the channel length. Using such expression for the drain current of $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ and assuming small values of $\lambda$ so that $\lambda V_{D S} \ll 1$, the relative variation of the current $I_{0}$, when $V_{\text {DDhigh }}$ varies between its minimum value $V_{\text {DDhighmin }}$ and its maximum value $V_{D D h i g h M A X}$ can be well approximated by

$$
\begin{equation*}
\frac{\Delta I_{0}}{I_{0}}=\lambda \frac{V_{\text {DDhighMAX }}-V_{\text {DDhighMIN }}}{\sqrt{k_{2} / k_{1}}-1} . \tag{4-13}
\end{equation*}
$$

From (4-13) we notice that, in order to have a small dependence of the current $I_{0}$ on the supply voltage, the ratio $k_{l} / k_{2}$ can not be too close to one and the channel length has to be chosen sufficiently large so that $\lambda$ is small enough. On the other hand, the ratio $k_{I} / k_{2}$ can not be too small otherwise an excessively large resistance $R$ should be required to have a small $I_{0}$, as can be seen from (4-8). A good tradeoff is found by setting $k_{1} / k_{2}=0.5$ and $R=2 \mathrm{M} \Omega$. Such a resistance was implemented in high resistive poly so that its area occupation is not too large compared to the area occupation of the entire transponder. Such a choice ensures, at the same time, a very good DC PSRR for the voltage regulator and an acceptable area occupation.

Since the reference voltage generator has two stable states, corresponding to the current given by (4-8) and to zero current, a start-up circuit (formed by $\mathrm{M}_{1 \mathrm{~S}}-\mathrm{M}_{3 \mathrm{~S}}$ ) is used to ensure that the former is reached.

### 4.4.1. Series Voltage Regulator

The series voltage regulator consists of a differential amplifier that compares the output voltage with the reference voltage and provides an error signal driving the gate of an NMOS transistor, in order to keep the output voltage equal to the reference voltage [5], [6]. A PMOS differential amplifier with an active NMOS load is used, and - in order to make its operation as independent as possible of supply voltage variations - is biased with the current $I_{0}$, previously generated, since the differential amplifier is not able to attenuate the variation of the bias current caused by the variations of the supply voltage $V_{\text {DDhigh }}$. Indeed, the effect of the variation of the bias current on the output regulated voltage, at DC, is given by

$$
\left.\frac{\Delta V_{R E G}}{\Delta I_{0}}\right|_{D C} \cong \frac{\frac{r_{d n} \| r_{d p}}{2} \frac{g_{m U} R_{L}}{1+g_{m U} R_{L}}}{1+\frac{r_{d n} \| r_{d p}}{2} g_{m} \frac{g_{m U} R_{L}}{1+g_{m U} R_{L}}} \cong \frac{1}{g_{m}}
$$

where, $r_{d n} \mathrm{e} r_{d p}$ are the drain-source resistances of the NMOS and PMOS transistors of the differential amplifier, $g_{m}$ and $g_{m U}$ are the transconductances of the sourcecoupled transistors, $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$, and of the pass transistor, $\mathrm{M}_{\mathrm{U}}$, respectively, and $R_{L}$ is the load resistance.

Since the bias current is very small, in order to ensure very small power consumption, the transconductance of the source-coupled transistors is small causing a strong dependence of the regulated voltage on the bias current. As a consequence, in order to obtain a good PSRR, it is necessary to bias the differential amplifier with a current almost independent of the supply voltage. As already said such a current is obtained by mirroring the reference current previously generated. Such a choice ensures that the bias current of the differential amplifier is almost independent of the supply voltage even if a simple current mirror is used to have a larger dynamic range. Furthermore, the channel lengths of the two source-coupled transistors, $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$, are set to be large enough ( $10 \mu \mathrm{~m}$ ) to reduce the channel length modulation effect and then make the differential amplifier almost independent of the supply voltage. In such a way a good PSRR with respect to the supply voltage $V_{D D h i g h}$ is achieved. The DC PSRR with respect to the supply voltage $V_{\text {DDlow }}$, instead, is given by,

$$
\begin{equation*}
\left.\frac{\Delta V_{R E G}}{\Delta V_{D D l o w}}\right|_{D C} \cong \frac{\frac{R_{L} / r_{d U}}{1+g_{m U} R_{L}+R_{L} / r_{d U}}}{1+\frac{r_{d n} \| r_{d p}}{2} g_{m}} \cong \frac{2}{g_{m} r_{d n} \| r_{d p}} \frac{1}{g_{m U} r_{d U}} . \tag{4-15}
\end{equation*}
$$

To minimize the expression in (4-15) we need to maximize the denominator, which is the product of the intrinsic gains of the differential amplifier and of the output stage: therefore, we choose a large channel width for both the output transistor $\mathrm{M}_{\mathrm{U}}(50 \mu \mathrm{~m})$ and the NMOS transistors of the active load ( $80 \mu \mathrm{~m}$ ). In such a way we can avoid the use of a second amplification stage that would lead to an increase in the power consumption. It is also important to ensure a high PSRR at the operation frequency of the reader-transponder system $(2.45 \mathrm{GHz}$ or $868 / 916$ MHz ), in order to drastically attenuate the effect of the ripple superimposed to $V_{\text {DDhigh }}$ and $V_{\text {DDlow }}$. In order to reach such objective, a capacitance $C_{u}$ of 5 pF is put at the output of the series voltage regulator. The stability is a critical aspect for such
kind of circuits because the pole associated with the output node could be quite low affecting the stability of the feedback amplifier. Indeed, since the output pole is given by $g_{m u} / C_{u}$, for small load current the output pole becomes quite small. For such reason a frequency compensation is required to avoid oscillation in the feedback amplifier, and is obtained by placing a capacitance $C_{C}$ of 50 pF at the output of the differential amplifier, in order to ensure that the pole associated with the capacitance $C_{C}$ is much lower than the output pole also for small load current, so that the amplifier has a dominant pole. With that choice of the capacitance $C_{C}$ the output pole is about 70 Hz and is more than two orders of magnitude lower than the output pole for a load current one hundredth of the maximum value ( $5 \mu \mathrm{~A}$ ). In order to reduce the area occupation on the chip, the capacitance $C_{C}$ and $C_{U}$ are implemented by using stack capacitors, which exploit the parallel between the poly2/poly1 capacitance and poly/substrate capacitance.

### 4.4.2. Supply Voltage Range

The minimum and maximum $V_{\text {DDlow }}$ required for the correct operation of the series voltage regulator are imposed by the output NMOS transistor, $\mathrm{M}_{\mathrm{U}}$, and are

$$
V_{\text {DDlowMIN }}=V_{R E G}+V_{D S s a t} \cong 0.63 \mathrm{~V}, \quad V_{\text {DDlowMAX }}=V_{R E G}+V_{D S M a x} \cong 5.6 \mathrm{~V},
$$

where, $V_{D S s a t}$ is the minimum drain-source voltage required to ensure the operation of $\mathrm{M}_{\mathrm{u}}$ in the saturation region and $V_{D S \max }$ is the maximum drain-source voltage that ensures the safe operation of $\mathrm{M}_{\mathrm{u}}$. The minimum and maximum $V_{\text {DDhigh }}$ required for the correct operation of the series voltage regulator are imposed by the reference voltage generator and are

$$
V_{\text {DDhighMIN }}=2 V_{\text {BEon }}+V_{D S 2 \text { sat }} \cong 1.4 \mathrm{~V}, \quad V_{\text {DDhighMAX }}=V_{\text {BEon }}+V_{G S 1}+V_{\text {CEMax }} \cong 6.3 \mathrm{~V},
$$

where, $V_{D S 2 \text { sat }}$ is the minimum drain-source voltage required to ensure the operation of $\mathrm{M}_{2}$ in the saturation region, $V_{\text {CEmax }}$ is the maximum drain-source voltage that ensures the safe operation of pnp transistors and $V_{\text {BEon }}$ is the base-emitter voltage when the diode is directly biased. As a consequence, the voltage at the output of the first stage of the voltage multiplier, which generates the supply voltage $V_{\text {DDlow }}$, is upper limited to a 1.4 V and the voltage at the output of the second stage of the voltage multiplier, which generates the supply voltage $V_{D D h i g h}$, is upper limited to 6.3 V. Furthermore, since the signal received by the transponder is an ASK modulated signal [7], the time constant of the voltage multiplier must be chosen large enough so that the two supply voltages $V_{D D h i g h}$ and $V_{\text {DDhigh }}$ remain almost constant when a symbol ' 0 ' is received and then no RF signal is present at the input of the transponder. For such reason, a large capacitance ( 500 pF ) is put at the two outputs of the voltage multiplier which is able to store the energy necessary to
ensure the correct operation of the series voltage regulator when there is no RF energy because a ' 0 ' is received.

### 4.4.3. Temperature Coefficient

Since the digital section is implemented in subthreshold CMOS logic, the static and dynamic currents in the logic gates are exponentially increasing with temperature, and so would be the power consumption if the supply voltage was regulated with respect to temperature variations. Such behavior would lead to an increase in the power consumption of the passive transponder and then to a reduction of the operating range. For such reason we choose to accept the negative temperature coefficient of $-2 \mathrm{mV} /{ }^{\circ} \mathrm{C}$ given by the p-n junction that generates the reference voltage. In such a way we have a significant advantage: the temperature coefficient of the supply voltage almost exactly compensates the temperature dependence of the subthreshold current, ensuring that the current in the logic gates, and therefore the performance and the power consumption of the subthreshold logic, are practically independent of temperature. The I-V characteristics of a MOS transistor in the subthreshold region can be approximated by (4-1). In subthreshold CMOS logic the delay of a gate is given by the time required to charge (discharge) its output capacitance with a current given by (4-1). Let $T_{P}$ be the propagation delay. If we consider for simplicity an inverter supplied by $V_{D D}$ and with an output capacitance $C_{O U T}$, we have $T_{P}=0.5 C_{\text {OUT }} V_{D D} / I_{D}$. By assuming $V_{D S} \gg V_{T}$, from ( 4-1 ) we can derive that, at temperature $T$, the current that charges and discharges $C_{\text {OUT }}$ is given by,

$$
\frac{I_{D}(T)}{I_{D}\left(T_{0}\right)}=A\left(\frac{T}{T_{0}}\right)^{2+\mu_{T}} \exp \left[\frac{V_{D D 0}-V_{t h 0}-(h+k) T_{0}}{m V_{T}}\right],
$$

where $A$ is a temperature independent coefficient, $\mu_{T}$ is the mobility temperature exponent [9], $V_{D D 0}$ and $V_{\text {th0 }}\left(V_{D D 0}<V_{\text {th0 }}\right)$ are the supply voltage and threshold voltage at the reference temperature $T_{0}, h$ and $k$ are the temperature coefficient of the supply voltage and the threshold voltage, respectively. If the supply voltage was regulated in temperature ( $h=0$ ), the two temperature dependent terms of (4-16) would increase with temperature leading to a strong variation of the current and then of the propagation delay. In the case of the proposed voltage regulator, the supply voltage has a temperature coefficient of $h=-2 \mathrm{mV} /{ }^{\circ} \mathrm{C}$, due to the $\mathrm{p}-\mathrm{n}$ junction that generates the reference voltage; in such a way, the exponential term of (4-16) decreases when temperature increases, leading to an overall decrease of the current. From (4-16) we can obtain that $I_{D}\left(100{ }^{\circ} \mathrm{C}\right) / I_{D}\left(0^{\circ} \mathrm{C}\right)$ is 0.619 for $h=-2 \mathrm{mV} /{ }^{\circ} \mathrm{C}$ and 2.532 for $h=0$.


Fig. 4-7: a) Ring oscillator supplied by a temperature independent supply voltage, b) Ring oscillator supplied by the proposed voltage regulator.

The relative variation of the propagation delay when the temperature varies from 0 to $100^{\circ} \mathrm{C}$ is given by,

$$
\begin{equation*}
\frac{\Delta T_{P}}{T_{P}\left(0^{\circ} \mathrm{C}\right)}=\frac{V_{D D}\left(100^{\circ} \mathrm{C}\right)}{V_{D D}\left(0^{\circ} \mathrm{C}\right)} \frac{I_{D}\left(0^{\circ} \mathrm{C}\right)}{I_{D}\left(100^{\circ} \mathrm{C}\right)}-1 \tag{4-17}
\end{equation*}
$$

and its value is $7.7 \%$ for $h=-2 \mathrm{mV} /{ }^{\circ} \mathrm{C}$ and $-60.5 \%$ for $h=0$. It is clear that better temperature compensation could be achieved by an ad-hoc regulator to perfectly cancel the temperature dependence but it would be much more complicate than a simple $V_{B E}$ regulator and then it would probably consume much more power. The proposed regulator, instead, despite its simplicity, allows us to achieve good temperature compensation with very small power consumption.
To validate such approach we can consider a ring oscillator implemented with three NAND gates (with the input terminals connected to operate as inverters) in subthreshold CMOS logic and comparing the performance obtained when the circuit is supplied by a DC voltage $V_{D D}=0.6 \mathrm{~V}$, independent of temperature (Fig. $4-7 \mathrm{a}$ ), and that obtained when it is supplied by the proposed voltage regulator (circuit shown in Fig. 4-7b), providing a temperature dependent DC voltage $V_{R E G}$. The dynamic power dissipated by a logic gate can be written as $P=C_{\text {OUT }} f V_{D D}^{2}$, where $C_{\text {OUT }}$ is the output capacitance, $f$ is the frequency. By assuming that we can neglect the variations of $C_{\text {OUT }}$ with the temperature, in the case of the circuit of Fig. 4-7a we have

$$
\begin{equation*}
\frac{\Delta P}{P} \cong \frac{\Delta f}{f} \tag{4-18}
\end{equation*}
$$

while in the case of the circuit of Fig. $4-7 \mathrm{~b}$ we have


Fig. 4-8 : Die Photograph (core).

$$
\begin{equation*}
\frac{\Delta P}{P} \cong 2 \frac{\Delta V_{R E G}}{V_{R E G}}+\frac{\Delta f}{f} . \tag{4-19}
\end{equation*}
$$

Since in a first approximation the operating range of the tag-reader system is inversely proportional to the square root of the power dissipated by the digital section [4], the relative variation of the operating range, when the temperature varies, is given by

$$
\begin{equation*}
\frac{\Delta r}{r}=-\frac{1}{2} \frac{\Delta P}{P} . \tag{4-20}
\end{equation*}
$$

Equation (4-20) allows us to verify how the performance of the tag-reader system, in terms of operating range, is affected by temperature variations in the two cases studied.

### 4.4.4. Experimental Results

The proposed voltage regulator was successfully implemented in AMS $0.35 \mu \mathrm{~m}$ BiCMOS with a die area occupation of $0.025 \mathrm{~mm}^{2}$. The die photograph is presented in Fig. 4-8. Measurements show that the mean output regulated voltage is 605 mV with a line sensitivity of $\pm 0.8 \mathrm{mV} / \mathrm{V}$ with respect to $V_{D D h i g h}$ and of $\pm 0.18 \mathrm{mV} / \mathrm{V}$ with respect to $V_{\text {DDlow }}$. The output regulated voltage as a function of $V_{D D l o w}$ and $V_{D D h i g h}$ is


Fig. 4-9: a) Output voltage vs. $V_{\text {ddlow }}$ for 3 values of $V_{\text {ddigh }}$ for an output current of 2 $\mu \mathrm{A}, \mathrm{b})$ Output voltage vs. Output current for $V_{\text {ddlow }}=1 \mathrm{~V}$ and $V_{\text {ddhigh }}=3 \mathrm{~V}$.


Fig. 4-10: a) PSRR for $\mathrm{V}_{\text {ddlow }}$, b) PSRR for $\mathrm{V}_{\text {ddhigh }}$ of the series voltage regulator.


Fig. 4-11: Output regulated voltage when a load current pulse is applied.
shown in Fig. 4-9a. Furthermore, the output voltage varies from 605.5 mV to 605 mV when the output current varies from 0 to $5 \mu \mathrm{~A}$, leading to an equivalent DC output resistance of about $100 \Omega$ (a few thousand times smaller than the equivalent load resistance). The output regulated voltage as a function of the output current is shown in Fig. 4-9b. Measurements also show a very good PSRR for both $V_{\text {DDlow }}$ and $V_{\text {DDhigh }}$ : Fig. 4-10 (right) shows, with respect to $V_{\text {DDhigh }}$, a DC PSRR of -58.5 dB and a PSRR at the RF operating frequency of -54 dB . Fig. 4-10 (left) shows, with respect to $V_{\text {DDlow }}$, a PSRR in DC of -78 dB and a PSRR at the RF operating frequency of -57 dB . The quiescent current is about 34 nA . Such a current is provided by $V_{D D h i g h}$ for the operation of the error amplifier and to generate the reference voltage. As a consequence, the quiescent power consumption, for the minimum supply voltage $V_{\text {DDhigh }}$ that allows the correct operation of the series voltage regulator, is about 48 nW . Furthermore, when applying a load current pulse from 0 to $5 \mu \mathrm{~A}$, a settling time of $480 \mu \mathrm{~s}$ and a maximum variation of the output regulated voltage of 45 mV are obtained, as shown in Fig. 4-11. In order to evaluate how fast the proposed series voltage regulator reacts to a variation of the RF power from no RF signal to a large value, supply voltage pulses from 0 to 3 V and from 0 to 1 V were applied to $V_{\text {DDhigh }}$ and $V_{\text {DDlow }}$, respectively, and the measurements show


Fig. 4-12: Oscillation frequency vs. temperature for the ring oscillators of Fig. 4-7.
that a settling time of about 1.5 ms is required. After such settling time required to power up the transponder, the storage capacitance is charged and the supply voltage remains almost constant during reception of the ASK modulated signal.

The measured temperature coefficient of the output voltage is $-2 \mathrm{mV} /{ }^{\circ} \mathrm{C}$ and is constant from 0 to $80^{\circ} \mathrm{C}$. The oscillation frequency of the two circuits of Fig. 4-7a and Fig. 4-7b is plotted as a function of temperature in Fig. 8. For the circuit of Fig. 4-7a the oscillation frequency has a large relative variation with temperature of about $4 \% /{ }^{\circ} \mathrm{C}$ and the power consumption has the same relative variation of about $4 \% /{ }^{\circ} \mathrm{C}$, in agreement with (4-18). In the case of Fig. 4-7b, the oscillation frequency has a very small dependence on the temperature of only $-0.08 \% /{ }^{\circ} \mathrm{C}$ and the power consumption has a relative variation of about $-0.74 \% /{ }^{\circ} \mathrm{C}$, in agreement with (4-19). From (4-20) we can derive that the relative variation of the operating range of the tag-reader system would be limited to $-2 \% /{ }^{\circ} \mathrm{C}$ and $0.37 \% /{ }^{\circ} \mathrm{C}$ for the circuits of Fig. 4-7a and Fig. 4-7b, respectively. As can be seen, the use of our noncompensated voltage regulator would ensure an overall temperature compensated performance of the subthreshold circuit.


Fig. 4-13: Block diagram of the proposed voltage regulator.

### 4.5. Constant-delay voltage regulator

A block diagram of the proposed voltage regulator is shown in Fig. 4-13. The core of the proposed voltage regulator is the current reference circuit that generates a current $I_{R E F}$ proportional to the generated supply voltage $V_{D D}$ and to the reference voltage $V_{0}$. Such a current is mirrored into a current to voltage converter and converted into a voltage, which is applied to the input of an amplifier that provides the supply voltage $V_{D D}$ so that the current $I_{0}$ in the PMOS, which has a gate-source voltage equal to $V_{D D}$, is equal to $I_{R E F}$. Then a series voltage regulator provides a regulated voltage $V_{R E G}$ equal to $V_{D D}$ used to supply the digital subthreshold circuit. If a CMOS inverter is connected to the output, when the PMOS conducts, its gatesource voltage is equal to $V_{R E G}$ and then its load capacitance will be charged with the current $I_{\text {REF }}$; its propagation delay is thus given by (4-2) with $I_{D}=I_{\text {REF }}$.

### 4.5.1. Current Reference Circuit

The current reference circuit is shown in Fig. 4-14. The two transistors $M_{1}$ and $M_{2}$ work in the triode region and their gate voltages are provided by the voltage doubler shown in Fig. 4-13. The two operational amplifiers $\mathrm{A}_{1}$ and $\mathrm{A}_{2}$ impose the drain-source voltages of the two transistors $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ equal to each other and to the reference voltage $\mathrm{V}_{0}$. The reference voltage $\mathrm{V}_{0}$ is generated by a voltage reference circuit that provides a reference voltage $V_{\text {REF }}$ and a voltage divider that provides the desired voltage $V_{0}=V_{\text {REF }} / \gamma$, where $\gamma$ is the partition coefficient. In order to ensure the operation of $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ in the triode region, such voltage has


Fig. 4-14: Current reference circuit.
been set to about 80 mV . Since the generated supply voltage $\mathrm{V}_{\mathrm{DD}}$ is few hundreds of mV , in order to ensure the operation of $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ in the strong inversion regime, the voltage $\mathrm{V}_{\mathrm{DD}}$ has been doubled before being applied to the gate of the two transistors. The I-V characteristic of a MOS in the triode region can be well approximated by,

$$
\begin{equation*}
I_{D}=k\left[\left(V_{G S}-V_{t h}\right) V_{D S}-V_{D S}^{2} / 2\right] \tag{4-21}
\end{equation*}
$$

By applying (4-21) to the transistor $\mathrm{M}_{1}$, which has a drain current equal to $I_{0}+I_{R E F}$, and to the transistor $\mathrm{M}_{2}$, which has a drain current equal to $I_{0}$, one can obtain,

$$
\begin{gather*}
I_{0}+I_{R E F}=k\left[\left(2 V_{D D}-V_{t h}\right) V_{D S 1}-V_{D S 1}^{2} / 2\right],  \tag{4-22}\\
I_{0}=k\left[\left(2 \alpha V_{D D}-V_{t h}\right) V_{D S 2}-V_{D S 2}^{2} / 2\right] . \tag{4-23}
\end{gather*}
$$

By assuming that $V_{D S 1}=V_{D S 2}$ and by subtracting (4-22) and (4-23), we can derive the expression of the current $I_{R E F}$, given by (4-24).


Fig. 4-15: Voltage reference generator.

$$
I_{R E F}=2 k(1-\alpha) V_{D D} V_{R E F} / \gamma .
$$

### 4.5.2. Voltage Reference Generator

The voltage reference generator is shown in Fig. 4-15. The voltage reference generator consists of a circuit that generates a bias current $I_{0}$ almost independent of the supply voltage $V_{D D} ; I_{0}$ is then injected into an active load $\left(\mathrm{M}_{7}-\mathrm{M}_{10}\right)$ to generate the reference voltage. In the current generator of Fig. 4-15, $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ operate in the subthreshold region, while $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$ are in the saturation region [10]. It can be shown that the expression for $I_{0}$ is [10]

$$
\begin{equation*}
I_{0}=\frac{m^{2} V_{T}^{2} k_{4}}{2}\left(\frac{N}{N-1}\right)^{2} \ln ^{2}\left(\frac{W_{2} / L_{2}}{W_{1} / L_{1}}\right)=\frac{\mu_{n} C_{o x} V_{T}^{2}}{2} h, \tag{4-25}
\end{equation*}
$$

where $N=\sqrt{k_{3} / k_{4}}$. The active load used to generate the reference voltage consists of two NMOS transistors, $\mathrm{M}_{7}$ and $\mathrm{M}_{8}$, which are $3.3 \mathrm{~V}-\mathrm{NMOS}$ with a threshold voltage $V_{\text {thL }}=0.45 \mathrm{~V}$ biased by $I_{0}$, and a voltage divider formed by $\mathrm{M}_{9}$ and $\mathrm{M}_{10}$, which are 5 V -NMOS with a threshold voltage $V_{\text {th } H}=0.75 \mathrm{~V}$. All transistors in the active load operate in the saturation region. The output voltage $V_{\text {REF }}$ is

$$
\begin{equation*}
V_{R E F}=(1-\beta) V_{t h H}+\beta V_{t h L}+\left[(1+\beta) \sqrt{\frac{h}{W_{7} / L_{7}}}-\sqrt{\frac{h}{W_{8} / L_{8}}}\right] V_{T}, \tag{4-26}
\end{equation*}
$$

where, $\beta=\sqrt{k_{10} / k_{9}}$. The output reference voltage so generated, has a value of about 1.3 V , is applied to a voltage divider that provides the voltage $\mathrm{V}_{0}$ required by the current reference circuit.

### 4.5.3. Sensitivity to Temperature Variations

By substituting (4-24) in (4-2), the propagation delay becomes,

$$
\begin{equation*}
t_{p}=\frac{C}{4 k(1-\alpha) V_{R E F} / \gamma} . \tag{4-27}
\end{equation*}
$$

Since $\alpha$ and $\gamma$ are partition coefficients, they can be considered, in a first approximation, independent of temperature and process variations (mismatch negligible). As a consequence, in order to have a zero temperature coefficient, the temperature coefficient of $k V_{\text {REF }}$ must be set to zero. This means that,

$$
\frac{\partial\left(k V_{0}\right)}{\partial T}=\frac{\mu_{T}}{T} V_{R E F}+\frac{\partial V_{R E F}}{\partial T}=0
$$

where we have used the BSIM3 model [9] for the temperature dependence of the mobility and $\mu_{T}$ is the mobility temperature exponent. From (4-28) we can derive the temperature coefficient of the reference voltage so that the temperature coefficient of the propagation delay is zero and then the performance of the digital subthreshold circuit is almost insensitive to temperature variations.

We can now proceed to calculate the temperature coefficient of the reference voltage $V_{\text {REF }}$. In a first approximation we can assume that the threshold voltage of a MOS transistor has a negative linear dependence on the temperature with a coefficient $K_{t H}$ for 5 V -NMOS transistors and $K_{t L}$ for 3.3 V -NMOS transistors. By differentiating the expression of $V_{R E F}$, given in (4-26), the temperature coefficient of the reference voltage is,

$$
\frac{\partial V_{R E F}}{\partial T}=-(1-\beta) K_{t H}-\beta K_{t L}+\frac{V_{R E F}-(1-\beta) V_{t h H}-\beta V_{t h L}}{T}
$$

By equating the temperature coefficient given by (4-29) to $-\mu_{T} V_{R E F} / T$, in order to satisfy (4-28), one can obtain,

$$
\beta=\frac{K_{t H} T-\left(1+\mu_{T}\right) V_{R E F}+V_{t h H}}{\left(K_{t H}-K_{t L}\right) T+V_{t h H}-V_{t h L}} .
$$

With such choice of $\beta$ the temperature coefficient of the voltage reference is positive but the overall temperature coefficient of the propagation delay is compensated leading to small sensitivity to temperature variations. From (4-30), it is clear that, in order to achieve the correct temperature compensation, we need to use in the active load transistors with different threshold voltages and different temperature coefficients.

### 4.5.4. Sensitivity to Process Variations

Subthreshold digital circuits are strongly affected by process variations, especially by threshold voltage variations. Because of the exponential dependence of the drain current on the threshold voltage shift, in the case of a constant supply voltage $V_{D D}$, a large sensitivity is achieved. From (4-1) for $V_{D S} \gg V_{T}$ and (4-2), we can derive the propagation delay sensitivity to process variations, as shown below,

$$
\frac{\partial t_{p}}{t_{p}}=-\frac{\partial I_{D}}{I_{D}}=-\frac{\partial \mu}{\mu}-\frac{\partial W}{W}+\frac{\partial L}{L}+\frac{1}{m V_{T}} \frac{\partial V_{t h}}{V_{t h}} .
$$

The process variations of $W$ and $L$ are negligible if a large value of $W$ and $L$ is used. Because of the presence of the coefficient $m V_{T}$, the dispersion of the threshold voltage due to process variations, which is about $20 \%$, is multiplied for a factor larger than 30 leading to huge sensitivity of subthreshold circuits to process variations.

In the case of the proposed voltage regulator, by assuming the process variations on $W$ and $L$ negligible, from (4-27) the propagation delay sensitivity to process variations is given by,

$$
\begin{equation*}
\frac{\partial t_{p}}{t_{p}}=\frac{\partial \mu}{\mu}+\frac{\partial V_{R E F}}{V_{R E F}} . \tag{4-32}
\end{equation*}
$$

From (4-26) we can derive the sensitivity of the reference voltage $V_{R E F}$ to process variations, as shown below,

$$
\begin{equation*}
\frac{\partial V_{R E F}}{V_{R E F}}=(1-\beta) \frac{V_{t h H}}{V_{R E F}} \frac{\partial V_{t h H}}{V_{t h H}}+\beta \frac{V_{t h L}}{V_{R E F}} \frac{\partial V_{t h L}}{V_{t h L}} . \tag{4-33}
\end{equation*}
$$

By assuming that the process variations on $V_{t h H}$ and $V_{t h L}$ are uncorrelated and that the relative variations of $V_{t h H}$ and $V_{t h L}$ are equal $\left(\partial V_{t h H} / V_{t h H} \cong \partial V_{t h L} / V_{t h L}=\partial V_{t h} / V_{t h}\right)$, the relative standard deviation of the reference voltage can be written as,

$$
\frac{\sigma_{V R E F}}{V_{R E F}}=\frac{\sigma_{V t h}}{V_{t h}} \sqrt{(1-\beta)^{2}\left(\frac{V_{t h H}}{V_{R E F}}\right)^{2}+\beta^{2}\left(\frac{V_{t h L}}{V_{R E F}}\right)^{2}} \cong 1.3 \frac{\sigma_{V t h}}{V_{t h}} .
$$

From (4-32), by assuming that the process variations on $V_{\text {REF }}$ and $\mu$ are uncorrelated, the standard deviation of the propagation delay is given by,

$$
\frac{\sigma_{t p}}{t_{p}}=\sqrt{\frac{\sigma_{\mu}^{2}}{\mu^{2}}+\frac{\sigma_{V R E F}^{2}}{V_{R E F}^{2}}} .
$$

In our process, (4-35)would be equal to 0.083 and then $3 \sigma / t_{p}$ would be $21.8 \%$, which is reasonable compared to a value larger than $300 \%$ when the regulator is not used.

### 4.5.5. Experimental Results

The proposed voltage regulator was implemented in AMS $0.35 \mu \mathrm{~m}$ CMOS process. The die photograph is shown in Fig. 4-16. In order to emphasize the advantages given by the proposed voltage regulator, at first, a CMOS inverter has been tested when supplied by the proposed voltage regulator and when supplied by a constant supply voltage and in both cases the rise time has been measured as a function of temperature. The supply voltage of the CMOS inverter in the case of a constant supply voltage has been chosen so that, at room temperature, the rise time of the inverter is equal in the two cases. Fig. 4-17 shows the behavior of the inverter in the two cases as a function of temperature. When the proposed regulator is used, the temperature coefficient of the rise time is $114 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$; in the case of a constant supply voltage the temperature coefficient is $5.5 \% /{ }^{\circ} \mathrm{C}$. The use of the proposed voltage regulator has reduced the sensitivity of the performance of a CMOS inverter to the temperature by 500 times. In order to test the sensitivity to process variations, the rise time of the inverter in the two cases has been tested for 20 different samples


Fig. 4-16: Die Photograph (core).


Fig. 4-17: Rise time vs. Temperature for a) an inverter supplied by the proposed voltage regulator $V_{D D}=2 \mathrm{~V}, \mathrm{~b}$ ) an inverter supplied by a constant supply voltage.


Fig. 4-18: Rise time in 20 samples for a) an inverter supplied by the proposed voltage regulator $\mathrm{V}_{\mathrm{DD}}=2 \mathrm{~V}, \mathrm{~b}$ ) an inverter supplied by a constant supply voltage.
from the same batch and different wafers and the results are shown in Fig. 4-18. Experimental results show that $3 \sigma_{t p} / t_{p}$ is $14.81 \%$ in the case of the inverter supplied by the proposed voltage regulator (Monte Carlo simulations account also for inter-batch variations and predict a larger $3 \sigma_{t p} / t_{p}$ of $21.8 \%$ ) and $368.5 \%$ in the case of a constant supply voltage.
The same measurements have been repeated with a ring oscillator and the oscillation frequency has been measured in both cases. Fig. 4-19 shows the behavior of the ring oscillator in the two cases as a function of temperature. In the case that the proposed regulator is used, the temperature coefficient of the rise time is $185 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$; in the case of a constant supply voltage the temperature coefficient is $20.5 \% /{ }^{\circ} \mathrm{C}$, which is more than 1000 times larger.

The ring oscillator in the two cases has been tested in 20 different samples and the results are shown in Fig. 4-19. Experimental results show that $3 \sigma_{t p} / t_{p}$ is $15.6 \%$ in the case of the ring oscillator supplied by the proposed voltage regulator (Monte Carlo simulations predict $3 \sigma_{\text {fosc }} / f_{\text {osc }}=21.6 \%$ ) and $406 \%$ in the case of a constant supply voltage.

The supply voltage range ensuring the correct operation of the proposed voltage regulator is from 1.8 V to 4.3 V . The minimum supply current measured at the


Fig. 4-19: Frequency vs. Temperature for a ring oscillator a) supplied by the proposed voltage regulator with $\left.V_{D D}=2 \mathrm{~V}, \mathrm{~b}\right)$ supplied by a constant supply voltage.


Fig. 4-20: Oscillation frequency in 20 samples for a) a ring oscillator supplied by the proposed voltage regulator $\mathrm{V}_{\mathrm{DD}}=2 \mathrm{~V}$, b) an inverter supplied by a constant supply voltage.
minimum supply voltage and at $0^{\circ} \mathrm{C} 2.9 \mu \mathrm{~A}$ and the maximum supply current measured at the maximum supply voltage and at $80^{\circ} \mathrm{C}$ is $3.4 \mu \mathrm{~A}$.
In the case of the CMOS inverter supplied by the proposed voltage regulator, when the supply voltage varies from its minimum to its maximum value, the rise time varies of $2.87 \%$ leading to a line sensitivity of $1.2 \% / \mathrm{V}$; the CMOS inverter supplied by a constant voltage has a line sensitivity of 130 times/V. In the case of the ring oscillator, instead, when the supply voltage varies from its minimum to its maximum value, the rise time varies of $1.9 \%$ leading to a line sensitivity of $0.77 \% / \mathrm{V}$; the ring oscillator supplied by a constant voltage has a line sensitivity of 5000 times $/ \mathrm{V}$. As a consequence, the proposed voltage regulator is almost insensitive to supply voltage variations ensuring good performance over the entire supply voltage range.

The area occupation on the chip is $0.44 \mathrm{~mm}^{2}$.

### 4.6. Conclusion

A voltage regulator for digital subthreshold circuits implemented in a standard CMOS process has been presented. The proposed regulator has been tested when providing power supply to a CMOS inverter and a ring oscillator operating in subthreshold regime. Experimental results show that the huge sensitivity of subthreshold circuits to temperature and process variations is drastically suppressed when the proposed voltage regulator is used and is reduced to the typical sensitivity of temperature compensated circuits operating in strong inversion. Moreover, the proposed regulator also fits the strict power requirements of subthreshold circuits drawing only few $\mu \mathrm{A}$ from the supply voltage.

### 4.7. References

[1] K. Finkenzeller, RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, $2^{\text {nd }}$ ed, Wiley and Sons, 1999, pp. 117-126, pp. 143-148 and pp. 183-186.
[2] T. Basten, L. Benini, A. Chandrakasan, M. Lindwer, J. Liu, R. Min, F. Zhao, "Scaling into Ambient Intelligence," Proc. DATE 2003, pages 76-81, March 2003.
[3] CEPT REC 70-03 Annex 1, ETSI EN 330 220-1.
[4] G. De Vita, G. Iannaccone, "Design Criteria for the RF section of UHF and Microwave Passive RFID Transponders," IEEE Transactions on Microwave Theory and Techniques, Vol. 53, No 9, pp 2978-2990, September 2005.
[5] V. Balan, "A Low-Voltage Regulator Circuit with Self-Bias to Improve Accuracy," IEEE Journal of Solid State Circuits, 38(2): 365-368, February 2003.
[6] G. A. Rincon-Mora and P. E. Allen, "A Low-Voltage, Low-Quiescent Current, Low Drop-out Regulators," IEEE Journal of Solid State Circuits, 33(1): 36-44, January 1998.
[7] ISO/IEC 18000-6, "Information technology - Radio frequency identification for item management - Part 6: Parameters for air interface communications at 860 MHz to 960 MHz ", 2004.
[8] K. N. Leung and P. K. T. Mok, "A CMOS Voltage Reference Based on Weighted $\Delta \mathrm{V}_{\mathrm{GS}}$ for CMOS Low-Dropout Linear Regulators," IEEE Journal of Solid State Circuits, 38: 146-150, January 2003.
[9] BSIM3v3 User's Manual. Available at: http://www-device.eecs.berkeley.edu/~bsim3/
[10] G. De Vita, G. Iannaccone, "A $300 \mathrm{nW}, 12 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ Voltage Reference in a Digital $0.35 \mu \mathrm{~m}$ CMOS Process," Proc. of VLSI Symposium, Honolulu, USA, June 2006.

## 5. NON-VOLATILE MEMORY

### 5.1. Introduction

An extreme attention to low power issues is key to the widespread adoption of passive UHF and Microwave RFID (Radio Frequency Identification) transponders in the field of identification systems and to their possible migration to ambient intelligence scenarios [1]. Indeed the practical use of passive transponders is determined by their operating range in realistic harsh conditions, which is a fraction of their operating range in free space, which in turn is determined by the minimum available power at the antenna required for transponder operation. Achieving a large operating range is particularly challenging in the European Union, where strict regulations limit the power transmitted by the reader to 500 mW EIRP [2], one eighth of the corresponding US limitation. In such conditions, in order to achieve an operating range of a few meters [3], the transponder must operate correctly with only few $\mu \mathrm{W}$ of regulated DC power supply.
A second consideration regards some recently presented scenarios of Ambient Intelligence [1], which often envisage a network of low-cost, low-data-rate transponders or transceivers. Present solutions are not typically low-cost (for example are based on Bluetooth transceivers), or are completely abstracted from the hardware implementation, and therefore have not addressed real cost issues. We believe that both for Ambient Intelligence and for the widespread use of RFID systems, the "hardware" is both the enabler and the limit, and that a much sounder approach would require to start from the transponder hardware design, focusing on the "low power" and "low cost" issues. In our view, RFID systems should be viewed as prototypes of Ambient Intelligence systems.
We have recently designed the RF section of an ultra-low-power passive transponder, with three main constraints: $i$ ) power consumption of both the analog and digital sections of the tag in the sub- $\mu \mathrm{W}$ regime, $i$ ) power efficiency of the RF/DC converter larger than $20 \%$ and iii) optimized modulation depth of the backscattered radiation in order to maximize the operating range [4]. In such a way, for a total DC power consumption of $1 \mu \mathrm{~W}$ and considering European regulations (EIRP $=500 \mathrm{~mW}$ ), an operating range of 3.4 m in the $2.45-\mathrm{GHz}$ ISM band and of 9.5 m in the UHF ISM band can be achieved [3], [4]. To match the constraint on power consumption, the digital section of the tag is implemented in subthreshold logic, with a standard $0.35 \mu \mathrm{~m}$ CMOS process, a supply voltage of 0.6 V , and a clock frequency smaller than 1 MHz .


Fig. 5-1: Tag Architecture.
In this paper, we focus on the possibility to introduce a 128 bit non-volatile embedded memory on the tag IC, implemented with a standard, i.e., single-poly, CMOS process (AMS $0.35 \mu \mathrm{~m}$ CMOS). The issue is particularly challenging again - for power constraints. The complete design of the flash memory, including the cell structure, the memory architecture and the memory circuits, has been performed with extreme attention to low power issues, ensuring a total power consumption of the tag below $1 \mu \mathrm{~W}$, during a read operation, and an acceptable degradation of the operating range during a write operation. To meet both requirements we use a read voltage equal to the supply voltage of the digital section ( $V_{D D}=0.6 \mathrm{~V}$ ) and we use Fowler-Nordheim program and erase.
The architecture of a passive RFID transponder with an on-board flash memory is shown in Fig. 5-1. A single stage voltage multiplier converts the input alternating voltage in a DC voltage required to supply the transponder. As can be seen, there are two series voltage regulators: the one at the bottom operates correctly within the read operating range providing a regulated voltage $V_{D D}=0.6 \mathrm{~V}$, used as a power supply for the finite state machine and - during a read operation - for the memory. The series regulator at the top operates correctly only when the tag is within the write operating range, and provides a voltage of 3 V that are further increased by the charge pump up to $V_{H I G H}=9 \mathrm{~V}$ and used for program/erase operations. The implementation of the RF front-end and of the voltage regulators is described elsewhere [4].

### 5.2. Memory Costraints

The most critical aspect of EEPROMs is represented by the power required to charge and discharge the memory cell floating gates, that is typically well above $1 \mu \mathrm{~W}$, and therefore may significantly reduce the operating range. The use of Fowler-Nordheim tunneling implies both low power and large write and erase times, in the order of a few tens of ms [5]; in the case of RFID systems the latter is not a severe limitation since data rates are typically rather low (10-40 kbps) [6] and the arguments of a write/erase operation are first stored on local registers and then -off-line - are transferred onto the non volatile memory, leaving the reader free to address other transponders. The time required to write the flash memory sets a limit to the minimum time during which the tag must stay within the write operating range of the reader, or, in other words, to the maximum relative reader-tag speed.
Another critical constraint is posed by the low supply voltage during a read operation. All EEPROM circuits must operate correctly with $V_{D D}$, and therefore must use low-voltage transistors, which require protection circuitry to prevent breakdown when a high voltage is applied in order to perform a write/erase operation. On the other hand, we can relax the constraints on memory access time, since passive RFID systems work with very low data rates, and on the chip area, since the memory has a very limited number of bits. This last consideration lets us pay the price in terms of area occupancy required to implement the memory in a standard CMOS process.

### 5.3. Memory Cell

A nonvolatile memory can be implemented in a standard CMOS process by realizing a floating node which preserves the stored charge [5], [6]. In its simplest form, an EEPROM is a MOS transistor having a floating gate, as shown in Fig. 5-2. The capacitive divider consisting of the capacitance $C$, associated to the PMOS transistor, and of the gate capacitance $C_{g}$ of the $\mathrm{M}_{\text {cell }}$ transistor, transfers a fraction of the voltage applied to the control terminal to the floating gate $F G$. To minimize power consumption, we use Fowler-Nordheim tunneling through the gate oxide of the NMOS transistor to charge and discharge the floating node. Such effect is obtained by applying relatively high field ( $\geq 10 \mathrm{MV} / \mathrm{cm}$ ) to the gate oxide, obtaining a tunneling current provided by the following approximate expression:

$$
\begin{equation*}
I \approx A(W L) E^{2} \exp (-B / E), \tag{5-1}
\end{equation*}
$$



Fig. 5-2: Memory Cell.
where $E$ is the electric field across the oxide, $W L$ is the gate area, $A=0.37 \mu \mathrm{~A} / \mathrm{V}^{2}$ and $B=223 \mathrm{MV} / \mathrm{cm}$. Values of $A$ and $B$ have been obtained by fitting 5-1) to the experimental I-V characteristic of a MOS capacitor with oxide thickness of 7.7 nm , as in the case of our CMOS technology.
In order to perform a write operation, a high voltage is applied to the floating node through the capacitive divider, the bit line is grounded, and source is floating. In such a way a high voltage drops between the gate and the channel, giving rise to a tunneling current that charges the floating gate. The floating source follows the drain voltage and ensures that no current flows from drain to source even if the bit line is not exactly at ground.
In order to perform an erase operation the high voltage is applied to the drain of the NMOS transistor, and the control terminal and the source are grounded. The high voltage drops across the drain and gate terminals and a tunneling current discharges the floating node [5]. No current flows from drain to source since the channel is not formed. The charge in the $F G$ results in a threshold voltage shift, as seen from the control terminal, which can be detected as a logic state by a sense amplifier. In our IC technology the maximum write/erase voltage is limited by the drain-source breakdown and has to be kept below 9 V [7]. The gate area of the NMOS transistor and the capacitive ratio $N \equiv C / C_{g}$ must be dimensioned in order to maximize the shift of the threshold voltage while keeping the power required during write/erase as small as possible, for a given high voltage. In Fig. 5-3a and Fig. 5-3b we show the results of numerical simulations of the time-dependent program operation where the tunnel current at a given time, given by (5-1), is a function of the potential of


Fig. 5-3: Results of the simulation of a program operation for an applied voltage on the control terminal of 9 V , a write time of 50 ms , and two values of the gate area ( 3.5 and $35 \mu \mathrm{~m}^{2}$ ). a) Threshold voltage shift as seen from the control terminal; b) Maximum power required for the operation.
the floating gate, that in turn depends on the charge accumulated in the floating gate up to that time. It is clear that the larger $N$, the larger is the voltage transferred to the floating gate and then the tunneling current and the charge variation at the floating node. But with increasing $N$ the PMOS capacitance increases too and then the variation of the threshold voltage seen from the control terminal tends to saturate for large $N$, as can be seen in Fig. 5-3a for a voltage $V_{H I G H}=9 \mathrm{~V}$ applied to the control terminal and a write time of 50 ms . The power required, instead, increases because of the increase of the tunneling current, as shown in Fig. 5-3b. We can see that for $N>20$ the shift of the threshold voltage seen from the control terminal increases very slowly while the power required increases quickly, and therefore we choose $N=20$. Since for a given $N$ the threshold voltage shift is independent of the gate area (as shown in Fig. 5-3a) while the power is strongly dependent on it (as shown in Fig. 5-3b), it is advisable to dimension the gate area of the NMOS transistor as small as possible, given the requirements of the sense
amplifier, and then to dimension the PMOS transistor accordingly.

### 5.4. Memory Architecture

We can see from Fig. 5-3a that with a gate area of $3.5 \mu \mathrm{~m}^{2}$ and a capacitance ratio $N=20,50 \mathrm{~ms}$ are sufficient to reach a threshold voltage of 2.6 V , which is sufficient for our needs. Furthermore, the maximum power required during the program/erase operation, corresponding to the initial instants, when the tunneling current is maximum, is about 30 nW . In order to write 128 bits in the memory two steps are required: an erase operation, to reset the bits that must be ' 0 ', and a write operation to set the bits that must be ' 1 '. Since the program and erase operations have the same duration [7], it is clear that updating simultaneously all the 128 cells implies a minimum time of 100 ms , that would correspond to the maximum power consumption, and therefore to a strong reduction of the operating range.
On the other hand, writing the 128 bits one by one would require only 60 nW , but a total writing time in excess of 10 s . A reasonable tradeoff is to write words of 16 bits simultaneously, requiring a total of 8 steps. In such a way the total write time is 0.8 s and the required power for program/erase is $0.96 \mu \mathrm{~W}$. In addition, we have to consider the power dissipated in the memory circuits during the program/erase operation, which leads to a total power consumption of few $\mu \mathrm{W}$ that allows us to achieve an operating range of several meters also during a write operation [3]. Since the memory has only 128 cells, we can use a single word line and a 3-bit decoder, resulting in a very simple circuit, while the total area occupation would still be acceptable.

### 5.5. Memory Circuits

### 5.5.1. Sense Amplifier

As we have already mentioned, the very low supply voltage of 0.6 V , requires us to implement the sense amplifier with low-voltage transistors and additional protection circuitry. The low $V_{D D}$ leads us to prefer a current-mode sense amplifier, as shown in Fig. 5-4 [8], that also lets us to avoid the problem of the control of the bit line voltage biasing, which is not suitable for very low power supply. All transistors in Fig. 5-4 are low-voltage transistors except for $M_{n 1}$ and $M_{n 2}$ that are high-voltage transistors: their use will be clarified later.
During a read operation, the word line goes to the read voltage ( 0.6 V ), the signal WE goes high, grounding the source of $\mathrm{M}_{\text {cell }}$, the signal Y goes high so that the cell is selected. Furthermore, during a read operation the output buffer is enabled while the input buffer is disabled and then the bit line is floating. The voltage on the bit line is set as small as possible, so that the high voltage transistor $\mathrm{M}_{\mathrm{n} 1}$ can have a


Fig. 5-4: Current mode Sense Amplifier.
gate-source voltage close to the read voltage and therefore correctly operates as a switch with an acceptable drain-to-source voltage (about 100 mV ). At the same time, the bit line voltage has to be large enough so that the $\mathrm{M}_{\text {cell }}$ transistor can provide a current sufficient for the correct operation of the sense amplifier. The current $I_{\text {cell }}$ is injected into the current mirror and compared with the current $I_{\text {ref }}$, generated by biasing the reference cell with an appropriate voltage $\alpha V_{D D}(0<\alpha<1)$. As a consequence, if the selected cell is erased the current comparator compares $I_{r e f}$ with a larger current $I_{\text {cell }}$ so that OUT goes to ' 1 '; if the selected cell is written the current $I_{\text {cell }}$ is much smaller than $I_{\text {ref }}$ and then the OUT goes to ' 0 '. The current $I_{\text {ref }}$ is set to only 6 nA leading to a static power during the read operation of only 0.36 nW per cell. If the cell is not selected, the pass transistor $\mathrm{M}_{\mathrm{n} 2}$ is disabled and then the power consumption is zero.
During an erase operation the word line is low, the signal WE is high, grounding the source of $\mathrm{M}_{\text {cell, }}, \mathrm{Y}$ is low, disabling the sense amplifier, the output buffer is disabled so that the output is floating, and the input buffer is enabled driving the bit line. If the cell must be erased, the input buffer applies a voltage $V_{\text {HIGH }}$ to the bit line, which largely drops between drain and gate of $\mathrm{M}_{\text {cell }}$ in order to erase the cell. Since the pass transistor $\mathrm{M}_{\mathrm{n} 1}$ is in the cut off region, almost all the voltage applied to the bit line drops between source and drain of $\mathrm{M}_{\mathrm{n} 1}$ protecting the current mirror from the high voltage. If the voltage applied to the bit line is zero the cell is not erased.
During a write operation the word line is at $V_{H I G H}, \mathrm{Y}$ is low and then the sense amplifier is disabled, the output buffer is disabled so that the output is floating, the


Fig. 5-5: Left: word line driver; Right: high voltage tri-state buffer.
input buffer is enabled driving the bit line, and the signal WE goes low so that the source of $\mathrm{M}_{\text {cell }}$ is floating. If the cell must not be written, the input buffer applies a voltage $V_{\text {HIGH }}$ to the bit line, so that the drain of $\mathrm{M}_{\text {cell }}$ is at $V_{H I G H}$, the source goes to the bit line voltage, since it is floating, and then the gate-source and gate-drain voltages of $\mathrm{M}_{\text {cell }}$ are low, preventing the writing of the cell. If the cell must be written, the input buffer applies a zero voltage to the bit line, so that the drain voltage of $\mathrm{M}_{\text {cell }}$ is grounded, the source goes to zero too, and then the gate-source and gate-drain voltages of $\mathrm{M}_{\text {cell }}$ are at the high voltage allowing the writing of the cell. It is important to note that the pass transistor $\mathrm{M}_{\mathrm{n} 1}$ has to be a high-voltage transistor in order to prevent the high voltage applied to the bit line to be transferred to the current mirror causing its breakdown.
Also transistor $\mathrm{M}_{\mathrm{n} 2}$ has to be a high-voltage transistor because its drain-source voltage can be at $V_{H I G H}$, as shown earlier. By using only two high-voltage transistors the sense amplifiers can perform the read operation with a very low power consumption and at the same time can work correctly during an erase or a write operation without causing the breakdown of the low-voltage transistors.

### 5.5.2. Word Line Driver

The word line driver has to provide a signal from zero to $V_{D D}$, during the read operation, and a signal that varies from 0 to $V_{H I G H}$ during a program/erase operation. The word line driver is shown on the left side of Fig. 5-5. The buffer in the upper branch is a tri-state buffer (shown in detail on the right side of Fig. 5-5 supplied by $V_{\text {HIGH }}$ and therefore implemented with high-voltage transistors; the
buffer in the lower branch is a tri-state buffer supplied by $V_{D D}$ and implemented with low-voltage transistors. The pass gate in the lower branch is supplied by $V_{H I G H}$ and therefore implemented with high-voltage transistors. During a read operation the high-voltage buffer is disabled and then its output is floating, while the lowvoltage buffer and the pass-gate are enabled transferring the input signal to the output. During a program/erase operation the low-voltage buffer and the pass gate are disabled, while the high voltage buffer is enabled setting the output voltage to $V_{H I G H}$ or zero, if the input voltage is $V_{D D}$ or zero, respectively. In such case the output high voltage drops on the pass gate, which is implemented with high-voltage transistors, ensuring the safe operation of the low-voltage buffer.
While the low-voltage buffer is a standard tri-state buffer supplied with the read voltage ( 0.6 V ), the high-voltage buffer has to be modified so that, during the read operation, when the tag is out of the write operating range and then $V_{H I G H}$ is close to zero, it ensures a floating output. Such behavior is obtained with the circuit shown in Fig. 5-5, on the right: it is a standard tri-state buffer implemented using highvoltage transistors, with the addition of the transistor $\mathrm{M}_{\mathrm{p} 1}$ and the $V_{C C}$ generator, that ensure that the output is floating during a read operation also in absence of high voltage (so for $V_{H I G H}=0$ ). $V_{\text {MEDIUM }}$ is generated by a voltage divider and is $\mathrm{V}_{\mathrm{HIGH}} / 2$. During an erase or program operation we have $V_{C C}=V_{\text {HIGH }}$ while during the read operation we have $V_{C C}=V_{\text {MEDIUM }}$.
The operating principle of the output buffer is the following: during the read operation, if $V_{\text {HIGH }}$ is close to zero, although the enable E is low, the pull up network of the NAND can not conduct and then its output voltage should also be zero. The presence of $\mathrm{M}_{\mathrm{p} 1}$ prevents $\mathrm{M}_{\mathrm{p} 2}$ to go into weak conduction; indeed, if the low voltage buffer imposes a high level $(0.6 \mathrm{~V})$, the voltage on the drain of $\mathrm{M}_{\mathrm{p} 2}$ is about 0.3 V , keeping $\mathrm{M}_{\mathrm{p} 2}$ in cut off. For high regulated values of $V_{H I G H}$ the pull up of the NAND goes in conduction and both $\mathrm{M}_{\mathrm{p} 1}$ and $\mathrm{M}_{\mathrm{p} 2}$ are in the cut off region. During a program/erase operation $V_{C C}$ is equal to the high voltage $V_{H I G H}$ and the enable E goes high. Transistor $\mathrm{M}_{\mathrm{p} 1}$ has a source-gate voltage $\mathrm{V}_{\mathrm{HIGH}} / 2$ and is therefore in the ON state, so that the input buffer can impose the voltage on the bit line according to the input IN . The circuit generating $V_{C C}$ is shown in Fig. 5-6 (left): it allows us to switch the output voltage between $V_{\text {MEDIUM }}$ and $V_{\text {HIGH }}$, according to a control signal RE, and must correctly also for very low $V_{\text {HIGH }}$. Since all transistors $\mathrm{M}_{1}, \mathrm{M}_{2}, \mathrm{M}_{3}$ and $\mathrm{M}_{4}$ are high-voltage pass transistors, they need to be driven by sufficiently high voltages, i.e., $V_{\text {MEDIUM }}$ for $\mathrm{M}_{1}$ and $\mathrm{M}_{4}$ and $V_{H I G H}$ for $\mathrm{M}_{2}$ and $\mathrm{M}_{3}$. Such choice ensures the correct operation of $\mathrm{M}_{2}$ and $\mathrm{M}_{4}$ for both values of the output voltage, preventing the conduction of the drain-bulk junction that would cause high power consumption. The blocks indicated with LS are Level Shifters: in Fig. 5-6 (right), the LS block is


Fig. 5-6: Left: Vcc generator; Right: circuit for the Level Shifter (LS).
shown that shifts the voltage from $V_{D D}$ to $V_{\text {MEDIUM }}$. For small values of $V_{\text {MEDIUM }}$, if the signal IN is low, transistor $\mathrm{M}_{8}$ conducts grounding the output voltage; if the signal IN is high, $\mathrm{M}_{7}$ brings to ground the gate of $\mathrm{M}_{6}$ but, since the supply voltage is small, $\mathrm{M}_{6}$ can not conduct and the output voltage will be an intermediate voltage between zero and $V_{\text {MEDIUM }}$ and then it is not sufficient to let the transistors of the $V_{C C}$ generator conduct. If the supply voltage is sufficiently high to allow the conduction of $\mathrm{M}_{6}$, the output voltage goes to the supply voltage. The same operation is also valid for the level shifter from $V_{\text {MEDIUM }}$ to $V_{H I G H}$ with the difference that all transistors are high-voltage transistors, in order to prevent the breakdown. Then, referring to the circuit of Fig. 5-6 (left), for small values of the supply voltage, the gate voltages of transistors $M_{1}, M_{2}, M_{3}$ and $M_{4}$ are not sufficient for their conduction and then the output voltage $V_{C C}$ will be about $V_{H I G H} / 2$, independently of RE, since $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$ are identical, and then equal to $V_{\text {MEDIUM }}$. It is clear that the situation of small values of $V_{H I G H}$ can happen only during a read operation. During a program/erase operation the voltage $V_{H I G H}$ is at 9 V . For values of $V_{H I G H}$ sufficiently high to ensure the conduction of all transistors in the level shifter and in the $V_{C C}$ generator, if RE is high (read operation), $\mathrm{M}_{1}$ and $\mathrm{M}_{2}$ conduct and then the output voltage is equal to $V_{\text {MEDIUM }}$; if RE is low (erase/write operation), $\mathrm{M}_{3}$ and $\mathrm{M}_{4}$ conduct and then the output voltage is equal to $V_{\text {HIGH }}$. As a consequence, during a read operation the voltage $V_{C C}$ is always equal to $V_{\text {MEDIUM }}$; during a program/erase operation, instead, it is equal to $V_{\text {HIGH. }}$. This is exactly the desired behavior, in order to ensure the correct operation of the tri-state buffer of Fig. 5-6.
The input buffer of Fig. 5-4 is identical to the tri-state buffer of Fig. 5-5, on the


Fig. 5-7: Die photograph.
right. The output buffer, instead, is a standard tri-state buffer supplied by $V_{D D}$.

### 5.5.3. Input and Output buffer and $Y$ decoder

The Y decoder is a simple 3-bit decoder, implemented with a NOR array, and some logics to address the memory. During a program/erase operation all its outputs are at grounded and then all the sense amplifiers are disabled.

### 5.6. Experimental Results

Single memory cells with different $N$ and the complete 128-bit memory have been implemented in AMS $0.35 \mu \mathrm{~m}$ CMOS. The die photograph of the 128 -bit memory is shown in Fig. 5-7. Each single cell has been tested by performing a write operation with different voltages and write times and by measuring the threshold voltage shift of the memory cell transistor with the following procedure: first the drain current of a fresh cell, when the read voltage of 0.6 V is applied on the control terminal, is measured; then a write operation is performed and the voltage of the control terminal required to obtain the drain current of the fresh cell is measured. The difference between the measured voltage and the read voltage of 0.6 V is the threshold voltage shift. Fig. 5-8a-c show the threshold voltage shift of the memory cell transistor as a function of the write voltage, for memory cells with different $N$ and for a write time of $25 \mathrm{~ms}, 50 \mathrm{~ms}$ and 100 ms , respectively. As clear from Fig. 5-8a-c, the threshold voltage shift increases with increasing $N$ with but saturates for large $N$, as noticed in section III. In addition, the larger the write time, the larger is the threshold voltage shift for a given $N$ but the smaller is the variation of the threshold voltage shift for a given variation of the write time, since the tunneling current roughly decreases exponentially as a function of time. Indeed, when the write time increases from 25 ms to 50 ms the threshold voltage shift increases by about 1 V ; when, instead, the write time increases from 50 ms to 100 ms the threshold voltage shift increases by less than 200 mV .


Fig. 5-8: Threshold voltage shift vs. the write voltage for a write time of a) 25 ms , b) 50 ms, c) $100 \mathrm{~ms} . \mathrm{N}$ is the ratio of control capacitor area to the transistor area.


Fig. 5-9: Threshold voltage shift vs. number of write cycles for a) write voltage of 9 V and program time of $50 \mathrm{~ms}, \mathrm{~b}$ ) write voltage of 9 V and program time of 100 $\mathrm{ms}, \mathrm{c}$ ) write voltage of 8.5 V and program time of 50 ms .

In order to evaluate the endurance of the memory cell, a sequence of write and erase operations has been performed in each cell. After each operation, the threshold voltage has been measured. Fig. 5-9a and Fig. 5-9b show the threshold voltage shift as function of the number of write cycles, for a write voltage of 9 V and a write time of 50 ms and 100 ms , respectively. Fig. 5-9a shows that the threshold voltage shift has a large degradation after 10k cycles, where the threshold voltage shift is about 1.8 V , while it is only about 0.8 V after 100 k cycles. Fig. $5-9 \mathrm{~b}$ shows that the threshold voltage shift starts degrading rapidly after 100 cycles and becomes almost zero after 20k cycles due to significant oxide stress. As a consequence, a write time of 50 ms is preferable to achieve a better endurance.
The endurance measurements were also repeated with a write time of 50 ms and a smaller write voltage of 8.5 V . In such a case we have a smaller initial threshold voltage shift, as shown in Fig. 5-8b, but, at the same time, we expect reduced oxide degradation. The measurement results are shown in Fig. 5-9c. From Fig. 5-9c, we can note that the memory cell has a smaller degradation but, because the smaller initial threshold voltage shift, the intrinsic endurance is worse than that achieved with a write voltage of 9 V . Indeed, the threshold voltage shift is 1.4 V after 10 k cycles and decreases to 0.6 V after 100k cycles.
The 128 -bit memory has been tested to measure the read and write time and the average power consumption during a read and erase/write operation. The bias voltage $\alpha V_{D D}$ of the reference cell has been set by using the following procedure. A cell has been written by applying to the word line 9 V for 50 ms and then it has been read by applying the read voltage of 0.6 V to the word line: with a voltage $\alpha V_{D D}$ of 0.6 V the terminal OUT in Fig. 5-4 goes to ' 0 ' because the current in the memory cell is much smaller than that in the reference cell. Now, the voltage $\alpha V_{D D}$ is decreased until the terminal OUT goes to ' 1 ': this means that the compared currents become so small that the current mirrors and the current comparator do not work correctly. Such value was measured to be about 50 mV . The $\alpha V_{D D}$ has been thus set to 70 mV to achieve the largest read window ensuring at the same time some margin for reliable read operation. Because of the oxide degradation, when the threshold voltage of the erased cell increases by $600 \mathrm{mV}-70 \mathrm{mV}=530 \mathrm{mV}$ (the threshold voltage of the erased cell becomes about 1 V ), the current in the erased memory cell during a read operation becomes equal to that in the reference cell causing a failure in the read operation: such condition determines the memory fail. From Fig. 5-9a, the threshold voltage shift of the erased cell becomes 1 V after about 10 K cycles. As a consequence, we expect that the endurance of the memory is about 10 K , which is acceptable for a flash memory. Such a limit is imposed by the fact that we must use a small read voltage of 0.6 V to be compatible with the rest of the digital section and to keep the power consumption low enough.


Fig. 5-10: Endurance of the memory measured in different cells of different samples for a write voltage of 9 V and a write time of 50 ms .

In Fig. 5-10, some experimental results about the endurance of the memory, obtained from 5 different samples and from 4 cells per sample, are shown. They were obtained by performing a sequence of write/erase operations and measuring after how many cycles the read operation of an erased cell gives an OUT of ' 0 ' instead of the correct value of ' 1 '.
From Fig. 5-10, we can see that endurance is larger than 10 k cycles in all the cells measured, as expected. To further increase endurance with the present circuit, one could use a smaller write/erase voltage. As can be seen in Fig. 5-9c, with a write voltage of 8.5 V we obtain a smaller program window, but also a reduced degradation of the threshold voltage of the erased cell, which is the limiting factor for endurance in the present circuit. Further improvements could be obtained by optimizing the sense amplifier.
Fig. 5-11a shows the timing diagram of a read operation. The time required to read a sector of 16 bits has been measured to be 0.96 ms . As a consequence, the time required to read all the 128 bits ( 8 sectors) is about 8 ms . In order to measure the average power during a read operation, we have performed a sequence of read


Fig. 5-11: Timing diagram for a) read operation, b) write operation.


Fig. 5-12: Power consumption during a write operation vs. number of simultaneously written cells.
operations in a sector and we have measured the average current provided by the read voltage generator. In the worst condition, in which all the 16 cells of the
selected sector are fresh and then the maximum current is drawn from the supply voltage, the measured power consumption is 230 nW .
In Fig. 5-11b we represent a sketch of timing diagram of an erase/write operation. The 16 bits of the selected sector are at first erased and then the new data are written leading to a total required time of 100 ms . As a consequence the time required to erase/write all the 128 bits ( 8 sectors) is about 0.8 s .
We have measured the average power for erase/write operations by performing a sequence of erase/write in a sector and by measuring the current provided by the write voltage generator. Fig. 5-12 shows the average power consumption as function of the number of the cells in which a symbol ' 1 ' must be written, .As can be seen it reaches $3.8 \mu \mathrm{~W}$ if 16 ' 1 ' must be written.
The area occupation is not critical even if large transistors are used to ensure good conduction in subthreshold regime, because the memory size is very small. Anyway, the area of chip occupied by the whole memory is about $1 \mathrm{~mm}^{2}$.

### 5.7. Conclusion

An embedded flash memory for passive RFID transponder ICs implemented with a standard CMOS process has been presented. The design conditions and possible optimizations for the memory cell have been studied in detail with particular attention at minimizing the power consumption. The criteria for the choice of write voltage and write time are aimed at maximizing the endurance of the memory and are discussed in detail. Experimental results show that the reading of a 128 -bit code can be performed in about 8 ms with a power consumption of only 230 nW ; the writing of a new 128 -bit code can be performed in 0.8 s with a power consumption of only $3.8 \mu \mathrm{~W}$. Such read and write times are compatible with the very low datarate of such systems. The low power consumption allows us to achieve an operating range of several meters during both write and read operations.

### 5.8. References

[1] T. Basten, L. Benini, A. Chandrakasan, M. Lindwer, J. Liu, R. Min, F. Zhao, "Scaling into Ambient Intelligence," Proc. DATE 2003, Munich, Germany, March 2003, pp. 76-81, 2003.
[2] CEPT REC 70-03 Annex 1, ETSI EN 330 220-1.
[3] G. De Vita, G. Iannaccone, "Design Criteria for the RF section of UHF and Microwave Passive RFID Transponders," IEEE Transactions on Microwave Theory and Techniques, Vol. 53, No 9, pp 2978-2990, September 2005.
[4] G. De Vita, G. Iannaccone, "Ultra Low Power RF Section of a Passive Microwave RFID transponder in $0.35 \mu \mathrm{~m}$ BiCMOS," Proc. ISCAS 2005, Kobe, Japan, May 2005.
[5] K. Ohsaki, N. Asamoto, S. Takagaki, "A Single Poly EEPROM Cell Structrure for Use in Standard CMOS Processes," IEEE Journal of Solid State Circuits, Vol. 29, No. 3, pp. 311-316, March 1994.
[6] ISO/IEC 18000-6, "Information technology - Radio frequency identification for item management - Part 6: Parameters for air interface communications at 860 MHz to $960 \mathrm{MHz} ", 2004$.
[7] A. Rantala, M. Sopanen, M. Aberg, "Implementation Experiments of Analog Nonvolatile Memory with a Standard $0.35 \mu \mathrm{~m}$ CMOS," Proc. NORCHIP 2004, Oslo, Norway, November 2004, pp. 71-74.
[8] A. Conte, G. Lo Giudice, G. Palumbo, A. Signorello, "A 1.35-V Sense Amplifier for Non Volatile Memories based on Current Mode Approach," Proc. ESSCIRC 2004, Leuven, Belgium, September 2004, pp. 471-474.

# 6. COMPLETE IMPLEMENTATION OF A PASSIVE TRANSPONDER 

### 6.1. Transponder architecture

We have implemented a read-only passive transponder that can operate in the ISM UHF and 2.45 GHz frequency ranges. For the demonstrator we have chosen a very simple architecture that minimizes power consumption and then maximizes the operating range.

In the communication from reader to transponder, an On Off Keying (OOK) modulation is used with a Pulse Width Modulation (PWM) coding. Thanks to the amplitude modulation, the received signal can be demodulated by using a simple rectifier with a proper time constant, since the frequency of the carrier and of the modulating signal differ by several orders of magnitude. Amplitude modulation of the signal transmitted by the reader allows the transponder to have on board just a simple demodulator instead of a coherent receiver, thereby drastically reducing power consumption. We choose a PWM coding with very different pulse duration between ' 0 ' and ' 1 ': in such a way a simple counter driven by a low accuracy clock is sufficient to demodulate the signal. With such coding scheme we do not need to implement a power-hungry Phase-Locked Loop (PLL) allowing us to keep the power consumption of the transponder in the $\mu \mathrm{W}$-regime.

In the communication from the transponder to the reader, an Amplitude Shift Keying (ASK) modulation is used with an FM0 coding and a data-rate of 40 kbps . The ASK modulation and the FM0 data coding have been used in compliance with the standard ISO 18000-6. A return link message consists of $n$ data bits preceded by a preamble. The preamble allows the reader to lock to the tag data clock, which has a low accuracy ( $10 \%$ ), and begin to decode the message. The preamble consists of 16 bits and contains multiple code violations (sequences of high and low levels not conforming to FM0 encoding) that act as a frame marker for the transition from preamble to data. The $n$ data bits consist of a sequence of bits corresponding to the ID code followed by the CRC field.

We implement the multiple access protocol ALOHA, which is the simplest possible and is efficient in the case of a small number of transponders to detect. It


Fig. 6-1: Architecture of a passive UHF/microwave transponder. FSM is the digital finite state machine
consists of transmitting the ID code at a random time instant until the ID code is correctly received by the reader. When this happens, the reader transmits an acknowledgment signal that turns off the detected transponder. Such protocol allows us to use a very simple digital section leading to very small power consumption.

The architecture of the passive transponder is shown in Fig. 6-1.

### 6.2. Analog section

The coupling element is typically a dipole or patch antenna. A voltage multiplier converts the input alternating voltage into a DC voltage which is used by a series voltage regulator to provide the regulated voltage required for the correct operation of the transponder. The voltage multiplier is matched with the antenna in order to ensure the maximum power transfer from the transponder's antenna to the input of the voltage multiplier. A backscatter modulator is used to modulate the impedance seen by the transponder's antenna during transmission. The RF section is then connected to the digital section, which typically is a very simple microprocessor or a finite state machine able to manage the communication protocol.

The input voltage $V_{\text {DDlow }}$, which has to provide almost all the power required for transponder operation, is generated by a single stage voltage multiplier, that is shown to provide the maximum power efficiency in the conversion from the RF


Fig. 6-2: Schematic of the single stage voltage multiplier.
energy at the transponder's antenna to the DC power supply at the output of the voltage regulator [1]. For the digital section, we consider an implementation based on subthreshold CMOS logic, with a standard $0.35 \mu \mathrm{~m}$ CMOS process, a regulated supply voltage $V_{R E G}=0.6 \mathrm{~V}$, and a clock frequency of 80 kHz , which enables sub- $\mu \mathrm{W}$ power consumption As stated in section 4 , a two-stage voltage multiplier is used: the first stage provides almost all the power required for the transponder operation and the second stage only provides the small power required for the error amplifier and the voltage reference generator. In such a way, power efficiency can be maximized. In the following sub-section, all the blocks of the analog section, shown in Fig. 6-1, are described in detail, except for the voltage regulator, whose operation has been described in detail in chapter 4.

### 6.2.1. Voltage multiplier

The voltage multiplier converts the input alternating voltage into a DC voltage. It must be able to operate at both UHF and 2.45 GHz . In accordance to the analysis done in chapter II, in order to maximize the power efficiency of the transponder, almost all the power required by the transponder is provided by the first stage; the second stage must only provide the few tens of nW required by the operational amplifier and the voltage reference generator, used in the voltage regulator.

Since we have used a standard CMOS process, the diodes in the voltage multiplier are implemented by using diode-connected MOS transistors. The DC voltage at the output of the first stage, $V_{\text {DDlow }}$, can be written as $V_{\text {DDlow }}=2\left(V_{0}-V_{t h}\right)$, where $V_{0}$ is the amplitude of the input alternating voltage and $V_{t h}$ is the threshold voltage of the MOS diodes. When the distance between reader and transponder corresponds to the maximum operating range, the DC output voltage of the first stage of the voltage multiplier is only slightly larger than 0.6 V and then rather close to the threshold voltage of MOS transistors. In this condition, power efficiency is drastically degraded. An approach to overcome such problem is to use
low-threshold-voltage transistors [2] or, in the case of a standard CMOS process, where low-threshold-voltage transistors are not available, other techniques to compensate the threshold voltage must be used to increase power efficiency [3], [4]. We have used a technique based on [3], to compensate the threshold voltage of MOS transistors used as diodes. For sake of simplicity, in the following a single stage voltage multiplier is described in detail and is shown in Fig. 6-2 . The 2-stage voltage multiplier is obtained from the single stage voltage multiplier by adding an identical second stage.

To improve the power efficiency, we need $i$ ) to minimize the parasitic capacitance at node A because it creates a leakage path from the input terminals and, $i i$ ) to compensate the threshold voltages of the MOS diodes $\mathrm{M}_{\mathrm{n} 1}$ and $\mathrm{M}_{\mathrm{p} 1}$, used in the voltage multiplier because they reduce the generated DC voltage $V_{\text {DDlow }}$.

The capacitors $C$ are poly1/poly2 capacitors and must be dimensioned so that their associated time constant is much smaller than the period of the input signal. If the previous condition is satisfied in the case of UHF transponder, it will be even more satisfied for a microwave transponder. The bottom plate of the capacitor C, placed at the input of the voltage multiplier, must be connected to the input alternating voltage $V_{I N}$ because the bottom plate parasitic capacitance is quite large, about one fifth of the capacitance C.. Furthermore, in order to reduce the parasitic capacitance at node A, a PMOS diode $\mathrm{M}_{\mathrm{p} 1}$ is used, instead of an NMOS diode, which has a smaller threshold voltage, so that the parasitic capacitance of the MOS diode $\mathrm{M}_{\mathrm{p} 2}$ used to compensate the threshold voltage of $\mathrm{M}_{\mathrm{p} 1}$ is connected in parallel with the output; if an NMOS diode was used the parasitic capacitance of the MOS diode used to compensate the threshold voltage would be connected to node A, degrading the power efficiency of the voltage multiplier.

In order to compensate the threshold voltage of the MOS diodes $M_{n 1}$ and $M_{p 1}$, the two diode-connected MOS transistors $\mathrm{M}_{\mathrm{n} 2}$ and $\mathrm{M}_{\mathrm{p} 2}$ are used. The two transistors are biased with a current $I^{*}$ obtained by mirroring the reference current generated in the voltage regulator. In such way, the transistors $\mathrm{M}_{\mathrm{n} 2}$ and $\mathrm{M}_{\mathrm{p} 2}$ are biased with a current independent of the DC voltage at the two outputs of the voltage multiplier. Moreover, such a current is smaller than 10 nA and then it is negligible in the calculation of the power efficiency. The two transistors $M_{n 2}$ and $M_{p 2}$ are dimensioned so that their gate-source voltage is very close to their threshold voltage. As a consequence the output DC voltage $V_{\text {DDlow }}$ has the expression shown below,

$$
\begin{equation*}
V_{\text {DDlow }} \cong 2 V_{0}-\left(V_{\text {thn } 1}-V_{G S n 2}\right)-\left(V_{t h p 1}-V_{G S p 2}\right) \cong 2 V_{0}, \tag{6-1}
\end{equation*}
$$

where $V_{0}$ is the amplitude of the input alternating voltage, $V_{t h n 1}$ and $V_{t h p 1}$ are the threshold voltages of $\mathrm{M}_{\mathrm{n} 1}$ and $\mathrm{M}_{\mathrm{p1} 1}, V_{G S n 2}$ and $V_{G S p 2}$ are the gate-source voltages of $\mathrm{M}_{\mathrm{n} 2}$ and $\mathrm{M}_{\mathrm{p} 2}$. In such a way the maximum DC output voltage is generated.


Fig. 6-3: Schematic of the ASK demodulator.

Simulations results show that a power efficiency of $39 \%$ is achieved, that is an input power of $2.56 \mu \mathrm{~W}$ is required to generate a DC output power of $1 \mu \mathrm{~W}$. Moreover, in order to generate the minimum output DC voltage $V_{D D l o w}$ required for the correct operation of the series voltage regulator $(630 \mathrm{mV})$, the amplitude of the input alternating voltage $V_{I N}$ must be as low as 350 mV . The compensation of the threshold voltage of MOS diodes used in the voltage multiplier allows us both to improve the power efficiency of the voltage multiplier and to relax the requirement on the quality factor of the power matching network.

The limiter consists of a series of diodes to limit the output DC voltage at the maximum value allowed by the dynamic range of the voltage regulator. A small resistance in series with the diodes is used to limit the current in the diodes when they go in conduction.

### 6.2.2. Demodulator

The schematic of the ASK demodulator is shown in Fig. 6-3. The demodulator is essentially a simple single stage voltage multiplier with a smaller time constant associated to the two capacitors $C$ with respect to the case seen above. The demodulator output must follow the variations of the PWM signal.
For the MOS diodes used in the demodulator, the same considerations done for the voltage multiplier are still valid. The two bias voltages $V_{\text {biasn }}$ and $V_{\text {biasp }}$ are generated as seen in the voltage multiplier. In such case the output resistance is implemented by exploiting the drain-source resistance of an NMOS transistor $M_{R}$, whose gate voltage is equal to $V_{\text {DDlow }}$. In such way the transistor $\mathrm{M}_{\mathrm{R}}$ operates as a voltage-


Fig. 6-4: Schematic of the ASK backscatter modulator.
controlled resistance. Indeed, when the power at the antenna increases, the DC voltages $V_{\text {DDlow }}$ and $V_{U}$ increase and then if the time constant of the output capacitance in the demodulator remained constant, the demodulator would not be able to follow the variations of the PWM signal anymore. By using a voltagecontrolled resistor, when $V_{\text {DDlow }}$ and $V_{U}$ increase, the resistance associated to $\mathrm{M}_{\mathrm{R}}$ decreases and then the time constant decreases as well. In such a way, the demodulator can follow the variations of the PWM signal for any power at the transponder's antenna. The output inverter, formed by $\mathrm{M}_{\mathrm{U} 1}$ and $\mathrm{M}_{\mathrm{U} 2}$, is supplied by the regulated voltage $V_{R E G}$, which is generated as shown in Fig. 6-2. Simulation results show that the power consumption of the demodulator, in the condition of minimum power at the input of the RF section, is about 250 nW . Such power must not be provided by the voltage regulator but directly by the antenna.

### 6.2.3. ASK backscatter modulator

The schematic of the ASK backscatter modulator is show in Fig. 6-4. The two NMOS transistors $\mathrm{M}_{\mathrm{M} 1}$ and $\mathrm{M}_{\mathrm{M} 2}$ are driven by two opposite signals. When the input voltage $V_{M O D}$ is high, $\mathrm{M}_{\mathrm{M} 1}$ is in cut-off region and $\mathrm{M}_{\mathrm{M} 2}$ conducts. In this case the output capacitance $C_{O U T}$ and the output resistance $R_{O U T}$ are given by,

$$
\begin{equation*}
C_{\text {OUT }} \cong C_{G D 2}+C_{D B 2}, R_{\text {OUT }} \cong r_{d 2}, \tag{6-2}
\end{equation*}
$$

where, $C_{G D 2}$ is the gate-drain capacitance of $\mathrm{M}_{2}, C_{D B 2}$ is the drain-bulk capacitance of $\mathrm{M}_{2}$ and $r_{d 2}$ is the drain-source resistance of $\mathrm{M}_{2}$. The value of the output resistance is chosen, in order to obtain a modulation depth in accordance with that derived in


Fig. 6-5: Block diagram of the digital section.
chapter 2 but in the case of an ASK modulation. In particularly, the value of the output resistance and capacitance are $650 \Omega$ and 128 fF , respectively. In this case the resistive part of the impedance seen by the antenna, $R_{l}$, is $R_{1}=R_{A} \| R_{\text {OUT }}=65 \Omega$, where $R_{A}$ is the antenna resistance, which is equal to $72 \Omega$ in the case of a dipole antenna.

When the input voltage $V_{M O D}$ is low, $\mathrm{M}_{\mathrm{M} 1}$ conducts and $\mathrm{M}_{\mathrm{M} 2}$ is in cut off region. In this case the output capacitance $C_{\text {OUT }}$ and the output resistance $R_{\text {OUT }}$ are given by,

$$
\begin{equation*}
C_{O U T} \cong C_{G D 1}+C_{D B 1}, R_{\text {OUT }} \cong r_{d 1} \tag{6-3}
\end{equation*}
$$

where, $C_{G D I}$ is the gate-drain capacitance of $\mathrm{M}_{1}, C_{D B I}$ is the drain-bulk capacitance of $\mathrm{M}_{1}$ and $r_{d l}$ is the drain-source resistance of $\mathrm{M}_{1}$. The two transistors have been dimensioned so that the output capacitances in the two cases are equal and the output resistance $r_{d l}$ is much larger than $r_{d 2}$. In this case, the value of the output resistance and capacitance are $13,2 \mathrm{k} \Omega$ and 128 fF , respectevily. In this case the resistive part of the impedance seen by the antenna, $R_{0}$, is $R_{0}=R_{A} \| R_{\text {OUT }} \cong 72 \Omega$.
As a consequence, the real part of the impedance seen by the antenna varies from $65 \Omega$ to $72 \Omega$ and the imaginary part of the impedance remains constant. The power consumption of the modulator is as low as 15 nW , since it works at very low frequency.


Fig. 6-6: Schematic of the random delay generator.

### 6.3. Digital section

The block diagram of the digital section is shown in Fig. 6-5. When the transponder is within the interrogation zone, the power at the transponder's antenna is sufficient to power on the transponder. In such condition, the 128 bit-ID code, contained in the ROM memory, is charged, in parallel, in the shift register and then, at a random time instant, generated by the random delay block, the ID code is serially transmitted after being encoded by the FM0 encoder block. The time required to transmit the 128 bit-ID code, with a data-rate of 40 kbps , is 3.2 ms . After each transmission of the ID code, the transponder waits for the acknowledgment sequence for 128 ms and, if the acknowledgment sequence is not detected by the ACK detector, the transponder retransmits the ID code at a random time instant; if the acknowledgment sequence is recognized by the ACK detector, the AND gate is disabled and then the transmission is turned off. The acknowledgment sequence is unique for all the transponders but, when the reader transmits the acknowledgment sequence, only the transponder which has transmitted the ID code received by the reader, is waiting for the acknowledgment sequence and then only such transponder can be turned off.


Fig. 6-7: Acknowledgement sequence format.

### 6.3.1. Random delay

The random delay circuit generates a positive edge at a random time instant. The schematic of the random delay circuit is shown in Fig. 6-6. It consists of a 4-bit-Pseudo-Random Binary Sequence (PRBS) generator that generates a binary number from 0 to 15 (from 0000 to 1111); such binary number is then converted to a delay by mean a counter supplied by a clock signal with a period of 3.2 ms , which is exactly equal to the time interval required to transmit the complete 128 bit-ID code with a data-rate of 40 kbps . As a consequence the random delay circuit can generate a random delay from 0 to 48 ms , after which the transmission of the ID code can start. When the transmission of the ID code is completed, the transponder waits 128 ms to receive the acknowledgment sequence and then, if the acknowledgment sequence has not been still received, a new random delay is generated.

### 6.3.2. Acknowledgement detector

The acknowledgement detector must decode the input PWM signal and compare the decoded digital signal with a given acknowledgment sequence, which is shown in Fig. 6-7. The acknowledgment sequence format consists of a sequence of bits 101, PWM coded, and separated to each other of $25 \mu \mathrm{~s}$. The bit ' 1 ' is coded with a pulse of duration $125 \mu \mathrm{~s}$ and the bit ' 0 ' is coded with a pulse of duration $25 \mu \mathrm{~s}$. When the negative edge is detected, a 3-bit-counter, supplied by a clock signal at 80 kHz , is turned on; when the positive edge is detected, the counter is turned off. If the output of the counter is smaller than 3 , which means that the pulse duration is smaller than $37.5 \mu \mathrm{~s}$, a bit ' 0 ' is detected; if the output of the counter is larger than 8 , which means that the pulse duration is larger than $100 \mu \mathrm{~s}$, a bit ' 1 ' is detected. Such thresholds, to decide if a ' 0 ' or a ' 1 ' has been detected, have been chosen, in order to take into account the accuracy of the clock signal, which can be worse than $10 \%$. If the sequence ' 101 ' is recognized, the output of the acknowledgement circuit goes low, interrupting the transmission of the ID-code.


Fig. 6-8: Block diagram of the proposed oscillator.

### 6.3.3. Clock generator

The clock generator is the most critical block in the digital section. The clock generator must have several requirements: low power, low voltage, low frequency, low process and temperature sensitivity, and full integrability. For such reason an RC-based oscillator has been implemented. RC oscillators are often used in micro controller, biomedical or other ASIC applications where the accuracy is not very important ( 1 to $10 \%$ ) and the frequency is quite low. Indeed, the accuracy of the oscillation frequency is affected by the accuracy of the resistor and capacitor used to define the frequency. Several implementations can be found in the literature [5], [6], [7], [8]. The RC oscillators are cheaper than crystal oscillators and do not require inductors. Since in digital CMOS process the tolerance on the value of resistors is larger than $30 \%$, most of the RC oscillators presented in the literature use an external resistor [6], [7], [8], which can have accuracy smaller than $1 \%$, allowing a frequency accuracy of few percents. In this paper, we present an oscillator without any external component and trimming, capable of 1 V supply voltage. Such oscillator is aimed to be used as a clock circuit in a passive microwave RFID transponder where an accuracy smaller than $15 \%$ is acceptable but a $\mu \mathrm{W}$ power consumption is required [9].

## Circuit Description

The block diagram of the proposed oscillator is shown in Fig. 6-8. It consists of a current generator circuit that provides a reference current $I_{r e f}$, used to charge and discharge a capacitor $C_{T}$, and two comparators that compare the voltage across such capacitor with two threshold voltages $V_{\min }$ and $V_{\max }$; the output voltages of the two comparators are then used, through an SR-flip flop, to drive the input switch. When the voltage across the capacitor becomes larger than $V_{\max }$, the output of the comparator 2 goes high, the flip flop is set, the switch goes to position 2 and the capacitor is discharged. When the voltage across the capacitor becomes smaller than $V_{\text {min }}$, the output of the comparator 1 goes high, the flip flop is reset, the switch goes to position 1 and the capacitor is charged. The oscillation frequency has the expression shown below,

$$
\begin{equation*}
f=\frac{I_{R E F}}{2 C_{T}\left(V_{\max }-V_{\min }\right)}, \tag{6-4}
\end{equation*}
$$

where all the parameters are indicated in Fig. 6-8. Both $V_{\min }$ and $V_{\max }$ are given by the gate-source voltage of a diode-connected MOS transistor biased by the reference current. The I-V characteristic of a MOS in the saturation region can be approximated by

$$
\begin{equation*}
I_{D}=\frac{\mu C_{o x}}{2} \frac{W}{L}\left(V_{G S}-V_{t h}\right)^{2}=\frac{k}{2}\left(V_{G S}-V_{t h}\right)^{2}, \tag{6-5}
\end{equation*}
$$

As a consequence, by using (6-5), the oscillation frequency can be written as:

$$
\begin{equation*}
f=\frac{\sqrt{I_{R E F}}}{2 C_{T} \sqrt{2 / k_{8}}(1-M)}, \tag{6-6}
\end{equation*}
$$

where $M=\sqrt{k_{8} / k_{7}}$. Since the mobility has a negative temperature coefficient, in order to achieve a low sensitivity to temperature variations, the reference current must have a positive temperature coefficient. The sensitivity of the oscillation frequency to process variations depends on the accuracy of the capacitor, on the variation of the mobility and of the reference current.
The most critical issue in the design of the proposed oscillator is the generation of the reference current, which must have a small sensitivity to process variations and a small value to keep the power consumption as small as possible. The parasitic


Fig. 6-9: Schematic of the current reference generator.
circuit elements and the intrinsic delays of the comparators and of the SR-flip flop are negligible because the oscillation frequency is quite low.

## Current Reference Generator

The schematic of the current reference is shown in Fig. 6-9. Transistors M2 and M3 operate in the subthreshold region. The I-V characteristic of a MOS in the subthreshold region can be well approximated by

$$
\begin{equation*}
I_{D}=\mu V_{T}^{2} \frac{W}{L} \exp \left(\frac{V_{G S}-V_{t h}}{m V_{T}}\right)\left[1-\exp \left(-\frac{V_{D S}}{V_{T}}\right)\right], \tag{6-7}
\end{equation*}
$$

The op-amp OA1 enforces the same voltage on the drain of M4 and M5 in order to have the same currents in the two branches of the current reference generator. The op-amp OA2 sets the drain voltage of M1 to the reference voltage $V_{\text {ref }}$. The generated current has the expression:


Fig. 6-10: Schematic of the voltage reference generator.

$$
\begin{equation*}
I_{r e f}=\frac{k_{3}}{2(1-N)^{2}} V_{r e f}^{2} \tag{6-8}
\end{equation*}
$$

where $N=\sqrt{k_{3} / k_{2}}$. In order to achieve a reference current with a positive temperature coefficient, as required for the temperature compensation of the oscillation frequency, the temperature coefficient of the reference voltage must be positive. Moreover, in order to minimize the sensitivity to process variations, the reference voltage must only depend on the ratios of parameters so that it is just affected by matching errors.

## Voltage Reference

The schematic of the voltage reference circuit is shown in Fig. 6-10. By assuming that the currents in the three branches of the circuit of Fig. 6-10 are identical, the voltage reference has the expression shown below,

$$
\begin{equation*}
V_{r e f}=\frac{R_{2}}{R_{1}} V_{T} \ln (n) \tag{6-9}
\end{equation*}
$$

where $n$ is the ratio of the emitter area Q 1 to the emitter area of Q 2 . As clear from (6-9), the reference voltage has a positive temperature coefficient and, when considering the process variations, it is only affected by matching errors.

## Temperature Coefficient

In a first approximation we can assume that the mobility has a temperature dependence given by [10],

$$
\mu=\mu_{0}\left(\frac{T}{T_{0}}\right)^{\mu_{T}}
$$

where $\mu_{0}$ is the mobility at the reference temperature $T_{0}, \mu_{T}$ is exponent mobility coefficient. By differentiating (6-6) and taking into account (6-8) and (6-9), we can derive the temperature coefficient of the oscillation frequency:

$$
\begin{equation*}
\frac{\partial f}{\partial T} \frac{1}{f}=\frac{\partial V_{\text {ref }}}{\partial T} \frac{1}{V_{\text {ref }}}+\frac{1}{2}\left(\frac{\partial \mu_{3}}{\partial T} \frac{1}{\mu_{3}}+\frac{\partial \mu_{8}}{\partial T} \frac{1}{\mu_{8}}\right)=\frac{\partial V_{\text {ref }}}{\partial T} \frac{1}{V_{\text {ref }}}+\frac{\mu_{T}}{T} . \tag{6-11}
\end{equation*}
$$

By differentiating (6-9), we obtain that

$$
\begin{equation*}
\frac{\partial V_{r e f}}{\partial T} \frac{1}{V_{r e f}}=\frac{1}{T} \tag{6-12}
\end{equation*}
$$

As a consequence, the relative temperature coefficient of the oscillation frequency is,

$$
\begin{equation*}
\frac{\partial f}{\partial T} \frac{1}{f}=\frac{1}{T}\left(\mu_{T}+1\right) \tag{6-13}
\end{equation*}
$$

Since in our IC technology, the exponent mobility coefficient is -1.3 , a theoretical temperature coefficient of about $1000 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ can be achieved, at room temperature.

## Process Variations Sensitivity

From (6-6), (6-8) and (6-9) we can derive the following expression of the oscillation frequency,

$$
\begin{equation*}
f=\frac{\sqrt{k_{3} k_{8}}}{2 C_{T} \sqrt{2}(1-M)(1-N)} V_{\text {ref }} . \tag{6-14}
\end{equation*}
$$

If the matching errors are neglected, $V_{r e f}, M$ and $N$ in (6-14) can be considered process independent, since $M$ and $N$ are $W / L$ ratios and $V_{\text {ref }}$ is generated from the supply voltage through a partition coefficient $\alpha\left(V_{r e f}=\alpha V_{d d}\right)$, which is only affected by matching errors. Therefore, the sensitivity of the oscillation frequency to process variations is due to the carrier mobility $\mu$ and to the accuracy of the capacitor $C_{T}$. To first order, from (6-14), we have:

$$
\frac{d f}{f}=\frac{d \mu}{\mu}-\frac{d C_{T}}{C_{T}},
$$

In a standard RC-oscillator, the standard deviation of the oscillation frequency depends on the accuracy of the resistor and the capacitor that determine the oscillation frequency. In the proposed oscillator, instead, as is clear from (6-15), the standard deviation of the oscillation frequency depends on the accuracy of a capacitor and of the mobility, which has a standard deviation much smaller than that of a resistor. Such solution allows us to implement the proposed oscillator in a fully integrated way, without any external component, achieving, at the same time, an accuracy good enough for several applications.

## Simulation Results

The proposed current reference circuit has been implemented in AMS $0.35 \mu \mathrm{~m}$ CMOS. Simulations show that the proposed oscillator generates a clock signal of about 80 kHz . The proposed circuit operates with a supply voltage of 1 V and consumes a maximum power consumption, at $80^{\circ} \mathrm{C}$, of about $1.12 \mu \mathrm{~W}$. The power consumption decreases to $1.04 \mu \mathrm{~W}$, at $0^{\circ} \mathrm{C}$. Fig. 6-11 shows the oscillation frequency as function of the temperature. From simulation results we can derive that the proposed oscillator has a temperature coefficient of about $842 \mathrm{ppm} /{ }^{\circ} \mathrm{C}$ over a temperature range from 0 to $80^{\circ} \mathrm{C}$, which is in good agreement with the theoretical results provided by (6-13). When the supply voltage varies from 1 to 1.5 V , the oscillation frequency varies from 80 kHz to 79 kHz , leading to a line sensitivity of $-2.5 \% / \mathrm{V}$.


Fig. 6-11: Oscillation frequency of the clock generator vs. temperature.


Fig. 6-12: Histogram of the oscillation frequency for 50 runs of the Montecarlo simulation.


Fig. 6-13: a) Power consumption vs. temperature, b) Histogram of the power consumption for 50 runs of the Montecarlo simulation.

In order to evaluate the sensitivity of the oscillation frequency to process variations, a Monte Carlo simulation has been performed and the results are shown in the histogram of Fig. 6-12. Simulation results show a mean value of the oscillation frequency of 80.85 kHz and a standard deviation of 3.19 kHz , leading to a $3 \sigma$ of about $11.85 \%$, which is good enough for many applications, such as the clock generator of passive microwave transponders, where an accuracy of $15 \%$ is typically acceptable, or in biomedical applications. Fig. 6-13a shows the variation of the power consumption as function of temperature. A Monte Carlo simulation has also been performed in order to evaluate the dispersion of power consumption due to the process variations. As is clear from Fig. 6-13a-b, the power consumption is always smaller than $1.14 \mu \mathrm{~W}$. The area occupation on the chip is $0.24 \mathrm{~mm}^{2}$. The performance of the proposed oscillator is compared with those of other designs already reported in the literature. The comparison is shown in Table 6-I. From Table 6-I we can note that the proposed oscillator has the lowest power-to-frequency ratio $(0.014 \mu \mathrm{~W} / \mathrm{kHz})$ and the lowest temperature coefficient. The sensitivity of the circuit to process variations is acceptable for several applications but larger than

Table 6-I: Comparison with other results reported in the literature.

|  | This work | Hwang <br> $[1]$ | Lasanen <br> $[2]$ | Bala [4] | Kakela <br> $[7]$ |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Technology | $0.35 \mu \mathrm{~m}$ <br> CMOS | $2 \mu \mathrm{~m}$ <br> CMOS | $0.35 \mu \mathrm{~m}$ <br> CMOS | $0.18 \mu \mathrm{~m}$ <br> CMOS | $3 \mu \mathrm{~m}$ <br> CMOS |
| Min Supply <br> Voltage (V) | 1 | 2 | 1 | 1.25 | 2.5 |
| Frequency | 80 kHz | $0.3-100$ <br> Hz | $100 \mathrm{kHz}-$ <br> 7 MHz | $6-24$ <br> MHz | 34.6 kHz |
| Max. power <br> consumption | $1.14 \mu \mathrm{~W}$ | $0.3 \mu \mathrm{~W}$ | $52 \mu \mathrm{~W}$ | 1.12 mW | $5.9 \mu \mathrm{~W}$ |
| Line <br> sensitivity | $-2.5 \% / \mathrm{V}$ | $\mathrm{N} / \mathrm{A}$ | $1.9 \% / \mathrm{V}$ | $\mathrm{N} / \mathrm{A}$ | $-2.3 \% / \mathrm{V}$ |
| TC | 842 <br> $\mathrm{ppm} /{ }^{\circ} \mathrm{C}$ | $\mathrm{N} / \mathrm{A}$ | $1.4 \% / \mathrm{V}$ | $\mathrm{N} / \mathrm{A}$ | $-3 \% / \mathrm{V}$ |
| Relative <br> Frequency | $11.85 \%$ | N/A | $5 \%$ | $4 \%$ | $13 \%$ |
| Ext R, C | No | No | Yes | Yes | Yes |
| Area $\left(\mathrm{mm}^{2}\right.$ ) | 0.24 | 0.281 | 0.09 | 0.14 | 0.1 |

that of other circuits considered in Table 6-I, all of which however have the drawback of requiring additional external components, which in our case is not required.

### 6.4. Conclusion

A low-frequency oscillator implemented in AMS $0.35 \mu \mathrm{~m}$ CMOS has been presented, which does not require any external component and can therefore be fully integrated. The proposed circuit is suitable for low power and low voltage applications in virtue of the supply voltage of 1 V and a power consumption of about $1.1 \mu \mathrm{~W}$. Power-to-frequency ratio and temperature sensitivity are significantly smaller than those obtained by comparable solutions presented in the literature. The proposed oscillator can be used in applications where a dispersion $(3 \sigma)$ of frequency due to process variations slightly larger than $10 \%$ can be accepted, such as passive RFID transponders or biomedical applications.

The proposed clock generator has been optimized in a successive implementation, achieving a power consumption as low as 708 nW with a standard
deviation due to process variations smaller than $15 \%$, as required for RFID transponders.

### 6.5. System performance

The power consumption of the digital section is strongly dominated by the power consumption of the clock generator. Since the digital section is quite simple, the static power consumption due to the leakage currents is quite small. Also the dynamic power consumption is negligible, since the clock frequency is very small. The power consumption of the digital section, excluding the clock generator is as small as 130 nW . As a consequence the total power consumption of the transponder is about 838 nW . The total power, $P_{R A D}$, that must be provided by the voltage multiplier is given by the power dissipated in the digital section, by the power dissipates by the voltage regulator and by the power dissipates by the ASK backscatter modulator. Such power is equal to about 900 nW .

The minimum power $P_{A N T}$ at the transponder's antenna, required for the correct operation of the transponder, is given by,

$$
\begin{equation*}
P_{A N T}=P_{D E M O D}+\frac{P_{R A D}}{\eta}, \tag{6-16}
\end{equation*}
$$

By assuming a power efficiency of the voltage multiplier of $39 \%$, the minimum power at the transponder's antenna is about $2.55 \mu \mathrm{~W}$. By assuming to use a dipole antenna, with such power consumption a theoretical operating range of 15.5 m and 41.7 m at UHF in Europe and US, respectively and of 5.5 m and 15.6 m at 2.45 GHz in Europe and US, respectively.

Such performance is much better than that of commercial prototypes listed in section 1.3.4 and are better than that of research prototypes listed in section 1.3.4.

### 6.6. References

[1] G. De Vita, G. Iannaccone, "Design Criteria for the RF section of UHF and Microwave Passive RFID Transponders," IEEE Transactions on Microwave Theory and Techniques, Vol. 53, No 9, pp 2978-2990, September 2005.
[2] J.-P. Curty, N. Joehl, C. Dehollain, M.J. Declercq, "Remotely Powered Addressable UHF RFID Integrated System," IEEE Journal of Solid State Circuits, Vol. 40, No. 11, pp. 2193-2202, November 2005.
[3] H. Nakamoto et al., "A Passive UHF RFID Tag LSI with 36.6\% Efficiency CMOS-Only Rectifier and Current-Mode Demodulator in $0.35 \mu \mathrm{~m}$ FeRAM Technology," ISSCC Dig. Tech. Papers, San Francisco, USA, pp. 310-311, February 2006.
[4] T. Umeda et al., "A 950 MHz Rectifier Circuit for Sensor Networks with 10mDistance," ISSCC Dig. Tech. Papers, San Francisco, USA, pp. 256-257, February 2005.
[5] C. Hwang, S. Bibyk, M. Ismail, B. Lohiser, "A Very Low Frequency, Micropower, Low Voltage CMOS Oscillator for Noncardiac Pacemakers," IEEE TCAS-I, Vol. 42, No. 11, pp. 962-966, 1995.
[6] K. Lasanen, E.R.-Ruotsalainen, J. Kostamovaara, "A 1-V, Self Adjusting, 5MHz CMOS RC-Oscillator," Proc. ISCAS 2002, Scottsdale, USA, Vol. IV, pp. 377-380, May 2002.
[7] T. O'Shaughnessy, "A CMOS, Self Calibrating, 100 MHz RC-Oscillator for ASIC Applications, " Proc. ASIC Conference 1995, Austin, USA, pp. 279-282, September 1995.
[8] F. Bala, T. Nandy, "Programmable High Frequency RC Oscillator," Proc. $18^{\text {th }}$ Int. Conf. on VLSI Design, Kolkata, India, pp. 511-515, January 2005.
[9] G. De Vita, G. Iannaccone, "Design Criteria for the RF section of UHF and Microwave Passive RFID Transponders," IEEE Transactions on Microwave Theory and Techniques, Vol. 53, No 9, pp. 2978-2990, September 2005.
[10] BSIM3v3 User's Manual. Available at:http://wwwdevice.eecs.berkeley.edu/~bsim3/.
[11]P. Kakela, T. Rahkonen, J. Kostamovaara, "A micropower RC oscillator for consumer ASIC applications," Proc. Electrotechnical Conf., Ljubljana, Slovenia, pp. 278-281, May 1991.

## 7. CONCLUSION

At first, we have presented an analysis and a discussion of the design options and tradeoffs for a passive microwave transponder. We have derived a set of criteria for the optimization of the voltage multiplier, the power matching network and the backscatter modulator in order to optimize the operating range, once fixed the datarate or to optimize the data-rate, once chosen the operative range. We have also shown that for RFID transponders requiring with a DC power of $1 \mu \mathrm{~W}$ for the digital section may reach an operating range of about 3.4 meters in the ISM 2.45 GHz band, and 9.5 meters in the ISM 868 MHz band, for a data-rate of several kbps and according to EU regulations, with a standard $0.35 \mu \mathrm{~m}$ process. Present US regulations would allow us to obtain an almost tripled operating range.

In order to match the strictly power requirements, the communication protocol between transponder and reader has been chosen in a convenient way, in order to make the architecture of the passive transponder very simple and then ultra-lowpower. In the communication from transponder to reader, the backscatter modulation, used to transmit data to the reader, allows us not to consume static power for the modulation. Moreover, the most simple anti-collision protocol, the ALOHA protocol, has been used, in order to make the digital section very simple and then to drastically reduce the power consumption. It is clear that the ALOHA protocol is efficient only for a few tens of transponder to be detected by the reader; more efficient anti-collision protocol can be used to quickly detect a large number of transponder but at the cost of larger power consumption and then smaller operating range. In the communication from the reader to the transponder, an adhoc coding scheme has been used to overcome synchronization and demodulation problems and then avoiding power-hungry circuits, such as the PLL. Such solution drastically limits the possibility of communication from the reader to the transponder to just a single command for turning off the transponder when it was detected. Some standards exist, i.e. EPC and ISO 18000, which have a large set of commands, but they require a complex and power hungry digital section to be implemented. As a consequence, all the choices done in the architecture of the transponder are lead by the necessity to achieve very low power consumption.

From the circuital point of view, the digital section has been implemented in subthreshold CMOS logic with very low supply voltage and clock frequency and each block of the analog section has been implemented to optimize the power consumption.

We have presented different solutions to supply power to the transponder, in order to keep the power consumption in the deep sub- $\mu \mathrm{W}$ regime and to drastically reduce the huge sensitivity of the subthreshold logic to temperature and process variations. Voltage reference generators based on the I-V characteristics of MOS
transistors have been used, rather than standard bandgaps, thereby reducing power consumption by more than one order of magnitude. Furthermore, ad-hoc series voltage regulators have been implemented to mitigate the sensitivity of subthreshold digital circuits to temperature and process variations, with a particular attention to the power consumption. Moreover, a low-voltage and low-power EEPROM in a standard CMOS process has been implemented. The memory exploits the Fowler-Nordheim tunnelling current to charge and discharge the floating gate avoiding hot carrier injection, which would lead to unacceptable power consumption, at the cost of a large access time, which is compatible with very low data-rate of RFID systems. Such memory has a power consumption of a few hundreds of nWs for a read operation and a few $\mu \mathrm{Ws}$ for a write operation. Such performance allows us to keep the power consumption during a read operation below one $\mu \mathrm{W}$ and to achieve an acceptable reduction of the operating range during a write operation.

Finally, we have presented the implementation of the entire passive transponder, operating in the UHF or microwave frequency range. Simulation results show that the total power consumption of the transponder is about 900 nW and the minimum power required at the transponder's antenna is as low as $2.5 \mu \mathrm{~W}$. Such expected power allows us to achieve operating ranges of 15.5 m and 41.7 m at UHF in Europe and US, respectively and of 5.5 m and 15.6 m at 2.45 GHz in Europe and US, respectively. Such operating ranges are much larger than that of others prototypes already reported in the literature or available in the market.

Such performances in terms of operating range, data-rates and cost, might open promising perspectives for the deployment of passive RFID systems even in Ambient Intelligence or Ubiquitous Computing scenarios, and in outdoor applications.

